State-of-the-Art Survey

LNCS 4568

Toru Ishida
Susan R. Fussell
Piek T. J. M. Vossen (Eds.)

Intercultural Collaboration

First International Workshop, IWIC 2007
Kyoto, Japan, January 2007
Invited and Selected Papers

Springer

Lecture Notes in Computer Science 4568

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Toru Ishida   Susan R. Fussell
Piek T. J. M. Vossen (Eds.)

Intercultural Collaboration

First International Workshop, IWIC 2007
Kyoto, Japan, January 25-26, 2007
Invited and Selected Papers

Springer

Volume Editors

Toru Ishida
Kyoto University
Department of Social Informatics
Yoshida-Honmachi, Kyoto 606-8501, Japan
E-mail: ishida@i.kyoto-u.ac.jp

Susan R. Fussell
Carnegie Mellon University
Human - Computer Interaction Institute, 3531 NSH,
5000 Forbes Avenue, Pittsburgh, PA 15213, USA
E-mail: sfussell@andrew.cmu.edu

Piek T. J. M. Vossen
Vrije Universiteit (VU) Amsterdam
De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands
E-mail: piek.vossen@irion.nl

Library of Congress Control Number: 2007931634

CR Subject Classification (1998): H.5.2, H.5.3, H.3-5, C.2, K.4.2, J.4, I.3

LNCS Sublibrary: SL 3 – Information Systems and Application, incl. Internet/Web and HCI

ISSN 0302-9743
ISBN-10 3-540-73999-8 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-73999-9 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media

springer.com

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 12101474 06/3180 5 4 3 2 1 0

Preface

The 1st International Workshop on Intercultural Collaboration, IWIC 2007, was held January 25–26, 2007, in Kyoto, Japan. The main theme of this workshop was intercultural collaboration, from both technical and sociocultural perspectives. To increase mutual understanding in our multicultural world, topics included collaboration support (such as natural language processing, Web, and Internet technologies), social psychological analyses of intercultural interaction, and case studies from field workers. This was a unique workshop in a world in which physical borders disappear rapidly and people and cultures are more and more on the move and in contact.

The workshop featured three prominent invited speakers: Christiane D. Fellbaum from the Department of Psychology, Princeton University, introduced the Global Wordnet activities, Yumiko Mori from NPO Pangaea presented communication components bridging children around the world, and Gary Olson from the School of Information, University of Michigan, explained the Science of Collaboratories Project. The workshop included 2 panels, 23 oral presentations, 19 posters, 6 demos, and excursions to world heritage sites in Kyoto. Participation in the workshop was by invitation only, but there were 103 professionals from 23 regions/countries: Australia, Canada, Chile, China, Cyprus, Denmark, Finland, France, Germany, India, Israel, Italy, Japan, Korea, Malaysia, The Netherlands, Romania, Russia, Sweden, Taiwan, Thailand, Tunisia and the USA.

To organize the workshop, three Program Co-chairs from different research fields worked together for the first time. Their expertise covered computer-supported collaborative work (CSCW), language technology and social informatics. The workshop itself became a case study in intercultural collaboration. To provide a different perspective of the workshop, each Program Co-chair noted a few paragraphs independently as follows.

The rise of the Internet and associated technologies for CSCW has made possible new collaborations among people from different national, cultural, and linguistic backgrounds. While these intercultural collaborations have the potential to bring great benefit to fields like business, science, and education, they are presently fraught with problems arising from linguistic and cultural differences. In order to address these problems and thereby more adequately support intercultural CSCW, we need a better understanding of how culture impacts communication and collaboration and we need new computer-mediated communication tools that enhance intercultural communication. The proceedings from this 1st International Workshop on Intercultural Collaboration provide insights on both of these fronts.

From the point of view of intercultural CSCW, the conference addressed many exciting topics, including the limitations of English as a common language, the potential of machine translation and related tools to better support collaboration between people with different native languages, the suitability of avatars and pictograms as alternatives to text-based communication, and the ways in which the broad range of cultures impact intercultural collaboration. The technologies described in this volume foreshadow an exciting new world in which anybody speaking any language can interact with anybody else. Taken together, the papers in this volume increase our understanding of how linguistic differences and cultural/national differences impact intercultural collaboration and help identify the key issues researchers will want to address in future research. (Susan R. Fussell)

Being an enabling technology, language technology is usually evaluated independently from end-user applications that incorporate them. What made this workshop extremely interesting was that it dealt with solving the real and complex problem of effective communication across different cultures and not just the perfect technology. A cross-cultural perspective does not just require a good translation, it requires successful communication. The tension between technology and effect in that case is even bigger. A failed language technology may not have any effect on the communication or it may have a fatal effect. Technically correct language technology may have the same effects.

From a language technology perspective, we are interested in building a universal index of meaning across languages and cultures, while maintaining cultural and linguistic differences in each language resource. We do not know how much common ground there is for interpretation and therefore we do not know what the differences are or how we can handle mismatches across cultures. Such an index will be a good starting point for exploring this and could become the basis for any cross-cultural communication system of the future. It will be a long and painful process that may take many decades to build, but we have to start somewhere. This workshop brought together people who help build such systems and people who deal with cultural mismatches on a daily basis; it helped to realize the full complexity of the problem. (Piek T.J.M. Vossen)

Intercultural collaboration is a goal-directed group activity based on intercultural communication. We set our research target on collaboration, not on communication, because we can clearly identify research issues when the goal of human group activities is defined. Goal-directed group activities can also be a target of optimization, and thus attract researchers who work on social and computer sciences. Obviously, the research area of intercultural collaboration encompasses CSCW and language technologies. To bridge the gap between these two research areas and fields of actual intercultural activities, however, we think the methodology of social informatics can be well used.

In the field of intercultural collaboration, various ideas have been tried out to overcome language barriers and to understand cultural diversity. For example, conversations between Chinese and Japanese are often done in English with the aid of written Chinese characters. To understand how humans manage language and cultural issues, researchers should join and observe intercultural activities. However, sites of intercultural activities are often far from research laboratories (both geographically and mentally). From the social informatics point of view, mutual learning between researchers and field workers is essential in developing new methodologies and technologies. In this workshop, we thus invited case studies from field workers as well as technical and scientific research papers to learn from experience and to share common issues in our community. (Toru Ishida)

This volume includes 29 invited and selected papers presented at the workshop. The papers are categorized into the following areas:

– Ontology for Language Resources and Services
– Developing Language Resources and Services
– Connecting Language Resources and Services
– Tools for Intercultural Collaboration
– Emotions in Intercultural Collaboration
– Analysis of Intercultural Collaboration
– Actual Fields of Intercultural Collaboration

During the workshop, we held discussions on how to create a new field under collaboration among CSCW, language technologies and social informatics. Our discussion will be continued at the next workshop, which will be held in 2008.

May 2007

Toru Ishida
Susan R. Fussell
Piek T.J.M. Vossen

Organization

IWIC 2007 was organized by the Language Grid Project, the National Institute of Information (NICT), Center of Excellence on Knowledge Society, Kyoto University, and the IEICE Special Interest Group for Intercultural Collaboration, and was supported by UNESCO.

Organizing Committee

General Co-chairs	Yasuhiro Katagiri (Future University-Hakodate, Japan), Hitoshi Isahara (NICT, Japan)
Program Co-chairs	Toru Ishida (Kyoto University, Japan) Susan R. Fussell (Carnegie Mellon University, USA) Piek T.J.M. Vossen (Vrije Universiteit Amsterdam, The Netherlands)
Local Organization Co-chairs	Rieko Inaba (NICT, Japan) Takashi Yoshino (Wakayama University, Japan)
Publicity Chair	Ahlem Ben Hassine (NICT, Japan)
Publication Chair	Yohei Murakami (NICT, Japan)
Finance Chair	Tomohiro Shigenobu (NICT, Japan)

Program Committee

Pushpak Bhattacharyya (Indian Institute of Technology (IIT), India)
Francis Bond (NICT, Japan)
Nicoletta Calzolari (Istituto Linguistica Computazionale, Italy)
Key-Sun Choi (Korea Advanced Institute of Science and Technology (KAIST), Korea)
Catherine Cramton (George Mason University, USA)
Christiane D. Fellbaum (Princeton University, USA)
Yoshihiko Hayashi (Osaka University, Japan)
Susan C. Herring (Indiana University, USA)
Reiko Hishiyama (Kyoto Women's University, Japan)
Chu-Ren Huang (The Institute of Linguistics, Academia Sinica, Taiwan)
Nancy Ide (Vassar College, USA)
Katherine Isbister (Rensselaer Polytechnic Institute, USA)
Chigusa Kita (Kansai University, Japan)
Manolis Koubarakis (Technical University of Crete, Greece)
Kyoung Jun Lee (Kyung Hee University, Korea)
Fang Li (Shanghai Jiaotong University, China)

Shigeo Matsubara (Kyoto University, Japan)
Michihiko Minoh (Kyoto University, Japan)
Akiyo Nadamoto (NICT, Japan)
Hideyuki Nakanishi (Osaka University, Japan)
Gary Olson (University of Michigan, USA)
Judy Olson (University of Michigan, USA)
Yuanchun Shi (Tsinghua University, China)
Virach Sornlertlamvanich (NICT, Thailand)
Dan Tufiş (The Romanian Academy/University "A.I.Cuza" of Iasi, Romania)
Kiyotaka Uchimoto (NICT, Japan)
Zili Zhang (Southwest University, China/Deakin University, Australia)
Qiping Zhang (Drexel University, USA)
Michael Zock (CNRS, Université de la Méditerranée, Université de Provence, France)

Local Organization Committee

Heeryon Cho (Kyoto University, Japan)
Donghui Lin (Kyoto University, Japan)
Naoki Miyata (Kyoto University, Japan)
Hiroki Morikawa (Kyoto University, Japan)
Masaki Goto (Kyoto University, Japan)
Rie Tanaka (Kyoto University, Japan)
Hironobu Miyaura (Kyoto University, Japan)
Kunikazu Fujii(Wakayama University, Japan)
Koji Fukui(Wakayama University, Japan)

Workshop Secretariat

Eri Tsunokawa (NTT-AT, Japan)

Sponsors

The Telecommunications Advancement Foundation
International Information Science Foundation
Tateisi Science and Technology Foundation
Inoue Foundation for Science

Table of Contents

Ontology for Language Resources and Services

Developing Language Resources and Services

Connecting Language Resources and Services

Tools for Intercultural Collaboration

Emotions in Intercultural Collaboration

Analysis of Intercultural Collaboration

Actual Fields of Intercultural Collaboration

Connecting the Universal to the Specific: Towards the Global Grid

Christiane Fellbaum[1,⋆] and Piek Vossen[2]

[1] Cognitive Science Laboratory, Princeton University
fellbaum@princeton.edu
[2] Irion Technologies and Faculty of Arts, Free University of Amsterdam
piek.vossen@irion.nl

Abstract. Despite their widespread use in Natural Language Processing applications, lexical databases and wordnets in particular do not yet contribute satisfactorily to the difficult problem of automatic word sense discrimination. Having built a number of lexical databases ourselves, we are keenly aware of still unresolved fundamental theoretical issues. In this paper we examine some of these questions and suggests preliminary answers concerning the nature of lexical elements and the conceptual-semantic and lexical relations that interconnect them. Our perspective is multilingual, and our goal is to formulate a proposal for a "Global Wordnet Grid" that will meet the challenge of mapping the lexicons of many languages in interesting and useful ways.

1 Introduction

Despite their widespread use in Natural Language Processing applications, lexical databases and wordnets in particular do not yet contribute satisfactorily to the difficult problem of automatic word sense discrimination. Having built a number of lexical databases ourselves, we are keenly aware of still unresolved fundamental theoretical issues. In this paper we examine some of these questions and suggests preliminary answers concerning the nature of lexical elements and the conceptual-semantic and lexical relations that interconnect them. Our perspective is multilingual, and our goal is to formulate a proposal for a "Global Wordnet Grid" that will meet the challenge of mapping the lexicons of many languages in interesting and useful ways.

The idea for a Global Wordnet Grid was born during the Third Global WordNet Conference in Korea (January 2006). This grid will be built around a set of concepts encoded as wordnet synsets in as many languages as possible and mapped to definitions in the SUMO ontology. We envision speakers from many diverse language communities creating and contributing synsets in their language. We initially solicit encodings for the nearly 5,000 Common Base Concepts used in many current wordnet projects.[1]

⋆ Work supported by the National Science Foundation and the Office of Disruptive Technology.

[1] Base Concepts are expressed by synsets that occupy central positions in the wordnet structures. They tend to express general concepts relatively high up in the hierarchies

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 1–16, 2007.

We anticipate cases of many-to-many mappings, where a given language will have more than one concept that covers the semantic space of a single Base Concept and vice versa. In other cases, a lexical item encoding a concept that is specific to a linguistic and cultural community will be included in the conceptual inventory shared by all languages, though there will be no corresponding lexemes in other languages.

Eventually, the Grid will represent the core lexicons of many languages in a form that allows further study of lexical and semantic similarities as well as disparities. Both research and applications will benefit from the Grid.[2]

2 Background: WordNet, EuroWordNet, Global WordNet

The Global Grid is a natural extension of the wordnets that have been built over the past decade. At the same time, we need to examine some fundamental assumptions that have guided past wordnets in the light of what we have learned. We begin with a brief review of the major wordnets.

2.1 WordNet

The Princeton WordNet is the first manually constructed large-scale lexical database that was widely embraced by the NLP community. WordNet was originally intended to test the feasibility of a model of human semantic memory that sought to explain economic principles of storage and retrieval of words and concepts. This model was based on the hierarchical organization of concepts expressed by nouns and the inheritance of properties (expressed by adjectives) and events (encoded by verbs) associated with these concepts. WordNet consists of four different semantic networks (one four each of the major parts of speech) that interrelated groups of cognitively synonymous words ("synsets") via lexical and conceptual-semantic relations. For details see [14], [7], [6].

The Princeton WordNet was designed and constructed with the goal of exploring the English lexicon, without a crosslinguistic perspective. Although it was not motivated by NLP needs, the WordNet model turned out to be useful for language processing. Consequently, wordnets started to be built in other languages.

2.2 EuroWordNet

Vossen was the first expansion of WordNet into other languages [23]. Along with the construction of lexical databases for (initially) four European languages,

and to be related to many other concepts. A comparisons of different wordnets led to a selection of English WordNet synsets that represent these concepts across a number of European languages. This selection is referred to as the Common Base Concepts [23].

[2] The Grid will be publicly and freely available and no proprietary claims are made by the contributors.

the EuroWordNet design deviated from that of the Princeton WordNet and contributed several fundamental innovations that have since been adopted by dozens of additional wordnets.

To increase the connectivity among synsets, a number of new relations were defined, in particular cross-part-of-speech relations. Furthermore, all relations were marked with a feature value indicating the combinations of relations (conjunctive or disjunctive) and their directionality.

The most important difference however was the multilingual nature of the database. Each wordnet was modeled after the Princeton WordNet, having its own spearate inventory of synsets and relations. In addition, the synsets of each language are linked via an "equivalence relation" to the InterLingual Index, or ILI. By means of the ILI, a synset in a given language can be mapped to a synset in any other language connected to the ILI. This design allowed the straightforward comparison of the lexicons of different languages both in terms of coverage, relations, and lexicalization patterns.

Initially, the EuroWordNet ILI was populated with the concepts (synsets) from Princeton WordNet. The reasons were mostly pragmatic — WordNet had a large coverage and was freely available. Furthermore, English was the language that was most familiar to all the European partners so judging equivalence was possible.

But several modifications and extensions of the ILI had to be considered. As WordNet was not designed as an ILI, it was often difficult to establish proper equivalence relations from the different languages. This was true even for languages that are closely related to English (like Dutch and German), and despite the fact that most European lexicons are marked by contemporary Anglo-American culture.

Compatibility between the EuroWordNet languages and the ILI with respect to lexical coverage and relations varied moreover depending on which of the two basic methods for building the European wordnets was followed:

– *Expand*: English synsets are translated into the target language and the relations are copied
– *Merge*: synsets are created for the target language, interlinked with the Princeton WordNet relations, and subsequently translated into English for mapping with ILI entries

The Expand Approach results in wordnets that are very close to the Princeton original, while the Merge Approach creates wordnets that often have a very different structure where synsets do not match straightforwardly.

2.3 Global WordNet

EWN was the first step towards the globalization of wordnets. Linguists and computer scientists in many countries started to develop WNs for their languages. Besides individual eforts, there are wordnets for entire geographic regions, such as BalkaNet [22] and the Indian WordNets (e.g., [21]). Currently, WNs exist for some 40 languages, including dead languages like Latin and Sanskrit. For information see www.globalwordnet.org.

The authors founded the Global WordNet Association (GWA), motivated by the desire to establish and maintain community consensus concerning a common framework for the structure and design of wordnets. Another goal is to encourage the development of wordnets for all languages and to link them such that appropriate concepts are mapped across languages. The multilingual wordnets allow one to compare the lexicons of different languages on a large scale, beyond the selected few lexemes that are often considered in the investigation of particular linguistic topics. Furthermore, the availability of global wordnets opens up exciting possibilities for crosslinguistic NLP applications.

3 Challenges

The goal of mapping the lexicons of genetically and typologically unrelated languages raises the question whether there exists a universal lexicon, an inventory of concepts that are lexically encoded (or potentially encodable) by all languages. Second, what kinds of concepts does such a universal lexicon cover and how large is the common core of lexicalized concepts for most or all languages? How do language-specific lexicalizations radiate out from the core? Conversely, we ask what the differences among the lexicons of diverse languages are, whether such differences are regular and systematic, and in which areas of the lexicon they are concentrated. For the cases where individual languages show lexical gaps, we ask whether these are attributable to grammatical and structural properties or to cultural differences.

These questions inevitably lead to another, more fundamental one. What constitutes a lexeme deserving of a legitimate entry in the databases? While even linguistically naive speakers have a notion of a "word," there is no hard definition of a word. A possible orthographic definition would state that strings of letters with an empty space on either side are words. While this would cover words such as *bank*, *sleep*, and *red*, it would wrongly leave out multiword units like *lightning rod*, itfind out, *word of mouth*, and *spill the beans* that constitute semantic and lexical units.[3] Clearly, a lexical unit will merit inclusion in a database when it serves to denote an identifiable concept. But as we shall see, this criterion is less than straightforward.

Assuming at least a working definition of word, the challenge is to arrange the words of a language into a structured lexicon. Although our starting point is the WordNet model, where lexically encoded concepts are interrelated to form a semantic network, we do not take it for granted that the WordNet relations are the most suitable to represent the structure of lexicons of English or other languages. More broadly speaking, we need to ask what constitutes a valid relation among words and concepts both in a given language and crosslingually.

[3] Note that the writing systems of many languages do not separate lexical units; clearly, this does not mean that these languages don't have words.

Finally, we explore the differences and communalities of semantic networks and ontologies. Given the notion of an ontology as a formal knowledge representation system, we ask how the lexicons of many diverse languages can be linked to an ontology such that reasoning and inferencing is enabled. Which relations should be encoded in the ontology and which ones are specific to one or more individual wordnets? Since each wordnet is also an (informal) ontology, incompatibilities between the wordnets and the formal ontology may arise. What do such mismatches tell us and what are the practical consequences for the use of wordnets for reasoning and inferencing and in Natural Language Processing?

4 What Belongs in a Universal Lexical Database?

Both formal, linguistic and informal, cultural criteria determine inclusion in the Global Grid; both turn out to be difficult to define.

Words and phrases that express available concepts must be included. *Availability* is the extent to which a word or phrase is *current* and *salient* within a language community. It affects the topics speakers talk about and the words they use to discuss these topics; it may well affect the way speakers view matters. While frequency and shared cultural background determine the degree of availability of a word or phrase, the *authority* of a speaker or a subgroup of speakers within a language community may have an effect on availability as well. For example, media have a significant influence on the words that are current; frequency counts for a given lexeme vary over time, as the newsworthiness of stories and topics grows and diminishes. Social groups determine availability and linguistic change, as studies of youth language have shown.

Each lexeme of a language is mapped onto a corresponding entry in the ontology. Languages that encode the same concept are thus mappable via the ontology. Adding the lexicons of many languages to the Global Grid will reveal which concepts are truly specific to one language only and which ones are lexicalized in other languages.

4.1 Culture-Specific Words and Concepts

In building a new wordnet and connectings to the English WordNet, one comes across cases where English has no corresponding lexicalized concept. Examples from the Dutch wordnet are the verb *klunen*, which refers to walking on skates over land to get from one frozen body of water to another. Because of different climatic, geographic, and cultural settings, this concept is specific to Dutch and not shared by many other languages (although it can be explained to, and understood by, non-Dutch speakers).

Another example is *citroenjenever*, which is a special kind of gin made with lemon skin. Unlike *klunen*, this *citroenjenever* might well be adopted by inhabitants of English-speaking countries and become a familiar concept.

Culture-specific concepts must be included in the ILI, although there may not be equivalence relations to any languages other than the one that lexicalizes such concept.

4.2 Availability and Salience

The Global Grid should include words and concepts that are available and salient in a linguistic community. This criterion may conflict with purely linguistic criteria for including words in a lexical database. Compound nouns present a case in point.

Standard lexical resources tend to follow the rule that compositional phrases like *dinner table* and *vegetable truck* need not be listed. But non-compositional compounds whose meanings is not the sum of the meanings of their components but where the entire compound is a semantic unit (*horseplay*,*ice luge*) must be included, as their meaning cannot be guessed even by competent speakers that are unfamiliar with these words or concepts. Non-compositionality is only one criteria for inclusion in a lexical database. Even seemingly transparent compounds like *table tennis* and *heart attack* are included in standard dictionaries (e.g., *American Heritage*), presumably because they encode frequent and salient concepts. Hence, these compounds are available to the language community, as ready-made expressions.

Compounds become established in a language community when they are frequent or salient and when their creators have a social standing that lends them what might be called "linguistic authority." This phenomenon can be seen in the areas of science and technology, popular entertainment and commercial branding, where people introduce new terms often with the explicit intention of adding them, along with a new concept, to the lexicon. An example is Dutch *Arbeidstijdverkorting.* Although its members, *Arbeid* ("work"), *tijd* ("time"), and and *verkorting* ("reduction") suggest a straightforward compositional meaning, this compound in fact denotes a special social arrangement invented in the 1980s to create jobs, whereby people got extra spare time in exchange for a reduced salary; the reductions were intended to hire additional workers and decrease unemployment.

Conversely, the following compounds found in today's news headlines are not to be found in any dictionary: *ministry hostages*, *celibacy ruling*, and *banana duty.* Such compounds are created on the fly, and in the context of current news stories, they are readily interpretable, yet their lifespan is limited by their newsworthiness; and only few such ad-hoc compounds will enter the lexicon on a long-term basis.

5 Lexical Mismatches as Evidence for Concepts

As in EuroWordNet, a word in any of the Grid languages will be mapped to the ILI. If the concept is also lexicalized in another Grid language, the two lexicalizations are mapped via their equivalence links to the same ILI entry. Mapping the lexicons of different languages quickly reveals cases where one language encodes a given concept and others do not. But more subtly, it shows up different ways of encoding a concept and raises the question as to what constitutes a word. We illustrate this point with a few speficic cases of semantically complex verbs.

Like nouns, new verbs are regularly formed by productive processes. Different languages have different rules for conflating meaning components. Some components are free morphemes, others are bound affixed. The concepts denoted by compound verbs in one language my be expressed by simplex morphemes in other languages. While one may not want to include complex verbs in one's lexicon based on the argument that they are productive and compositional, the existence of corresponding monomorphemic lexemes in other languages argues for the conceptual status of complex verbs and hence their crosslinguistic inclusion in a multilingual resource.

5.1 Accidental Gaps

Fellbaum and Kegl examine the English verb lexicon in terms on WordNet hierarchies [8]. They argue that English has a non-lexicalized concept "eat a meal," with its own subordinates (*dine, lunch, snack,..* and distinct from the sense of "eat" that denotes the consumption of food and has a number of manner subordinates (*nibble, munch, gulp,...*). Here, the gaps are postulated on the basis of the two semantically distinct verb groups specifying manners of eating. We assume that such gaps are language-specific and that other languages may well have distinct lexicalizations for the two superordinate *eat* concepts.

In fact, a comparison of English and Dutch verbs of cutting reveals a similar crosslinguistic asymmetry. The English verb *cut* does not specify the instrument for cutting something. Only its troponyms do: *snip, clip* imply scissors, *chop* and *hack* a large knife or an axe, etc. Dutch does not have a verb that is underspecified for the instrument, and speakers select the appropriate verb based on the default instrument, which also expresses the manner of cutting (*knippen* "cut with scissors or a scissor-like tool", *snijden* "cut with a knife or knife-like tool", *hakken* "to cut with an axe, or similar tool"). Thus, the lack of a Dutch superordinate verb seems accidental rather than universal.

5.2 Argument Structure Alternations

In some languages, verbal affixes change both the meaning and the argument structure of the base verb. For example, German "be-" is a locative suffix that allows the Location argument to be the direct object. Thus, verbs like *malen* (paint) and *spruehen* (spray) when prefixed with *be-* obligatorily take the entity that is being painted or sprayed (the "Location") as their direct object:

1. Sie bemalte/bespruehte die Wand (mit Farbe)
2. She painted/sprayed/sprayed the wall (with paint)

When the material (the "Locatum") is the direct object, the verb is in its base form:

1. Sie malte/spruehte Farbe an die Wand.
2. She painted/sprayed paint on the wall.

In English, there is no formal difference between the two meanings of such verbs, and it could be overlooked were it not for data from languages like German. However, the structure of the English WordNet forces one to reflect the differences by assuming two distinct senses that members of two different superordinates. The Location variants are manners of *cover* and the Locatum variants are manners of *apply*. A better way of representing the close semantic relation between such verb pairs would be by means of a "Perspective" relation.

6 Perspective

Both the *paint* and *spray* sentences given above can refer to one and the same event.[4] The difference between the sentence could be referred to as one of "perspective." To illustrate what we mean by perspective, we give another example, this one involving two lexically distinct verbs.

Converse pairs like *buy* and *sell* (that are encoded as kinds of semantic opposition in the Princeton WordNet) express the actions of different participants in the same event, a sale in this case.[5] While the verbs and the corresponding nouns each merit their own lexical entries, we want to represent them as encodings of different perspectives on the same event. We propose to do this in the ontology.

SUMO currently distinguishes two processes as well: "Buying" and "Selling." As in FrameNet, both events are a subclass of "FinancialTransaction" and have the same axiom that expresses a dual perspective. The SUMO-KIF representation ([15], [16]) expresses a mutual relation between two statements; one statement in which the Agent of Buying (entity x) obtains something from someone (entity y) that bears the role ORIGIN, and another statement where entity y is the Agent of the Selling process and where the entity x bears the role of DESTINATION.

The ontology thus encodes both entities as agents. A more compact encoding would be one where the two verbs *buy* and *sell* are linked to the same process and the argument structure of each verb can be co-indexed with the entities in the axiom.

Converse and reciprocal events may be encoded very differently across languages. For example, Russian has two different verbs corresponding to English *marry*, depending on whether the Agent is the bride or the groom. And whereas English encodes the difference between the activities of a teacher and a student in two different verbs, *teach* and *learn*, French uses the same verb, *apprendre*, and encodes the distinction syntactically.

[4] It has been suggested that the Location/Locatum alternation in English is accompanied by a subtle semantic difference; Anderson states that the Location alternant implies a "holistic" reading whereby the Location is completely affected [1]. In the first sentence, this would mean that the wall is completely covered with paint. This claim has been challenged, however.

[5] Baker et al. 1998 capture this difference by referring to two different Frame Elements — the Buyer and the Seller — of a single frame [3].

Referring to the event (sale, marriage, etc.) in the ontology allows equivalence mappings to the different languages; the encoding of distinct verbs and roles is then confined to the lexicons of each language.

7 Relations in the Global Grid

We anticipate that some lexical and semantic relations will reside in the ontology while others will be restricted to individual languages. It is an open question, subject to the investigation of of a sufficiently large number of lexicons, which relations will be encoded and where. We cite a few specific cases that must be considered.

7.1 Capturing Semantic Differences Via Language-Internal Relations

Some languages regularly encode semantic distinctions by means of morphology. For example, Slavic languages systematically distinguish between two members of a verb pair; one verb denotes an ongoing event and the other a completed event. English can mark perfectivity with particles, as in the phrasal verbs *eat up* and *read through*. By contrast, Romance languages tend to mark aspect by different verb conjugations on the same verb but make no distinction on the lexical level.

In Dutch, aspectual verbs can be created by prefixing a verb with *door*:

– doorademen, dooreten, doorfietsen, doorlezen, doorpraten
– (continue to breathe, eat, bike, read, talk)

An aspectual relation could be introduced for these languages that links verb synsets expressing different aspects of a given event.[6]

Another example are words marked for biological gender. While *teacher* in English is neutral and underspecified with respect to gender, many profession nouns in German, Dutch, and the Romance languages are not. In Dutch, the morphologically unmarked form *leraar* is masculine and the marked form *lerares* is feminine. While masculine and feminine nouns map to the corresponding nouns in languages that draw this distinction, both map onto a single noun in languages like English.

8 Ontology

The study of ontology goes back at least to Aristotle's "Metaphysics," and, as the name implies, is concered with what exists, i.e., what concepts and categories there are in the world and what the relations among them are. Under

[6] Note that these cases cannot be accommodated with the classical WordNet relations, such as troponymy. The aspectually marked verbs do not make encode manners of either the activity verbs (*eat, read*) or of aspectial verbs like *finish* or *complete*.

this definition, WordNet is an ontology, in that it records both the concepts and categories that a language encodes and the relations among them, including the hyponymy and meronymy relations proposed by Aristotle. For this reason, WordNet is often called a "lexical ontology."[7]

Ontology has another meaning in the context of AI and Knowledge Engineering, where it is the formal statement of a logical theory. For AI systems, what "exists" is that which can be represented. A formal ontology contains definitions that associate the names of entities in the universe of discourse (e.g., classes, relations, functions, or other objects) with human-readable text describing what the names mean, and formal axioms that constrain the interpretation and well-formed use of these terms (see e.g., [9]). One such ontology is SUMO.

9 SUMO

SUMO, a Suggested Upper Merged Ontology [15], allows data interoperability, information search and retrieval, automated inferencing, and natural language processing. SUMO has been translated into various representation formats, but the language of development is a variant of KIF, a version of the first order predicate calculus.

SUMO consists of a set of concepts, relations, and axioms that formalize a field of interest. As an upper ontology, it is limited to concepts that are meta, generic, abstract or philosophical and hence general enough to address a wide range of domains at a high level. SUMO provides a structure upon which ontologies for specific domains such as medicine and finance can be built; the mid-level ontology MILO [17] bridges SUMO's high-level abstractions and the low-level detail of domain ontologies.

SUMO consists of 1,000 terms and 4,000 definitional statements in first order logic language SUO-KIF (Standard Upper Ontology Knowledge Interchange Format). It is also translated into the web ontology language OWL. SUMO has natural language generation templates and a multi-lingual lexicon that allows statements in SUO-KIF and SUMO to be expressed in many languages. SUMO has been fully mapped to the English WordNet and to wordnets in many other languages as well. Synsets map to a general SUMO term or to a term that is directly equivalent to a given synset. New formal terms are defined to cover a greater number of equivalence mappings and the definitions of the new terms depend in turn on existing fundamental concepts in SUMO. SUMO could replace the ILI created for EWN and become the ontology for all wordnets linked to the Princeton WordNet; it is currently the ontology for Arabic WordNet [4]. For example, if the Arabic word sense for *shai* ("tea") is exhaustively defined by relations to SUMO terms, this definition can replace an equivalence relation currently encoded between the Arabic synset *shai* and the English synset *tea* in WN. If there are equivalence relations from wordnets in other languages to the

---

[7] See also, for example, the Ontolinguistic research program at the University of Muenchen [20].

same PWN synset, these synsets can be linked to the same SUMO definitions, as described by Pease and Fellbaum [18].

10 The Architecture of the Grid

Using a natural language as an ILI to link the lexical and conceptual inventories of diverse languages may introduce biases and prevent the adequate representation of concepts in such languages as they must be mediated via the language of the ILI. To avoid this, the Global Wordnet Grid database will comprise a language-neutral, formal ontology as its ILI. This ontology will differ in some important respects from the ILI in EuroWordNet, which is a list of unstructured concepts derived from English WordNet:

1. The list of primitive concepts is not based on the words of a particular language but on ontological observations.
2. The concepts are related in a type-hierarchy and defined with axioms.
3. It is possible to define additional complex concepts using KIF expressions and primitive elements.

A central question that we addressed in this paper is, which concepts should be included in the ILI-ontology? We noted that the ILI-ontology must be able to encode all concepts that can be expressed in any of the Grid languages. However, the ILI need not provide a linguistic encoding — a label — for all words and expressions found in the Grid languages. We saw that many lexicalizations are transparent and systematic while others are non-compositional or seemingly ad-hoc.

We assume a reductionist view and require the ILI-ontology to contain the minimal list of concepts necessary to express equivalence across languages and to support inferencing.

Following the OntoClean method ([10], [11]), identity criteria can be used to determine what is the minimal set of concepts in all cultures where the Grid languages are used. These identity criteria determine the essential properties of entities that are instances of these concepts:

1. *rigidity*: to what extent are properties of an entity true in all worlds? E.g., a person is always a "man" but may bear a Role like "student" only temporarily; "man" is a rigid property while "student" and "father" are anti-rigid.[8]
2. *essence*: which properties of entities are essential? For example, "shape" is an essential property of "vase" but not an essential property of the clay it is made of.
3. *unicity*: which entities represent a whole and which entities are parts of these wholes? An "ocean" represents a whole but the "water" it contains does not.

The identity criteria are based on certain fundamental requirements. These include that the ontology be descriptive and reflect human cognition, perception, cultural imprints and social conventions [13].

[8] See also Carlson's discussion of individual vs. stage level predicates [5] and Pustejovsky's discussion of Roles [19].

The work of Guarino and Welty has demonstrated that the WordNet hierarchy, when viewed as an ontology, can be improved and reduced [10,11]. For example, roles such as AGENTS of processes are anti-rigid. They do not represent disjunct types in the ontology and complicate the hierarchy.

Consider the hyponyms of *dog* in WordNet, which include both types (races) like *poodle*, *Newfoundland*, and *German shepherd*, but also roles like *lapdog* and *herding dog*. Germanshepherdhood is a rigid property, and a German shepherde will never be a Newfoundland or a poodle. But German shepherds may be herding dogs.

The ontology would only list the types of dogs (dog races): *Canine* → *PoodleDog; NewfoundlandDog; GermanShepherdDog,...* If a language lexicalizes a role such as *herding dog*, the type hierarchy of the ILI is not extended, but a KIF expression is created that defines the word. An informal paraphrase of such a definition could be: (instance x Herding dog) ⇆ ((instance x Canine) and (agent x Herding)), where we assume that Herding is a process defined in the type hierarchy as well.

The type/role distinction will clear up many cases where we find mismatches or partial matches between English words and words from other languages. Earlier evaluations of mismatches in EuroWordNet [24] suggest that most mismatches can be resolved using KIF-like expressions and avoiding an extension of the type hierarchy in the ILI with new categories. Gender lexicalizations, differences in perspective, aspectual variants, etc. usually do not represent new types of concepts but can be defined with KIF expressions as well, relating them explicitly to concepts that are types.

When words in the Grid languages suggest new types, the ontological criteria can be used to decide on extensions of the type hierarchy. This is the case not only for culture-specific concepts but also for other kinds of lexicalization differences. For example, the specific ways of cutting lexicalized in Dutch are actually distinct types of processes. In this case, Dutch would be the source for the extension of event types, as the English lexicalization remain too abstract.

In summary, the proposed ILI-ontology has the following characteristics:

1. It is *minimal* so that Terms are distinguished by essential properties only (reductionist)
2. It is *comprehensive* and includes all distinct concept types of all Grid languages
3. It allows the definition of all lexicalizations that express non-essential properties of the types using KIF expressions
4. It is *logically valid* and usable for inferencing

In EuroWordNet, equivalence relations currently vary considerably. Some wordnets only have "exact" equivalence, while others also allow "near equivalence" and have many-to-many relations among synsets and the corresponding concepts in the ILI.

The ILI-ontology we propose here will be more explicit about the meaning of the equivalence relation. Because the ontology is minimal, it will be easier to establish precise and direct equivalences from Grid languages to the ontology and likewise equivalence across languages. The multilingual Grid database will thus consist of wordnets with synsets that are either simple names for ontology types in the type hierarchy or words that relate to these types in a complex way, made explicit in a KIF expression. Note that if two Grid language wordnets create the same KIF expression, they state equivalence without an extended type hierarchy.

11 Towards the Realization of the Global Grid

We propose to take the SUMO ontology as a starting point for three reasons:

– It is consistent with many ontologies and ontological practice.
– It is has been fully mapped onto WordNet.
– Like WordNet, it is freely and publicly available.

SUMO, an upper ontology, is by far not rich or large enough to replace the Princeton WordNet as an ILI-ontology. The current mapping of SUMO to WordNet will be taken as a starting point; most of these mappings are subsumption relations to general SUMO types. The first step is therefore to extend the SUMO type hierarchy so that it becomes as rich as WordNet with respect to disjoint types. Note that not all synsets from WordNet are necessary. In fact, all WordNet synsets must be reviewed with respect to the OntoClean methodology [11] so that only rigid (and semi-rigid) concepts are preserved. All remaining synsets must be defined using KIF expressions as described earlier. For example, the English word *watchdog* would get a simple KIF expression like: ⇆ ((instance x Canine) and (role x GuardingProcess)), where x co-indexes with the referent of the noun.

Subsequently, other languages that have already established equivalence relations with WordNet can replace these with the improved mappings to SUMO, which can be copied from the Princeton WordNet. For example, Dutch *poedel* and Japanese *pudoru* will become simple names for the type ⇆ ((instance x Poodle), because they are equivalents of WordNet synset *poodle*. Similarly, Dutch *waakhond* and Japanese *banken* wold be linked to the same KIF expression as both are equivalent to *watchdog* in WordNet; the KIF expression can simply be copied.

In other cases, the equivalence relations to WordNet may require some revision as it is now possible to express certain subtle distinctions between concepts expressed in the Grid languages and corresponding ones in English that could not be expressed in the EuroWordNet model. For example, the Dutch verb *bankdrukken* will be related to the English noun *bench press*, meaning *a weightlifting exercise*, because there is no corresponding verb synset in WordNet. The part-of-speech mismatch does not allow a direct match between the Dutch and English synsets. The ontology will include a process "BenchPress" that is not marked for part of speech; both the English noun and the Dutch verb can be linked to this same process. This does not prevent us from indicating further differences, such as aspectual meaning.

Importantly, this design makes it unnecessary to write separate KIF expressions for ontological concepts in each language — most expressions can be linked via their relations to English synsets and revisions are required in some cases only.

However, the situation will arise where synsets in Grid languages cannot be mapped to WordNet. In those cases, the concepts represented by these synsets need to be checked for adherence to OntoClean. This step may result in extensions to the type hierarchy in some cases; in other cases, the wordnet builders need to write a KIF expression clarifying the particular concept's relation to the ontology. For example, The Dutch noun *straathond* (street dog), which is not mapped to WordNet, can be defined relatively easily: ⇆ ((instance x Canine) and (habitat x Street)), following the model of similar expressions such as *watchdog* and *herding dog*.

Synsets that are not disjunct types usually have a relatively straightforward semantic structure and the KIF expressions in many cases can be copied from similar synsets that can be identified in the ontology by browsing through the hierarchy of roles and processes.

We are aware that highly specific concepts, restricted to a given culture, may be present difficulties for providing an exhaustive and satisfactory KIF expression. A solution is to provide a definition, or "gloss," in the Grid language, a corresponding English gloss, and the most specific superclass in the type hierarchy of the ontology.

New types could be created and built in a Wiki environment. Initially, full definitions are not necessary; it is more important to have a comprehensive list of type candidates that become more precisely defined by the community in the course of the Grid construction. Furthermore, the possibility should be explored to allow the creation of KIF expressions via a simple interface or questionnaire that makes such expressions accessible to linguists and speakers of a language unfamiliar with ontology.

12 Conclusion

The Global Wordnet Grid can only be realized in a collaborative framework among builders of wordnets from many diverse linguistic and cultural backgrounds. Its development will undoubtedly include several steps and many rounds of refinement. Throughout the development of the Global Wordnet Grid, we expect discussion and the need for revisions as more languages join and the coverage for each language increases. Mapping the lexicons of many diverse languages, and the cultural notions they encode, is bound to be a long and painful process, but also a worthwhile one. The result will be a unique database that allows for a better understanding among people from different linguistic and cultural backgrounds and opens up new possibilities for research and applications.

We think it is important that such a database is built on a large scale and that it is based on a diverse set of languages and cultures. The languages will form the empirical evidence and basis for the construction of a truly universal index.

References

1. Anderson, S.R.: On the Role of Deep Structure in Semantic Interpretation. Foundations of Language 7, 387–396 (1982)
2. Apresyan, J.D.: Regular Polysemy. Linguistics 142, 5–32 (1973)
3. Baker, C., Fillmore, C.J., Lowe, J.B.: The Berkeley FrameNet project. In: Proceedings of the COLING-ACL, Montreal, Canada (1998)
4. Black, W., Elkateb, S., Rodriguez, H., Alkhalifa, M., Vossen, P., Pease, A., Bertran, M., Fellbaum, C.: The Arabic WordNet Project. In: Proceedings of the Conference on Lexical Resources in the European Community, Genoa, Italy (2006)
5. Carlson, G.: Reference to Kinds in English. Garland Press, New York (1980)
6. Fellbaum, C.: The English Verb Lexicon as a Semantic Net. International Journal of Lexicography 3, 278–301 (1990)
7. Fellbaum, C. (ed.): WordNet: An Electronic Lexical Database. MIT Press, Cambridge, MA (1998)
8. Fellbaum, C., Kegl, J.: Taxonomic Structure and Object Deletion in the English Verbal System. In: de Jong, K., No, Y. (eds.) Proceedings of the Sixth Eastern States Conference on Linguistics, pp. 94–103. Ohio State University, Columbus, OH (1989)
9. Gruber, T.R.: A translation approach to portable ontologies. Knowledge Acquisition 5, 199–220 (1993)
10. Guarino, N., Welty, C.: Identity and subsumption. In: Green, R., Bean, C., Myaeng, S. (eds.) The Semantics of Relationships: an Interdisciplinary Perspective, Kluwer Academic Publishers, Dordrecht (2002)
11. Guarino, N., Welty, C.: Evaluating Ontological Decisions with OntoClean. Communications of the ACM 45(2), 61–65 (2002)
12. Levin, B.: English Verb Classes and Alternations. University of Chicago Press, Chicago, IL (1993)
13. Masolo, C., Borgo, S., Gangemi, A., Guarino, N., Oltramari, A.: WonderWeb Deliverable D18 Ontology Library. Laboratory For Applied Ontology - IST-CNR, Trento (2003)
14. Miller, G.A. (ed.): WordNet. Special Issue of the International Journal of Lexicography, 3 (1990)
15. Niles, I., Pease, A.: Towards a Standard Upper Ontology. In: Proceedings of the 2nd International Conference on Formal Ontology in Information Systems, Ogunquit, Maine (2001)
16. Niles, I., Pease, A.: Linking Lexicons and Ontologies: Mapping WordNet to the Suggested Upper Merged Ontology. In: Proceedings of the International Conference on Information and Knowledge Engineering, Las Vegas, Nevada (2003)
17. Niles, I., Terry, A.: The MILO: A general-purpose, mid-level ontology. In: Proceedings of the International Conference on Information and Knowledge Engineering, Las Vegas, Nevada (2004)
18. Pease, A., Fellbaum, C.: Formal Ontology as Interlingua. In: Huang, C.R., Prevot, L. (eds.) Ontologies and Lexical Resources, Cambridge University Press, Cambridge (2007)
19. Pustejovsky, J.: The Generative Lexicon. MIT Press, Cambridge, MA (1995)
20. Schalley, A.C., Zaefferer, D. (eds.): Ontolinguistics. Mouton de Gruyter, Berlin (2007)
21. Sinha, M., Reddy, M., Bhattacharyya, P.: An Approach towards Construction and Application of Multilingual Indo-WordNet. In: Proceedings of the Third Global Wordnet Conference, Jeju Island, Korea (2006)

22. Tufis, D. (ed.): The BalkaNet Project. The Romanian Journal of Information Science and Technology 7, 1–248 (2004)
23. Vossen, P. (ed.): EuroWordNet. Kluwer, Dordrecht (1998)
24. Vossen, P., Peters, W., Gonzalo, J.: Towards a Universal Index of Meaning. In: Proceedings of ACL-99 Workshop, Siglex-99, Standardizing Lexical Resources, University of Maryland, College Park, Maryland, June 21-22, pp. 81–90 (1999)

Towards a Conceptual Core for Multicultural Processing: A Multilingual Ontology Based on the Swadesh List

Chu-Ren Huang, Laurent Prévot, I-Li Su, and Jia-Fei Hong

Institute of Linguistics, Academia Sinica
Taipei, Taiwan
{churen,prevot,isu,jiafei}@gate.sinica.edu.tw

Abstract. The work presented here is situated in the broader project of creating of multilingual lexical resources with a focus on Asian languages. In the paper, we describe the design of the upper-level we are creating for our multi-lingual lexical resources. Among the current efforts devoted to this issue our work put the focus on (i) the language diversity aiming at massively multi-lingual resource, and (ii) the attention devoted to the ontological design of the upper level.

Keywords: ontology, lexical resource, multilinguality, Swadesh list.

1 Introduction

The work presented is this paper is situated in the broader project of creating of multilingual lexical resources with a focus on Asian languages. When approaching the domain of lexical resources and their use in Natural Language Processing comes the question of the task repartition between the lexical and conceptual levels. This has been investigated for a long time under the light of philosophical or formal knowledge representation principles [5,17]. Recently, the “ontological trend” generated an important amount of work concerning these ontologies from both knowledge engineering and computational linguistics perspectives. These projects often differ radically in the way they handle the ontology-lexicon interface but a common concern they share is the design of an upper level for the resource: EuroWordNet [20,21], SIMPLE ontology [8], Ω (Omega) [16], OntoSem [12], SUMO-WN [14], OntoWordNet [4]. Among these projects, OntoSem is the first to have multilinguality explicitly on their agenda, while neither SUMO-WN nor Onto-WordNet has explicit design for multilinguality.

In this paper, we describe the design of the upper-level we are creating for our multi-lingual lexical resources. Our work put the focus on:

– the language diversity it is covering,
– the attention devoted to the ontological design of the upper level.

As for the language diversity, the multi-national project (“Developing International Standards of Language Resources for Semantic Web Applications”) [19]

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 17–30, 2007.

in which the work take place regroups Japanese, Thai, Italian and Taiwanese teams for creating a multilingual resource aligned with the Princeton WordNet [3]. Moreover, within our team we benefit of the input from other languages such as Bangla, Malay, Taiwanese, Cantonese and Polish. There are two levels of development for these languages: for the languages represented by the project members (Japanese, Thai, Chinese and Italian) the goal is to create basic but significant core lexicons that can be compared to the Base Concept of EuroWordNet [20] and Global WordNet Grid[1]. About the other languages we only collected their respective Swadesh lists [18] in the perspective of building a minimal but massively multilingual lexical resource.

About the ontological aspect, in spite of some efforts devoted to the design of the upper level, the existing resources were not clear enough about their ontological commitment which was making them somehow difficult to compare. We aim here to compare them more thoroughly under the light of the recent works in formal ontology in order either to select the most appropriate model for the upper-level of multilingual lexical resources.

After this introduction, the next section will present the different methodologies that can be used for selecting a core lexicon while the section 4 describe the compilation of the Swadesh list for the languages considered. Then the section 3 describes our experiments for designing a prototype for the core upper level. We then investigate how the coverage of our resource can be extended (section 6).

2 Approaches for Designing a Core Lexicon

Traditional approaches considered for establishing a compact list of basic terms (or *core lexicon*) can be divided into two categories according to their criteria for selecting the terms: *semantic primacy* and *frequency*. In addition in this section we propose a third way to be explored: the *universality* criterion.

2.1 Frequency Criterion

The first intuitive idea for selecting a core lexicon is to uses statistical information such as word frequency. However, this naive approach of simply taking the most frequent words in a language is flawed in many ways. First, all frequency counts are corpus-based and hence inherit the bias of corpus sampling. For instance, since it is easier to sample written formal texts, words used predominantly in informal contexts are usually underrepresented. Second, frequency of content words is topic-dependent and may vary from corpus to corpus. Last, and most crucially, frequency of a word does not correlate to its conceptual necessity, which should be an important, if not only, criteria for core lexicon. The definition of a cross-lingual basic lexicon is even more complicated. The first issue involves determination of cross-lingual lexical equivalences. That is, how to determine

[1] See http://www.globalwordnet.org/gwa/gwa_grid.htm

that word `a` (and not `a'`) in language A really is word `b` in language B. The second issue involves the determination of what is a basic word in a multilingual context. In this case, not even the frequency offers an easy answer since lexical frequency may vary greatly among different languages. The third issue involves lexical gaps. That is, if there is a word that meets all criteria of being a basic word in language A, yet it does not exist in language D (though it may exist in languages B, and C). Is this word still qualified to be included in the multilingual basic lexicon?

A recent elaboration [22] proposed to use the notion of *distributional consistency* rather than crude frequency. This measure provides better result than other statistically based approaches but it requires balanced corpus of significant size to be applicable. Such corpora are only available for few languages and we would like to have a method that could be used with languages deprived from extensive resources.

2.2 Semantic Primacy Criterion

To answer about the lack of consideration of the "conceptual necessity" of the terms selected by the *frequency* approach, it is natural to consider more foundational work concerning knowledge organization. The idea of this approach is to determine a list of concepts that are semantic primitives (or atoms) that cannot be easily defined from other concepts. These concepts are located in the upper part of various hierarchical models. Each new distinction is made on the base of clear different semantic features. The main problem of these semantic primitives is their abstractness that make them rarely lexicalized (e.g 1stClassEntity in *EuroWordNet* top-level [20], NonAgentiveSocialObject in *Dolce* [11] or SelfConnectedObject in *SUMO* [13])[2]. These upper-levels will have a role to play in the design of our resource but they are not so useful for the constitution of the lexical core we are thinking of. We would like the basic building block of our resource to come from linguistic source, corresponding to the "linguist" ontology builder types described by Eduard Hovy in [6].

2.3 Swadesh List or the Universality Criterion

The lack of resources for most of languages led us to consider the Swadesh list [18] (reproduced as an appendix) as a potential core lexicon. The Swadesh list has been developed by Moriss Swadesh in the fifties for improving the results of quantitative historical linguistics. His list remains as a widely used vocabulary of basic terms. The items of the list are supposed to be as universal as possible but are not necessarily the most frequent. The list can be seen as a least common denominator of the vocabulary. It is therefore mainly constituted by terms that embody human direct experience. The list is 207 items long and is composed by the totality of the 200-item Swadesh first list, plus 7 terms coming from a 100-item list that Swadesh proposed later. This list is available for a

[2] In this paper we use SmallCaps font for concepts and `TypeWriter` font for terms.

great number of languages and its inclusion in the resources being collected in the context of the Rosetta project[3] warrants the quality and the maintenance of the resource. Moreover the Swadesh list items have been selected for their universality. Although quite different from the semantic primacy, this criterion ensure some kind of linguistic primacy that we are interested in.

These characteristics qualify the list has an interesting starting point for building a core lexicon in many different languages and for establishing easily the translation links. However, the methodology for establishing the list (essentially dictated by Swadesh's field work) introduces several issues that we have to deal with.

First, although made of lexical atoms, nothing prevents many other potential atoms to be discarded simply because of their lack of relevance for lexico-statistic purposes. This issue is specially important because it forbids us, when trying to propose a structure for the list, to think the Swadesh list as a definitive list of concept. As a consequence, a room for subjective appreciation remains open for introducing new concepts in the list.

The second issue results also from the initial purpose of the list. To be usable in field work context, the list concerns only direct human experience and avoids completely other foundational domains. On the other hand there is a richness for verbs describing human everyday activities that do not require modern tools.

Finally the Swadesh list, by its nature, has been established for spoken language in the context of face-to-face interaction.

3 Experiments: Designing a Core Ontology from the Swadesh List

3.1 The Experiments on Chinese

The Chinese Swadesh list was obtained by consulting with the Academia Sinica Chinese Wordnet group. One or more Chinese Wordnet entry for each item of the list were obtained, and non basic readings were eliminated. Subsequently, we obtain automatically the concept distribution of the items in SUMO taxonomy through SINICABOW,[4] a resource developed at the Academia Sinica which combines the Chinese wordnet, the Princeton WordNet and SUMO [13].

3.2 The Experiments with English

About the English list we studied three different ways for building a taxonomy out of the simple list:

A. Keep the structure as minimal as possible by not adding any further (generalizing) concept in the list.
B. Keep the structure as minimal a possible but also try to get a reasonable organization from a knowledge representation viewpoint.
C. Simply align the terms to SUMO ontology [13] and prune the result.

[3] See http://www.rosettaproject.org/ for more information.
[4] See [7] and *http://bow.sinica.edu.tw/* for more information.

The first experiment (A) was not very conclusive since the list itself only includes very few words that are situated at different specificity level. The Swadesh items are indeed typically situated at the basic or generic level of categorization specificity [2]:p82. It is therefore expectable that they do not present a lot of taxomomic relations among them.

For the experiments (B) and (C), we proceeded in two steps:

1. Disambiguate the Swadesh List items by associating each of them to with one WordNet synset.
2. Create the taxonomy.

For (B), the taxonomy was created manually in a bottom-up fashion, by grouping the terms into more general categories while trying to keep the taxonomy as intuitive and minimal as possible. It resulted in about 220 classes organized in a preliminary taxonomy. Some generalization levels are missing since too few Swadesh items were corresponding to these areas.

In the case of the SUMO version (C), once we got the WordNet synsets the further mapping to SUMO was immediate once thanks to the mapping proposed in [14].

As for the technical aspects, our mapping operations have been done semi-automatically under Protégé[5] and more specifically with the help of ONTOLING[6] plug-in. The existing resources we used were WordNet 2.1 and an OWL transaltion of SUMO.[7]. The results of the experiences are available in OWL format on this website.[8]

4 Comparing Lexicalization Patterns

In addition of Chinese and English we compiled the Swadesh list for Bangla, Malay, Cantonese and Taiwanese from native speakers (students and colleagues). The universality aim of the Swadesh list was confirmed in this experiments. The Swadesh list is extremely well covered in the languages we studied. The only item that was said to not be lexicalized is `stab` in Bangla that is translated by `thiknagro-ostro-die-aghat-kora` which means literally *hit-with-a-sharp-instrument*.

An interesting issue has been raised by the Malay data in which several Swadesh items received the same Malay equivalent:

– `kaki`: both `foot` and `leg`
– `perut`: both `belly` and `gut`
– `hati`: both `heart` and `liver`
– `jalan`: both `road` and `walk`

[5] For more information, visit *http://protege.stanford.edu/*

[6] For more information, see [15] and visit *http://ai-nlp.info.uniroma2.it/software/OntoLing/*

[7] Available at *http://www.ontologyportal.org/translations/SUMO.owl.txt*

[8] See *http://www.sinica.edu.tw/∼prevot/Swadesh/*

For example `perut` corresponding to both `belly` and `gut` in the Swadesh list but for which the most natural translation is `stomach` can be compared to *nari-vuri* (`gut` again) in Bangla which is a compound word made of *nari* (small intestine) and *vuri* (large intestine).

These examples emphasizing the complexity of the lexical-conceptual relation. But might also raise the issue of cultural specificities at the conceptual level itself. In order to not loose such information our conceptual upper level cannot come only from experiments done in a given language or by a team of ontologists from a given culture. Such an language independent structure is likely to emerge in the upper level but we should not impose it as a starting point.

The structure presented in [21] also supports such a view. In Wordnets there are words and synsets (or word senses). The synsets across languages are not the same and their semantic organization differs. These synsets structures constitute real language-dependent lexical ontologies. It is worthwhile to consider these sense organizations coming from the languages before wrapping them up in an already established ontologies usually developed mono-culturally.

5 The Problems Encountered So Far

5.1 Function Words

A significant amount of words in the list (28 out of 207) are pronouns, demonstratives, quantifiers, connectives and prepositions. These words do not play a direct role in a taxonomy of the entities of the world. Unsurprisingly, many of them are absent from both WordNet and SUMO (e.g `you, this, who, and,...`). Quantifiers are present in WordNet (in adjectives) but placing them in the taxonomy is a thorny issue. SUMO-WordNet grouped them mysteriously under the EXISTS concept together with concepts such as `living`. About prepositions, some are present in WN (e.g `in`) but some other not (e.g `at`). In the beginning phase of the project, we simply decided to isolate all these words and to defer the discussion about them for later.

5.2 Ambiguities

The success of the Swadesh list is partly due to its under-specification and to the liberty it gives to compilers of the list. The absence of gloss results in genuine ambiguities, although some of them are partially removed through minimal comments added in the list (e.g `right` (correct), `earth` (soil)) and the implicit semantic grouping present in the list. More complex cases include terms like `snow` or `rain` that may refer to a meteorological phenomenon or to a substance. In such cases we allowed ourselves to integrate both meanings in the taxonomy (e.g SNOWSUBSTANCE *is-a* SUBSTANCE, SNOWFALL *is-a* PHENOMENON).

In this precise case, this ambiguity might be resolved by considering the semantic grouping that is sometimes proposed in the list (here `snow` appear together with `sky, wind, ice` and `smoke`).

For dealing with polysemy the solution could be to position the given polysemous *synsets* under several ontological concepts. However, placing in a taxonomy a term under two incompatible concepts results in an inconsistent resource. A way to deal with this problem, is to have only a few core meanings (ideally one) and derive the other senses from a richer relation network including other relations than hyperonymy. The generative lexicon [17] is an illustration of this possibility where the simple taxonomic link is replaced by four different relations.

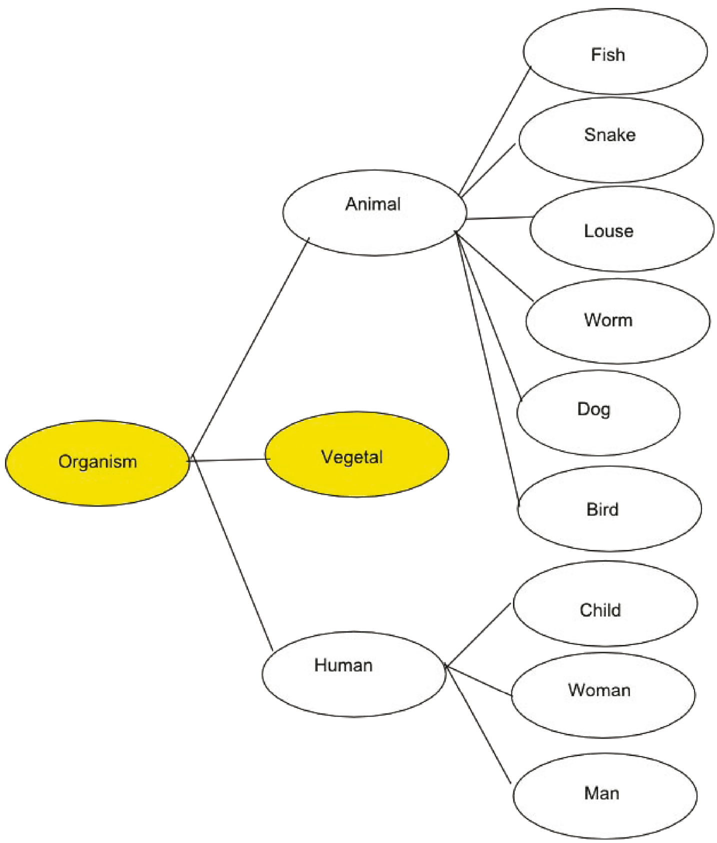


Fig. 1. Manual bottom-up taxonomy extrapolation (Method B), Shaded classes added from Swadesh list

5.3 Granularity Heterogeneity and More General Categories

Here the methodology chosen (A, B or C) introduces different issues. In the case of A, we actually did not succeed in identifying a structure where the nodes will be lexicalized by the items of the list. At best we get clusters than can be grouped under a general concept though extremely vague relation. For example, `sea, lake` and `river` might be grouped under WATER with option A. But so can be `rain, snow, ice` or `cloud` and why not having also `wet` and `drink`, `swim` or even `fish`. All these terms are *associated* with `water` but that do not qualify them for being equally positioned under WATER as a concept in a taxonomy. An on-going extension of WordNet concerns the addition of these loose links between the terms [1]. According to this study, such relations could remain unlabeled. However a step further could consist in identifying more precisely the nature of these "associations". For example, many of these terms refer to entities *constituted-by* WATER, others are *physical-state* of WATER or activities involving WATER. But adding these precisely links drive us away from the initial stage of our project.

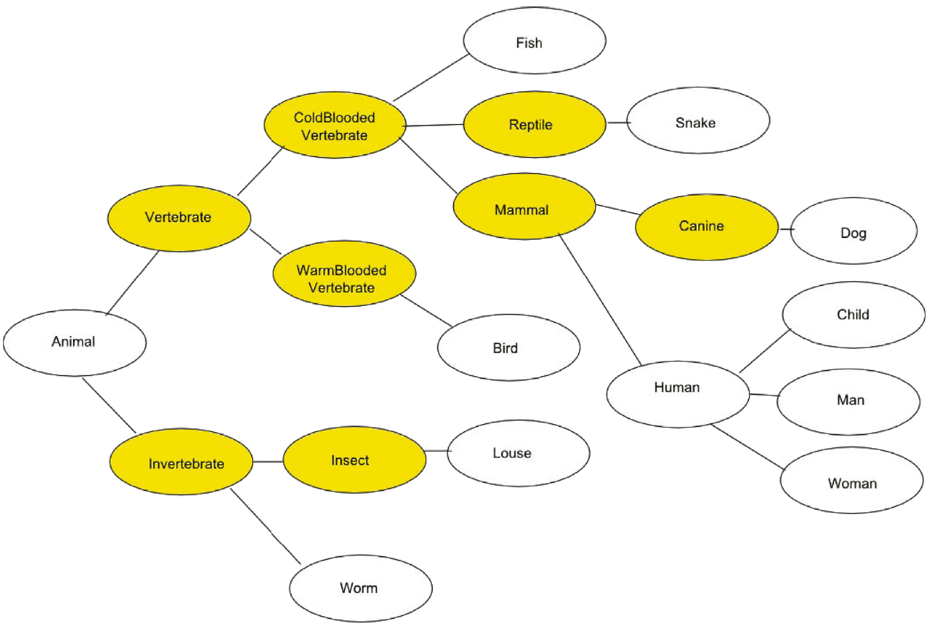


Fig. 2. SUMO alignment, pruning and trimming (Method C), Shaded classes added from Swadesh list

The options B and C actually takes us a step further away by introducing many new categories for disambiguating the terms and for accounting for intermediates levels such as BodyParts, Processes... (See Fig. 1 and 2).

The ontologies resulting from (B) and (C) experiments, (B) has much flatter structure than (C). The ontology coming from the SUMO filtering includes actually a lot of intermediate levels that are not necessary for classifying satisfactorily the Swadesh items. As a consequence, the resulting structures need to be pruned and trimmed as illustrated in the figure 2.

5.4 Conceptual Discrepancies?

The last issue concerns the discrepancies about the world conceptualization between on the one hand direct human experience viewpoint and on the other hand the modern scientific viewpoint. For example, which relations we will retain in our ontology for terms such as `sun, star` and `moon`. There is no such term as `satellite` in the list and nothing indicates that this concept is relevant for a direct human experience viewpoint. For now, while following the options B and C, we made the less committing choice. In this example we placed all the terms in question under the AstronomicalBody SUMO concept and under SkyObject for our own taxonomy proposal.

When turning to cross-cultural studies, it becomes clear there are different lexical (ar potentially conceptual) organizations for a given domain. See for example, the case of body parts in Rossel Island [9] or the one of geographical objects in Australia [10]. More examples (perhaps lexical only) are coming from the lexical gaps that are frequent as it has been noticed in the different lexical multilingual resources projects [20].

These issue highlight again the need to separate the lexical (*words*), semantic (*word senses* or *synsets*) and the ontological level (*formal concepts*). When facing a difference in lexical organization like the one observed in Malay. The word `perut` corresponds both to `belly` and `intestine`, there are four options for dealing with such a situation:

1. `perut` as a word having one "ambiguous" meaning ⟨*perut*⟩[9] corresponding to two different concepts in the ontology Belly and Intestine. Here a meronymy relation might holds between the two concepts –*Part-of*(Intestine,Belly)– which could explain the polysemy.
2. `perut` as a word having two meanings $\langle perut_1 \rangle$, $\langle perut_2 \rangle$, respectively mapped to Intestine and Belly. This will correspond to the SUMO-WordNet case since the sense division in WordNet is extremely fine grained.
3. `perut` as a word having one meaning ⟨*perut*⟩ requiring the creation of a new concept in the ontology corresponding roughly to the Belly and the Intestine together. In some case it might be difficult to define the new category on the base of other categories. Still we can assume as a working assumption that starting from a foundational ontology all new concepts are definable in terms of the ones already present. This model results in two types of categories in the ontology: stated (from axioms) and inferred (theorems). These apparently technical considerations are related to the discussions on the nature of the categories of the mental lexicon. These categories include both stable learned categories and other ones that need to be computed.

[9] In the rest of the paper the synsets will be written in italic within these brackets ⟨*synset*⟩.

4. Finally, `perut` as a word having a vague meaning ⟨*perut*⟩ corresponding to an ontological concept with a weak characterization that might catch more word senses from different languages and that would not quite fit in the picture if we restrict too precisely the intented meaning of our vocabulary. Ontology is seen as a expression of shared ontological commitments as precise as possible for avoiding misunderstandings. Therefore, ontologists probably want to rule out this last option, keep these complexities at the "linguistic" level and start using ontology with the concepts and their definitions precisely established.

Lexical gaps do not implies conceptual gaps and in many case it might be tempting to handle multilingual lexical discrepancies with complex mapping described (cases 1,2). However, we should not discard completely the case 3 and 4 specially when dealing with very different cultures.

6 Extending the Word List

As emphasized earlier in this paper, the Swadesh list offer an interesting starting point for an universal core lexicon but is not enough by itself. In this section we compare the coverage of the Swadesh list with the one of the Base Concept Set [20,21] as it is proposed by the Global WordNet Association[10]. Since both the Swadesh list and BCS are linked to SUMO, we are in position to compare the repartition of their mappings to SUMO.

The BCS synsets are mapped to 928 SUMO concepts, the most frequent mapping are presented in the table 1. In the table, the number of mappings from the Swadesh list is also provided. The number in parentheses corresponds to the number of indirect mappings (e.g `cut` is mapped to Cutting but Process is a parent of Cutting). In this example, although Process did not receive any mapping, many Swadesh items are classified under it. Below the double line of the table is included the SUMO concepts hosting as significant number (according to the list size) of Swadesh list mappings without being among the most common host for Base Concept synsets.

In both case the SubjectiveAssessmentAttribute is the most frequently mapped. In the case of the Swadesh list mapping we found all the adjectives that present a certain degree of subjectivity (e.g `bad, new, dirty...`) (See the documentation for this concept in figure 3). We expect that once a more comprehensive model for *qualities* (or properties) proposed in the ontology many adjectives should find a more satisfactory place in the model.

More significant are the SUMO categories for which the repartition of the BCS was strikingly different from the one of the Swadesh list. All categories related to Artifact, Device (and therefore TransportationDevice) are totally absent from Swadesh list but among the most frequent mappings for the BCS. In addition of this expected result, IntentionalPsychologicalProcess, Communication, Radiating Sound, Group, Position, SocialRole and Text are also almost absent from the Swadesh list. This suggest the direction in which

[10] See http://www.globalwordnet.org/gwa/gwa_base_concepts.htm

Table 1. SUMO concepts receiving the more mapping from BCS and Swadesh

SUMO Concept	Mapping BCS	Mapping Swadesh
SubjectiveAssessmentAttribute	338	18
IntentionalProcess	93	1 (11)
Process	84	0 (64)
Motion	78	7 (25)
Device	70	0 (0)
Artifact	64	1 (2)
Communication	62	1 (2)
IntentionalPsychologicalProcess	51	0 (2)
RadiatingSound	46	0 (1)
BodyPart	42	16 (31)
Putting	41	0 (0)
Removing	40	1 (1)
Region	39	1 (9)
TimeInterval	38	2 (2)
Group	37	0 (0)
ShapeAttribute	36	4 (4)
Position	36	0 (0)
Text	35	0 (0)
TransportationDevice	34	0 (0)
Human	33	1 (4)
Increasing	32	1 (1)
part	32	0 (0)
SocialRole	32	0 (0)
Touching	20	4 (4)
Organ	10	8 (8)
Impelling	8	4 (4)
ColorAttribute	2	5 (5)

```
(documentation SubjectiveAssessmentAttribute "The &%Class of
&%NormativeAttributes which lack an objective criterion for their
attribution, i.e. the attribution of these &%Attributes varies from
subject to subject and even with respect to the same subject over
time. This &%Class is, generally speaking, only used when mapping
external knowledge sources to the SUMO.  If a term from such a
knowledge source seems to lack objective criteria for its attribution,
it is assigned to this &%Class.")
```

Fig. 3. Documentation for SubjectiveAssessmentAttribute SUMO concept

we sould extend the Swadesh list items in priority: social objects, mental objects, artifacts and communication.

Finally, some of the frequent Swadesh list item are not well represented for BCS. In the *ColorAttribute*the Swadesh list includes the five primary colors, BCS only includes ⟨*shade*⟩ and ⟨*colored*⟩. Another oddity is the the presence of

$\langle wife \rangle$ but not $\langle husband \rangle$ in BCS while both are present in Swadesh list. These unexpected holes in the BCS shows that even a small list like the Swadesh might present some interesting suggestions for the design of the core lexicon.

7 Conclusion and Future Work

In this paper we investigate the idea of using the Swadesh list as a central resource for developing massively multilingual resources. Among the variety of on-going efforts on the development of multilingual resources, our project put the focuse on:

- Language diversity: the languages we consider are not only European languages as EuroWordNet [20] or SIMPLE [8].
- Solidity of the ontology: We do not want to commit our upper level too early to an existing ontology. We prefer to carefully compare the existing proposals, trying to understand the design choices for determining which ones are the most pertinent for the upper-level of a multilingual lexical resource.

As for the Swadesh list, we identified some limitations for this resource and emphasized some benefits of its usage. More precisely, it can be used as a good starting point for developing a linguistic ontology of direct human experience for a great number of languages. Such a resource is useful:

- (i) *per se*, for comparing different versions of the different lexical organizations (if there is more than one) and investigate the hypotheses of the relativist/universalist debate.
- (ii) as a first step for constituting a more applicative core lexicon for direct human experience that can should be integrated with core lexicon to form a full lexicon.

About (i), it is clear that more empirical experiments are needed in order to establish the structure underlying the list. An interesting approach could be to start with unlabeled semantic relations as described in [1] and later try to specify these relations according to their semantics.

About (ii), the Swadesh list, being limited to direct human experience and established in a spoken language context, has to be efficiently complemented by basic concepts of foundational knowledge areas (such as ARTEFACTS) for increasing its interest as a resource for NLP. Another important aspect is the integration of other relations than the taxonomic one in order to address the polysemy issue as described in 5.2. About this last point we are currently encoding the meronomic relations as well as various participation relations of objects to processes (thematic roles relations).

Finally for many languages other resources are simply not available. Developing such a micro-lexicon can be taken as a seed for developing a more significant lexicon around it. The idea is to have a simple upper level that is not disproportionated compare to the size of the lexicon but that can be seen as a first step toward the integration of a new languages in the Semantic Web. In this perspective the experiment B (manual extrapolation from a core lexicon) might be easier to perform than the integration of the terms in an already existing complex resource (experiment C).

To test further these ideas, we are currently working in collaboration with colleagues in Vietnam and the Philippines for extending the experiment to more languages.

Acknowledgment. We would like to thank the people involved in this project in Taipei, Kamrul Hasan, Sophia Lee, Siaw-Fong Chung, Tian-Jian Jiang, Katarzyna Horszowska Yong-Xiang Chen and the NEDO project members that provide the Swadesh lists. This research was carried out through financial support provided under the NEDO International Joint Research Grant Program (NEDO Grant).

References

1. Boyd-Graber, J., Fellbaum, C., Osherson, D., Schapire, R.: Adding dense, weighted connections to wordnet. In: Proceedings of the Third International WordNet Conference (2006)
2. Croft, W., Cruse, A.: Cognitive Linguistics. Cambridge University Press, Cambridge (2004)
3. Fellbaum, C. (ed.): WordNet: An Electronic Lexical Database. MIT Press, Cambridge (1998)
4. Gangemi, A., Navigli, R., Velardi, P.: The ontowordnet project: extension and axiomatisation of conceptual relations in wordnet. In: Meersman, R., Tari, Z., Schmidt, D.C. (eds.) CoopIS 2003, DOA 2003, and ODBASE 2003. LNCS, vol. 2888, Springer, Heidelberg (2003)
5. Hobbs, J.R., Croft, W., Davies, T.R., Edwards, D., Laws, K.I.: Commonsense metaphysics and lexical semantics. Computational Linguistics 13(3-4), 241–250 (1987)
6. Hovy, E.: Methodologies for the reliable construction of ontological knowledge. In: Dau, F., Mugnier, M.-L., Stumme, G. (eds.) ICCS 2005. LNCS (LNAI), vol. 3596, Springer, Heidelberg (2005)
7. Huang, C.-R., Chang, R.-Y., Lee, S.-B.: Sinica BOW (bilingual ontological wordnet): Integration of bilingual WordNet and SUMO. In: 4th International Conference on Language Resources and Evaluation (LREC2004), Lisbon (2004)
8. Lenci, A., Bel, N., Busa, F., Calzolari, N., Gola, E., Monachini, M., Ogonowski, A., Peters, I., Peters, W., Ruimy, N., Villegas, M., Zampolli, A.: Simple: A general framework for the development of multilingual lexicons. International Journal of Lexicography (2000)
9. Levinson, S.C.: Parts of the body in yeli dnye, the papuan language of rossel island. Language Sciences 28, 221–240 (2006)
10. Mark, D.M., Turk, A.G.: Landscape categories in yindjibarndi: Ontology, environment, and language. In: Kuhn, W., Worboys, M.F., Timpf, S. (eds.) COSIT 2003. LNCS, vol. 2825, Springer, Heidelberg (2003)
11. Masolo, C., Borgo, S., Gangemi, A., Guarino, N., Oltramari, A.: Wonderweb deliverabled18, ontology library (final). Technical report, LOA-ISTC, CNR (2003)
12. McShane, M., Zabludowski, M., Nirenburg, S., Beale, S.: Ontosem and simple: Two multi-lingual world views. In: ACL 2004 Workshop on Text Meaning and Interpretation, Barcelona, Spain (2004)

13. Niles, I., Pease, A.: Towards a standard upper ontology. In: Welty, C., Smith, B. (eds.) Proceedings of the 2nd International Conference on Formal Ontology in Information Systems (FOIS-2001), Ogunquit, Maine (2001)
14. Niles, I., Pease, A.: Linking lexicons and ontologies: Mapping wordnet to the suggested upper merged ontology. In: Proceedings of the 2003 International Conference on Information and Knowledge Engineering (IKE 03), Las Vegas, Nevada (2003)
15. Pazienza, M.T., Stellato, A.: The protégé ontoling plugin - linguistic enrichment of ontologies. In: Gil, Y., Motta, E., Benjamins, V.R., Musen, M.A. (eds.) ISWC 2005. LNCS, vol. 3729, Springer, Heidelberg (2005)
16. Philpot, A., Hovy, E., Pantel, P.: The omega ontology. In: Proceedings of ONTOLEX' 2005 (2005)
17. Pustejovsky, J.: The generative lexicon. MIT Press, Cambridge (1995)
18. Swadesh, M.: Lexico-statistical dating of prehistoric ethnic contacts: With special reference to north american indians and eskimos. Proceedings of the American Philosophical Society 96, 452–463 (1952)
19. Takenobu, T., Sornlertlamvanich, V., Charoenporn, T., Calzolari, N., Monachini, M., Soria, C., Huang, C.-R., YingJu, X., Hao, Y., Prevot, L., Kiyoaki, S.: Infrastructure for standardization of asian language resources. In: Proceedings of ACL-COLING (2006)
20. Vossen, P. (ed.): EuroWordNet: a multilingual database with lexical semantic networks. Kluwer Academic Publishers, Dordrecht (1998)
21. Vossen, P.: Eurowordnet: a multilingual database of autonomous and language-specific wordnets connected via an inter-lingual-index. International Journal of Lexicography 17/2 (2004)
22. Zhang, H., Huang, C.-R., Yu, S.: Distributional consistency: A general method for defining a core lexicon. In: Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC2004) (2004)

Appendix: The Swadesh List

i thou he we you they this that here there who what where when how not all many some few other one two three four five big long wide thick heavy small short narrow thin woman man human child wife husband mother father animal fish bird dog louse snake worm tree forest stick fruit seed leaf root bark flower grass rope skin meat blood bone fat egg horn tail feather hair head ear eye nose mouth tooth tongue fingernail foot leg knee hand wing belly guts neck back breast heart liver drink eat bite suck spit vomit blow breathe laugh see hear know think smell fear sleep live die kill fight hunt hit cut split stab scratch dig swim fly walk come lie sit stand turn fall give hold squeeze rub wash wipe pull push throw tie sew count say sing play float flow freeze swell sun moon star water rain river lake sea salt stone sand dust earth cloud fog sky wind snow ice smoke fire ashes burn road mountain red green yellow white black night day year warm cold full new old good bad rotten dirty straight round sharp dull smooth wet dry correct near far right left at in with and if because name

Conceptual Framework of an Upper Ontology for Describing Linguistic Services

Yoshihiko Hayashi

Graduate School of Language and Culture, Osaka University
1-8 Machikaneyama-cho, Toyonaka-shi, Osaka 560-0043, Japan
hayashi@lang.osaka-u.ac.jp

Abstract. This paper presents a conceptual framework of an upper level ontology for describing linguistic web services. Such an ontology is urgently required to develop an efficient language infrastructure on which a tailored linguistic service can be realized by combining existing general linguistic services and/or recently developed community-based language resources. The ontology will serve as a semantic foundation for the descriptions of the technical components that are inevitably referred to in composite service composition. It will also facilitate the wrapper generation processes that are unavoidable when a linguistic service is incorporated into the language infrastructure. The proposed ontology states that processing resources that may utilize associated language resources, such as lexicons or corpora, are in themselves, a linguistic service. This paper also proposes a taxonomy of processing resources and static language resources and develops a sub-ontology for abstract linguistic objects, such as meaning, expression, and description.

Keywords: service ontology, linguistic service, language resource, processing resource, NLP tool, language infrastructure.

1 Introduction

Several types of linguistic services are currently available on the Web, including text translation and dictionaries. A variety of natural language processing (NLP) tools is also available and public. These tools can be converted into so-called web services, provided that appropriate programs for abstracting the details of each NLP tool are also available. In addition to these, a number of community-based language resources targeting particular domains of application have been developed, and some of them are ready for dissemination. A composite linguistic service tailored to a particular user's requirements would be composable, if there were an infrastructure on which elemental linguistic services, such as NLP tools, and associated language resources could be efficiently combined. The *Language Grid*[1] was proposed as such a language infrastructure [10], and has been thoroughly researched and developed [13].

Ishida [10] argues that the language grid has two dimensions: 'horizontal' and 'vertical.' The horizontal dimension concerns useful combinations of existing

[1] http://langrid.nict.go.jp/

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 31–45, 2007.

linguistic services/language resources. The vertical dimension is associated with specific intercultural communications observed in a variety of intercultural collaboration activities. Such activity often requires linguistic services/language resources that are fine-tuned to the domain.

A composite linguistic service that combines required components from the two-dimensional space could be composed if there were an infrastructure on which the necessary components could be discovered and efficiently combined. For this to be possible, these components must be properly described, and the semantics of the descriptions must be defined by a mutually understood ontology. Such an ontology will also facilitate the wrapper generation processes that are unavoidable during incorporation of existing general linguistic services or dissemination of recently developed community-based language resources.

The remainder of this paper is organized as follows. Section 2 first discusses possible technical configurations of a linguistic web service, and then introduces the general architecture of the language grid infrastructure. Section 3 summarizes what is necessary for a linguistic service ontology, and sets out desiderata for it. The very top level of the ontology is also described in this section. Some of the core parts of the proposed top level ontology are detailed in Section 4, including the taxonomies of core elements, and the relations among linguistic expressions, meanings, and descriptions. Section 5 summarizes the works relevant to the topic. Concluding remarks are presented in section 6.

2 Architecture of the Language Infrastructure

Murakami et al. [13] classifies users of the language grid into *language service providers,* who publish their own atomic language services, and l*anguage service users,* who construct and deploy composite language services by combining existing atomic language services. This classification can be somewhat confusing, because the term *user* usually implies 'end user.' Therefore, in this paper, we classify users of the language grid into *linguistic service providers* and *linguistic service end users*. A linguistic service provider disseminates a linguistic service on the language grid, regardless whether the service is atomic or composite. On the other hand, a linguistic service end user consumes the linguistic services deployed on the language grid, presumably by employing computational tools designed and implemented for assisting intercultural communications. This paper mainly discusses issues with respect to the linguistic service providers.

In this paper, we do not consider the issues of intellectual property rights, terms-and-conditions, and pricing policy of software programs or language resources; although we recognize that these are crucial in actually operating a language infrastructure.

2.1 Technical Configuration of a Linguistic Web Service Site

Suppose you are a potential linguistic service provider willing to provide some useful linguistic function on the language infrastructure by operating a linguistic service on your site. The two possible scenarios are summarized below.

- Utilizing an external web-based linguistic service: For now, a number of web-based linguistic services, such as text translation and dictionary access, are available. The average web user can easily access these services with his/her web browsers. You, as a linguistic service provider, can provide a web service version of such an externally existing linguistic service. To do this, you have to implement a delegation program that formulates a query in the syntax designated by the external service and extracts necessary information from the returned HTML code. Given a situation where a variety of linguistic services are available, this could be a way to provide a linguistic service to the language infrastructure. In this case, the program on your site functions as an adaptor that adjusts detailed differences between the external web-based services and the standard interface defined in the infrastructure. We call this type of program a *wrapper*[2].
- Running an NLP function at your site: In the second scenario you run an NLP function, such as text translation (NLP system) or morphological analysis (NLP tool), at your site. You can develop such an NLP function on your own. In this case, your program has to be compliant with the standard interface in the infrastructure; the core NLP function of the program has to be wrapped so that it functions as a standard web service. Alternatively, you can install a third-party NLP tool or NLP system at your site, and wrap it as a web service. You would be able to install a NLP tool that is available on an R&D institute's web site, and an NLP system if you acquired the license from the vendor. This scenario's significant difference from the former one is that you have to provide both a wrapper program, and a computational resource for running the NLP tool/system, regardless of whether it is in-house or developed by a third party. In addition to these NLP functions, we also consider access functions for language resources as a type of linguistic services. Language resources here mean data-only static resources such as lexicons, or corpora. A prominent example of static language resources is a dictionary source data. For example, we can often acquire dictionary source data by downloading or purchasing it. WordNet[3] [6] and EDR [5] Electronic Dictionary are examples of this. A static language resource always has to be equipped with an access function to be utilized in the language infrastructure. The language resource access function also has to be wrapped to be used as a standard web service.

2.2 General Architecture of the Infrastructure

Based on the discussion above, the fundamental technical components in a language infrastructure, particularly in the language grid, are: (a) external web-based services, (b) on-site NLP core functions, (c) static language resources, and (d) wrapper programs.

Figure 1 is a schematic diagram of the general architecture of the language grid. The technical components listed above are deployed as shown in the figure. Computational nodes in the language grid are classified into the following two types as described in [13].

---

[2] Murakami et al. [13] mentions this type of service providers as *wrapper implementers*.

[3] The online search service is provided at http://wordnet.princeton.edu/perl/webwn

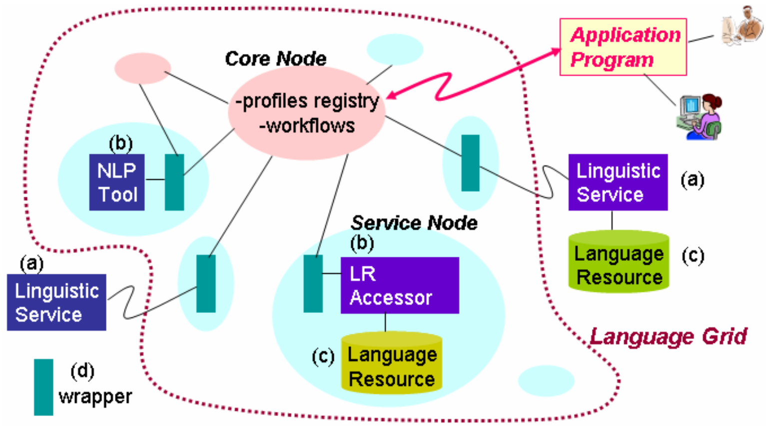


Fig. 1. General architecture of the language grid

- Service nodes: These accommodate atomic linguistic services based on the configuration discussed in the previous subsection. That is, they can provide functionalities of the NLP tool/system running on a node, or they can simply have a wrapper program that consults an external web-based linguistic service.
- Core nodes: These maintain a repository of the known atomic linguistic services and provide service discovery functionality to the possible users/applications. They also maintain a workflow repository for composite linguistic services, and are equipped with a workflow execution engine.

3 Conceptual Framework of the Upper Ontology

3.1 Conceptual Framework

The two necessities for a linguistic service ontology can be summarized as below.

- Composition of composite linguistic services: A composite linguistic service that is tailored to particular user requirements should be composable by humans now and automatically composable in the near future. To make this possible, each technical component has to be described properly before service composition time. For example, if a process pipeline is to be composed of cascading existing linguistic services, the author has to check whether the input constraints for some linguistic services in the pipeline are compatible with the output of the immediately preceding linguistic service. The ontology should provide sufficient vocabulary for stating such input/output constraints. Furthermore, the preconditions, conditions that must hold when a service is being invoked, and the effects of the change of the states brought about by the service invocation of a linguistic service, should also be properly defined for any possible future configuration that utilizes automatic planning.
- Efficient wrapper generation: Wrapper programs have to be properly prepared and deployed in the language infrastructure in order to incorporate both existing

linguistic services and recently developed language resources that are being disseminated. A linguistic service ontology will facilitate such a process. For example, a linguistic service provider who wants to disseminate a new language resource can benefit from the ontology; the relevant part of the taxonomy in the ontology provides guidelines for placing the service in the whole universe of linguistic services. Furthermore, the infrastructure may be able to provide a 'skeleton code' of the wrapper program for a major linguistic service type that has a position in the taxonomy. The service provider would be able to develop his/her own wrapper program code by simply refining the provided skeleton program.

Based on the above discussion about the necessities for a linguistic service ontology, the most important desideratum for it is that it be able to specify the input/output constraints of a linguistic service properly. The taxonomy of linguistic service can be at least partly designed by classifying the input/output constraints. Further desiderata for the ontology can be summarized as follows.

- Vocabulary to describe language resources and linguistic objects: The linguistic service ontology has to provide a vocabulary for describing both processing elements and static language resources that are utilized by linguistic services. Furthermore, it should be able to encode linguistic objects on a variety of levels, such as lexicon entry, linguistic annotation, and even more abstract objects such as those that denote meaning.
- Conforming to technical standards: The linguistic service ontology should be compliant with relevant technical standards, especially for its formal representation, given that the major concern with the language infrastructure is interoperability of the technical components.

3.2 The Top Level of the Ontology

We have developed the upper part of the service ontology so far and have been working on detailing some of its core parts. We have adopted OWL [2] as the working description language for the moment. We use Protégé[4] as the main tool for working with the service ontology.

Figure 2[5] shows the top level of the proposed linguistic service ontology. In the figure and alike in this paper, a rectangle denotes a class, and a directed edge connecting classes indicates the relation between the classes. The label attached to an edge represents the type of the relation; among them, black edges with 'isa' labels depict the class hierarchy.

The topmost concept is `NL_Resource`, which is partitioned into `ProcessingResource`, and `LanguageResource`. Here, as in GATE [1], processing resource refers to programmatic or algorithmic resources, while language resource refers to data-only static resources such as lexicons or corpora. Although the details have not yet been worked out, meta-information for these two types of resources should be available. The innate relation between these two classes is: a

[4] http://protege.stanford.edu/

[5] The diagrams shown in this paper have been automatically produced by the Protégé OntoViz plugin; the detailed arrangement of the elements in a figure was not in the author's control.

processing resource can use language resources. This relationship is specifically introduced to properly define linguistic services that are intended to provide language resource access functions.

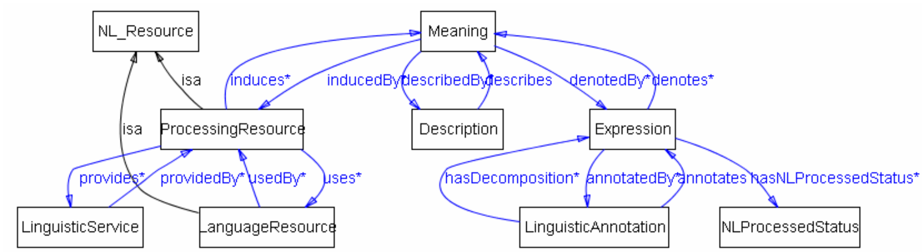


Fig. 2. Top level of linguistic service ontology

As shown in the figure, `LinguisticService` is provided by a processing resource, stressing that any linguistic service is realized by a processing resource, even if its prominent functionality is accessing language resources in response to a user's query. It also has the meta-information for advertising its non-functional descriptions.

The fundamental classes, `Expression`, `Meaning`, and `Description` and the innate relations among them are also illustrated in Fig. 2. These play crucial roles in defining functionalities of some types of processing resources and associated language resources. As shown in Fig. 2, a language expression may denote a meaning, and the meaning can be further explained by a description, especially for human uses. These classes and the relations among them will be detailed in 4.3.

In addition to these, `NLProcessedStatus` and `LinguisticAnnotation` are important classes. NLP status is introduced especially to represent the so-called IOPE (Input-Output-Precondition-Effect) parameters of a linguistic processor, which is a subclass of processing resource. The results of a linguistic analysis are encoded as instances of a linguistic annotation class, and the instances are attached to the instance of the expression class. That is, an expression can be annotated by linguistic annotations. Respecting the recent direction of 'NLP as document annotation' [9, 12], we also consider that any linguistic analysis is additive. Issues with these two classes in the context of our linguistic service ontology are discussed in the next section.

4 Some Details of the Upper Ontology

4.1 Taxonomy of NL Resources

A natural language resource can be either a language resource or a processing resource depending on its nature, as shown in Fig. 2. Figure 3 shows the taxonomy of the language resource class.

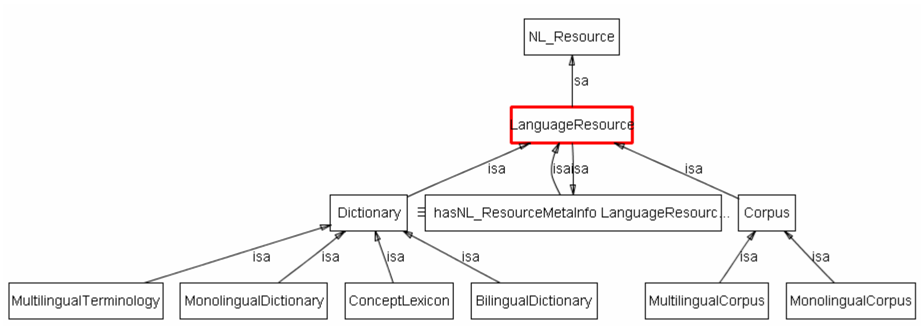


Fig. 3. Taxonomy of language resources

As shown in Fig. 3, the language resource class is partitioned into subclasses for `Corpus` and `Dictionary`. The corpus class is divided into, for the moment, `MultilingualCorpus` and `MonolingualCorpus`. The immediate subclasses of the dictionary class are: (1) `MonolingualDictionary`, (2) `BilingualDictionary`, (3) `MultilingualTerminology`, and (4) `ConceptLexicon`. The major instances of (1) and (2) are so-called machine-readable dictionaries (MRDs). Many of the community-based language resources should fall into (3), including multilingual medical term lists. For subclass (4), we consider the computational concept lexicons, which have a WordNet-like structure[6].

Figure 4 illustrates the top level of the processing resource taxonomy: it consists of the following four subclasses, which take into account the input/output constraints of processing resources and the language resources they utilize.

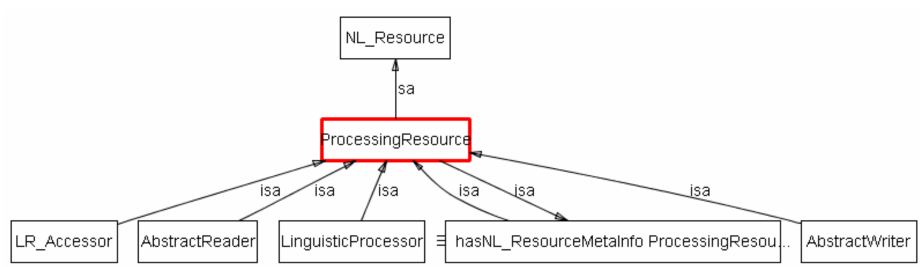


Fig. 4. Taxonomy of processing resources

- `AbstractReader`: This class is introduced to describe computational processes that convert non-textual representation (e.g. speech) into textual representation (character strings). Speech-To-Text systems (STT; speech recognizer) are an obvious subclass of the abstract reader class. For a speech recognizer, its input should have “audio” as its main MIME type.

[6] This means that a whole concept system is organized as a network with a set of synset nodes, and the lexical semantic links that connect the synset nodes.

- AbstractWriter: This class covers a class of computational processes whose processing direction is the opposite of the abstract reader class. This class is introduced to describe text to non-text conversion. Text-To-Speech (TTS) is the representative example. In parallel to the STT, TTS's output should be audio.
- LR_Accessor: This class is introduced to describe language resource access functionalities. It is first partitioned into CorpusAccessor and DictionaryAccessor, depending on the type of language resource it accesses. Figure 5 shows the upward view of the DictionaryAccessor taxonomy. As shown in the figure, the input to a language resource accessor is a query (LR_AccessQuery), and the output of an accessor is a kind of 'dictionary meaning', which is divided into further sub-classes by referring to the taxonomy of a dictionary. Also stated in the figure is that a dictionary accessor uses dictionaries.

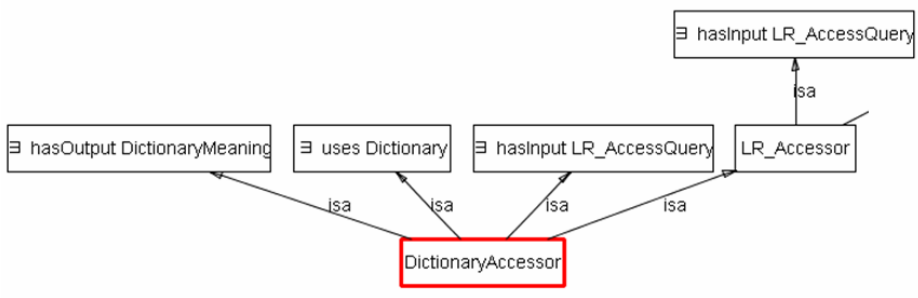


Fig. 5. Taxonomy of dictionary accessors (upward)

- LinguisticProcessor: This class is introduced to represent NLP tools/systems. Figure 6 shows the upward view of the taxonomy of LinguisticProcessor. As clearly stated in the figure, the IOPE parameters of a linguistic processor are constrained to the corresponding classes. The linguistic processor class is first partitioned into Transformer and Analyzer[7]. Figure 7 develops the taxonomy of the transformer subclass. Although it is not shown in Fig.7, the transformer class is subclass of an anonymous class whose hasEffect property is constrained to Transformed class, which is a subclass of NLPStatus class. A transformer somehow transforms the input expression into another expression: the meaning of the output could be either different[8] from the one denoted by the input (MeaningNonPreservingTransformer), or the same (MeaningPreservingTransformer), and the language of the input/output for a meaning preserving transformer could be either different (Translator) or the same (Paraphraser). In Fig. 7, just to give an idea of how a low-level class that will represent a concrete software is defined, classes such as J-to-E_Translator_1 and J-to-J_Paraphrasor_1 are introduced.

---

[7] As seen from the current taxonomy, we have not yet considered any NLP tools that generate an expression from a representation at some level of linguistic representation.

[8] However, we still do not have such a 'meaning transformer'.

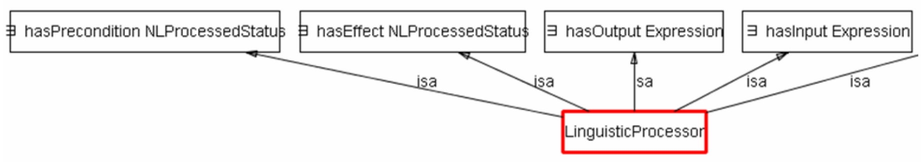


Fig. 6. Taxonomy of linguistic processors (upward)

4.2 Linguistic Analyzer and NLP Status

Figure 8 develops the taxonomy of the analyzer class. While it is not depicted in the figure, the input/output constraints of a linguistic analyzer are specified by the

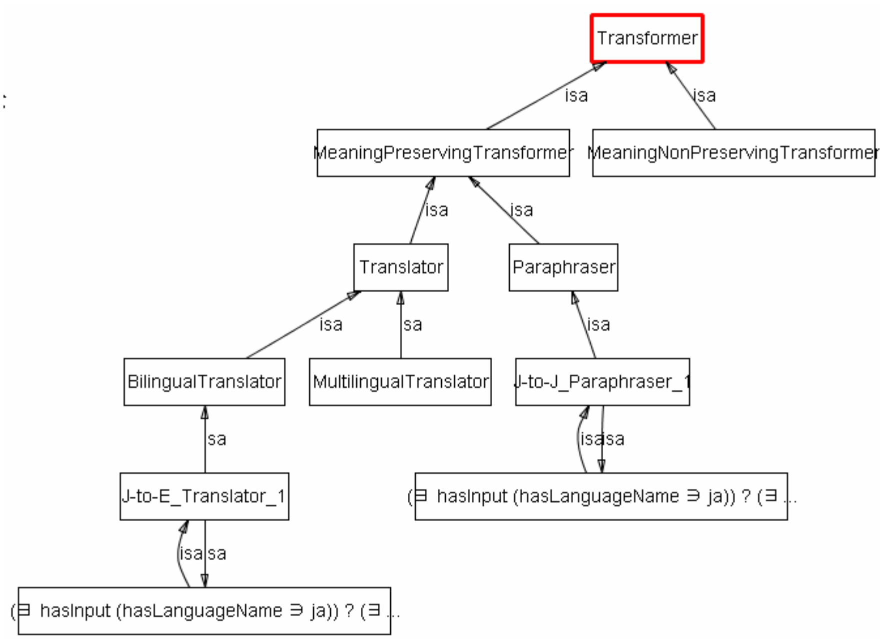


Fig. 7. Taxonomy of transformers

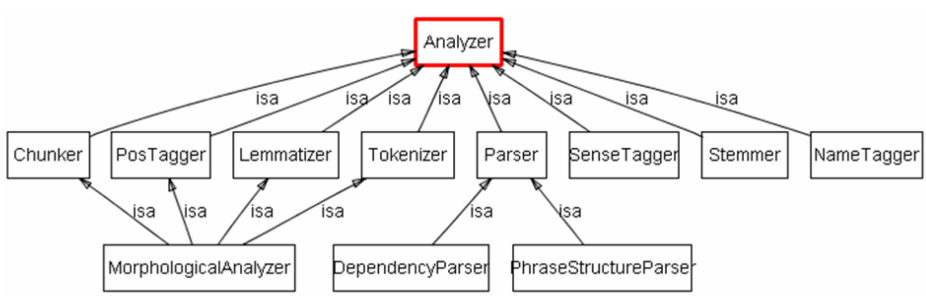


Fig. 8. Taxonomy of linguistic analyzers

`Expression` class, while its precondition/effect parameters are defined by `NLProcessedStatus` class, as inherited from the linguistic processor class. The details are also not shown in this figure, these constraints are further restricted with respect to the taxonomy of the processing resource. We also assume that any linguistic analyzer adds some linguistic annotations to the input, as proposed elsewhere [9, 12]. That is, an analyzer working at a certain linguistic level (or 'depth') adds the corresponding level of annotations to the input. In this sense, any natural language expression can have a layered/multiple linguistic annotation. To make this happen, a linguistic service ontology has to appropriately define a sub-ontology for the linguistic annotations by itself or by incorporating some external standard. The linguistic annotation class is currently much too underspecified to import the international standards that will be specified in the near future. For the moment, it can only accommodate faceted attribute-value pairs and morph-syntactic annotations, such as constituent decompositions and syntactic dependencies.

Figure 9 illustrates our working taxonomy of NLP status. Note that, in this figure, only the portion related to linguistic analyzer is detailed, and the sub-taxonomy rooted by `Transformed` is not shown. Benefits from the NLP status class will be twofold: (1) as a part of the description of a linguistic analyzer, we assign corresponding instances of this class as its precondition/effect parameters, (2) any instance of the expression class can be concisely 'tagged' by instances of the NLP status class, according to how 'deeply' the expression has been linguistically analyzed so far. Essentially, such information can be retrieved from the attached linguistic annotations. In this sense, the NLP status class might be redundant. Tagging an instance of expression in that way, however, is reasonable: we can define the input/output constraints of a linguistic analyzer concisely with this device. This should be effective, even when we adopt, for example, OWL, which is restricted in its expressiveness, as the ontology language.

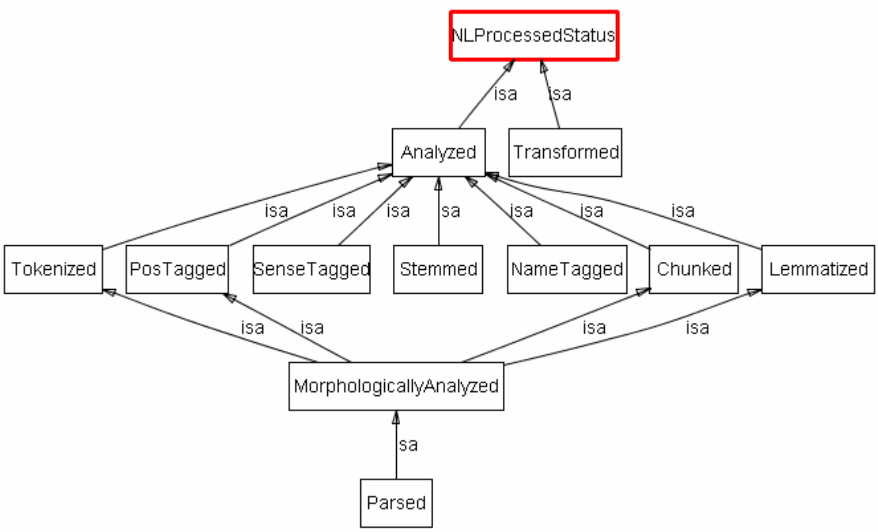


Fig. 9. Taxonomy of NLP states

Each subclass in the taxonomy represents the type or level of a linguistic analysis, and the hierarchy depicts the relations among them. For example, if an expression has been parsed, it would already have been morphologically analyzed because parsing usually requires the input to be morphologically analyzed beforehand. Just as easily imagined, the taxonomy is partly isomorphic to the type for the linguistic analyzers shown in Fig. 8.

Note that the taxonomy in Fig. 9 is only partial and preliminary. The arrangement of the subclasses within the hierarchy may end up being far different, depending on the languages considered[9] and the actual NLP tools that are at hand. For example, the notion of 'chunk' is different from language to language. Also, NLP tools are essentially idiographic. However, if we go too far in this direction constructing a taxonomy would be meaningless and we would forfeit reasonable generalities. We probably need to seek a modest taxonomy that is sufficient for describing actually available NLP tools.

4.3 Expression, Meaning, and Description

Any actual natural language expression of arbitrary length can be represented as an instance of the `Expression` class. We also introduced the `Meaning` class in order to represent some kind of meaning that may be denoted by an expression. Here we are not going into the philosophical question: "what is a meaning?" Instead, we simply assume that any linguistic expression denotes 'some' meaning. Sometimes just the existence of the meaning class instance is sufficient for expressing that "there is surely some meaning, but the details are not clear." This kind of placeholder is useful when we are to represent two different expressions that share the same meaning, which is a characteristic of the meaning preserving transformer. The meaning class will be further refined by introducing subclasses based on an abstract meaning representation formalism, or more concrete information presented in a language resource such as a dictionary or even in more general knowledge resource such as an encyclopedia. Currently we only have `DictionaryMeaning` as a subclass of the meaning class, as discussed in the next subsection.

The `Description` class is introduced to encode additional information that may be useful for a human user to better understand the meaning associated with an expression. That is, an instance of the description class is always attached to an instance of the meaning class. Because, for the moment, we only have the dictionary meaning as a subclass of the meaning class, we only have `DictionaryDescription` as a subclass of the description class. The dictionary meaning class and the dictionary description class are further refined in concordance with the taxonomy of the dictionary class.

Figure 10 summarizes the innate relations among the three classes described above. Note that the dictionary entry class, a subclass of expression class, is introduced for representing a dictionary entry, which is defined as a set of: a canonical written form, a pronunciation, possibly a part-of-speech, and a word sense number.

---

[9] Linguistic analyses can be highly language dependent: this point was made by Thierry Declerck.

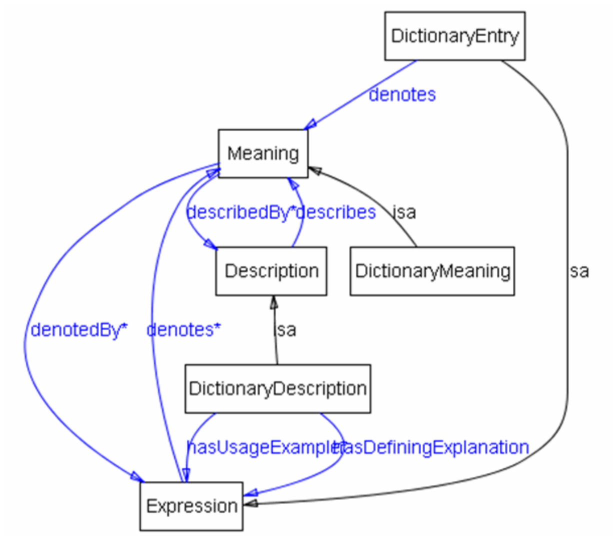


Fig. 10. Relations among `DictionaryEntry`, `DictionaryMeaning`, and `Dictionary Description`

4.4 Dictionary Meaning, Dictionary Description, and Dictionary Model

Following the previous discussion, one question might arise: which information items in a dictionary entry should be considered elements of the meaning class, and which elements of the description class? This question is important, because the issue is closely related to the topic of dictionary modeling [7, 8], which has been considered crucial in the context of lexical resource interoperability.

Our answer to the question is summarized in Table 1, which classifies the relations between the dictionary class and the information items with respect to the meaning/description distinction. The fundamental idea here is that a set of expressions that (seem to) denote a single meaning or a lexicalized concept, is classified as an element of the meaning class, whereas other information items are classified as elements of the description class. For example, as shown in Fig.10, usage example and gloss (defining expression) are given by instance of the expression class. For concept lexicons like WordNet, we also consider the associated lexical semantic relations to be part of the meaning class, rather than part of the description class.

This distinction is in concordance with the three-layered (lemma/meaning/concept) dictionary meta-model recently proposed by Hayashi and Ishida [8]. The lemma layer there is represented by the dictionary entry class here, whereas items associated with meaning and concept layers there are represented by instances of the meaning and the description classes as summarized in Table 1.

Table 1. Classification of dictionary information items

Dictionary Class	Meaning	Description
Monolingual Dictionary	a set of synonyms	explanations, usage examples, etymological information, etc.
Bilingual Dictionary Multilingual Terminology	a set of translations	usage examples, derivational forms, collacations, etc.
Concept Lexicon	synset, lexical semantic relations	gloss, usage examples

5 Related Work

There are several research areas that are relevant to the present topic. A general discussion on language resources and the semantic web that can be found in Declerck et al. [4] stresses the need for a language infrastructure that allows us to share both NLP tools and language resources by means of semantic web technologies.

The study with the approach most similar to our own is Klein and Potter's [12]. They have developed an ontology for NLP services by focusing on two types of users: NLP researchers and users of domain-specific text mining services. While they also presented an upper ontology for processing resources, they focus mainly on NLP analysis tools, including neither NLP application systems nor language resource accessors. Thus they have not thoroughly developed a sub-ontology for language resources or relations among expression, meaning, and description.

This paper introduced processing resource taxonomy, but only partially. LT-World (Language Technology World) [11] portal site[10] has successfully provided a natural language resource taxonomy that offers fine-grained and substantial coverage. In contrast, the ACL NLSR (Natural Language Software Registry)[11] [3] introduces a taxonomy that is relatively coarse, probably for the sake of resource dissemination.

Shifting to the topic of software architecture for NLP tool composition/integration, GATE [1] and the Heart-of-Gold middleware architecture [15] are most relevant to our work. Although both are well-established software architectures and have already achieved some success, both still lack an ontological foundation in the context of distributed semantic web services.

Needless to say, there is an enormous amount of work dedicated to standardization of linguistic objects with respect to interoperability of linguistic resources. Among them, we should refer in particular to LAF (Linguistic Annotation Framework) [9], DCR (Data Category Registry) [14], and LMF (Lexical Markup Framework) [7] as the prominent initiatives.

---

[10] http://www.lt-world.org/
[11] http://registry.dfki.de/

6 Conclusion

This paper proposed an upper ontology for describing linguistic web services, along with language resources being involved. Although the proposed linguistic service ontology successfully defined a number of first class objects and the innate relations among them, it must be further refined and extended through a cycle of descriptive work for specific NLP tools/systems and the language resources that are being considered in actual domains of application. Such descriptive efforts will reveal technical issues still to be addressed. For the moment, it is obvious that we should explore a better and more proper taxonomy for the linguistic analyzers and the corresponding NLP states. As discussed in 4.2, seeking a modest taxonomy that is both sufficiently detailed but also general enough to capture reasonable generalities of classes of linguistic analyzers is a pressing issue. In addition, detailing the schema for representing linguistic annotations and lexicon entries is crucial. In this regard, we should carefully observe the ongoing activities of international standardization, such as LAF/DCR or LMF. We also plan to seek opportunities to cooperate with the relevant initiatives such as Lirics[12] or Clarin[13] to further explore ontology-based language infrastructure architectures.

Acknowledgments. This research was carried out as a part of the “R&D promotion scheme funding international joint research” promoted by NICT. The author expresses his gratitude to Thierry Declerck (DFKI) for his helpful discussions, particularly on the linguistic-related issues. Thanks also go to Yohei Murakami (NICT) for his technical discussions that focused on language grid architecture.

References

1. Cunningham, H., Maynard, D., Bontcheva, K., Tablan, V.: GATE: A Framework and Graphical Development Environment for Robust NLP tools and applications. In: Proc. of ACL2002, pp. 168–175 (2002)
2. Deborah, L.M., Harmelen, F.V. (ed.): OWL Web Ontology Language Overview. W3C Recommendation (2004), http://www.w3.org/TR/2004/REC-owl-features-20040210/
3. Declerck, T.: Extending NLP Tools Repositories for the Interaction with Language Data Resource Repositories. In: Proc. of ACL/EACL Workshop on Sharing Tools and Resources for Research and Education (2001)
4. Declerck, T., Buitelaar, P., Calzolari, N., Lenci, A.: Towards A Language Infrastructure for the Semantic Web. In: Proc. of LREC2004, pp. 1481–1484 (2004)
5. Electronic EDR Dictionary Technical Guide (2003), http://www2.nict.go.jp/kk/e416/EDR/
6. Fellbaum, C. (ed.): WordNet: An Electronic Lexical Database. MIT Press, Cambridge (1998)
7. Francopoulo, G., George, M., Calzolari, N., Monachini, M., Bel, N., Pet, M., Soria, C.: Lexical Markup Framework (LMF). In: Proc. of LREC2006, pp. 233–236 (2006)
8. Hayashi, Y., Ishida, T.: A Dictionary Model for Unifying Machine Readable Dictionaries and Computational Concept Lexicons. In: Proc. of LREC2006, pp. 1–6 (2006)

[12] http://lirics.loria.fr/
[13] http://www.mpi.nl/clarin/

9. Ide, N., Romary, L.: International standard for a linguistic annotation framework. Journal of Natural Language Engineering 10(3-4), 211–225 (2004)
10. Ishida, T.: Language Grid: An Infrastructure for Intercultural Collaboration. In: Proc. of SAINT2006, pp. 96–100 (2006)
11. Jörg, B., Uszkoreit, H.: The Ontology-based Architecture of LT World, a Comprehensive Web Information System for a Science and Technology Discipline. Leitbild Informationskompetenz: Positionen - Praxis - Perspektiven im europäischen Wissensmarkt (2005)
12. Klein, E., Potter, S.: An Ontology for NLP Services. In: Proc. of LREC Workshop on a Registry of Linguistic Data Categories within an Integrated Language Resource Repository Area (2004)
13. Murakami, Y., Ishida, T., Nakaguchi, T.: Infrastructure for Language Service Composition. In: Proc. of SKG2006 (2006)
14. Romary, L. (eds.): Towards a Data Category Registry for ISO TC37 (2004), http://www.tc37sc4.org/new_doc/ISO_TC_37-4_N133_DCR_for_TC37.pdf
15. Schäfer, U.: Middleware for Creating and Combining Multi-dimensional NLP Markup. In: Proc. of the EACL2006 Workshop on Multi-dimensional Markup in NLP (2006)

Formal Ontology of 'Cultures' and 'Ethnic Groups' Based on Type Theory and Functional Programming

Norihiro Ogata

School of Language and Culture, Department of Language and Information,
Osaka University
1-8 Machikaneyama, Toyonaka, Osaka, Japan
ogata@lang.osaka-u.ac.jp

Abstract. This paper will propose *formal ontology* [7] of 'cultures' and 'ethnic groups' by *type theory* with *record types*, which is implemented by *functional programming* such as *Objective Caml* [12]. The structured entities are represented by data structures called `qualia`, `culture`, and `ethnicGroup`, based on fine-grained lexical semantics [16,19]. As a result, we can not only specify the conceptualizations relating to each culture and ethnic group for knowledge sharing, but also we can define functions treating and transforming `qualia`, `culture` and `ethicGroup` by using a type-theoretical framework including implementation by functional programming.

Keywords: Formal Cultural Ontology, Type Theory, Functional Programming, qualia, culture, ethnic group.

1 Introduction

Recent rapid endangerments of ethnic groups, minority languages, and traditional cultures, and great discoveries of archaeological cultures require not only the descriptions or records of them but also the frameworks of knowledge sharing among the researchers across these fields. This paper will propose *formal ontology* [7], a formal specification of conceptualizations for knowledge sharing, of 'cultures' and 'ethnic groups' by type theory with *record types* [17], which is implemented by *functional programming* [12] such as *Objective Caml*.

The basic idea is that cultures and ethnic groups can be regarded as not only simple entities but also structured entities of conceptualizations (traits, facets or patterns) including 'meaning' or 'concepts' depending on each culture and ethnic group, which must be formally specified for knowledge sharing, comparisons of cultures, comparisons of ethnic groups, and comparisons of languages.

For example, in the common Western culture, a *dog* is conceptualized as a pet or companion animal, but in some regions, especially in the *Akha* culture,

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 46–60, 2007.

it is conceptualized as a stock. In some culture, *ice* is conceptualized as frozen water, whereas in another culture, it is conceptualized as melting thing. On the other hand, in many cultures, a *dog* and a *wolf* are differentiated, whereas in the biology both denote the almost same kind *canis lupus*[1] Similarly, *ice*, *water*, *steam* are differentiated in many cultures, whereas in the chemistry they denote the same thing *H2O*. Such an inconsistency between culture-dependent conceptualizations and scientific conceptualizations can be found easily, as Frege or Kripke [10] pointed as *puzzles*: e.g., *The Morning Star*, *The Evening Star* and *Venus*.

Such differences of conceptualizations of a concept among cultures and ethnic groups affects its verbal expressions. Therefore, the WordNet [1] or other generic data base of concepts cannot be directly used for formal specifications of conceptualizations in each culture and ethnic group. On the other hand, we necessitate to identify *correspondent concepts* across cultures and ethnic groups in order to compare the cultures, languages, ethnic groups.

Therefore, the architecture proposed by this paper assigns two types of entities to a concept: simply typed entity and structured entity. The simply typed entities play roles of the ID of a concept, whereas the structured entities specify the conceptualization of a concept in a culture or ethnic group.

The structured entities are represented by a data structure called `qualia`, whose type is a *record type* and are based on fine-grained lexical semantics such as Moravcsik's [16] *aitiational schemes* or Pustejovsky's [19] *qualia structures*. Similarly each culture is modelled as a structured entity of type of a record-type `culture` and each ethnic group as a structured entity of type of a record-type `ethnicGroup`. The data structures of types of `culture` or `ethnicGroup` specify the basic important concepts about specifications of cultures and ethnic groups, which are usually described in ethnographies, such as social organizations, ethics (laws, taboos, matrimonies, etc.), beliefs (religions, myths, medicine, astronomy, meteorology, technology, etc.), livelihoods (food-getting, food/goods productions, dwelling, etc.), eventualities (ceremonies, festivals, etc.), which are accessible by an inference rule of record types. Furthermore, the relating from `culture` to `ethnicGroup` and vice versa is defined and generated by functions between them.

As, a result, by using a type-theoretical framework including implementation by functional programming, we can not only specify the conceptualizations relating to each culture and ethnic group for knowledge sharing, but also define effective operations on them.

In section 2, a type theory with record types and basic types will be proposed. In section 3, a record type `qualia` will be proposed. In section 4, a record type of culture, `culture`, will be defined as a record-type. In section 5, a record type of ethnic group, `ethnicGroup`, will be defined as a record-type. In section 6, generation functions from `ethnicGroup` to `culture` and from `culture` to `ethnicGroup` will be defined by functional programming.

---

[1] Strictly, a dog, called *canis lupus familiaris* in the biology, is a subspecies of a wolf, called *canis lupus* in the biology.

2 A Type Theory for Ontologies of Cultures and Ethnic Groups

2.1 The Basic Types of Simple Entities

Adding to basic types of the usual functional programming such as `int` (integer), `bool` (boolean), `string`, and so on, a basic type `e` (entity), which is introduced into formal semantics of natural language by Montague [14] is introduced, as follow[2]:

Definition 1 (type e)

```
type e = NilE|E of string|K of string|Stage of string * int
|Group of string|Stuff of string|STUFF of string * e
|Part of e * string|Any of string * int|An of string * int
|Artifact of string * e * int|ARTIFACT of string|STAGE of string
|Culture of string|EG of string|Of of e * e|Pl of string;;
```

`NilE` means the undefined entity of `e`, `Any(x,`n`)` means an arbitrary object of `x` whose ID is n, and `An(x,`n`)` means a certain object of `x` whose ID is n. Following the ontologies of Carlson's [4] and Link's [13], `e` has subentities such as *individuals* (e.g., `(E "John")`, i.e., a person *John*), *kinds* (e.g., `(K "dog")`, i.e., a kind *dog*), *stages* (e.g., `Stage("child",201)`, i.e., a *child* whose ID is 201 and `STAGE("child")`, i.e., *child* (of human)), *groups* (e.g., `(Group "the Beatles")`, i.e., a music band called *the Beatles*), *stuffs* (e.g., `(Stuff "water")`, i.e., a matter called *water*; `STUFF("meat",(K "dog"))`, i.e., the meat of dog), *parts* (e.g., `Part((K "dog"),"tail")`, i.e., the part of dog called its *tail*)[3]. Furthermore, to treat conceptualizations relating to cultures and ethnic groups, type `e` has subentities such as *artifacts* (e.g., `Artifact("CD",ARTIFACT("Ballades by Chopin"),304)`, i.e., the CD of the Ballades by Chopin whose ID is 304), *ethnic groups* (e.g., `(EG "Mongol")`, i.e., the ethnic group called *Mongol*), and *cultures* (e.g., `(Culture "Mongol")`, i.e, the culture of the Mongol people).

Adding to type `e`, types such as `dimension` (≈ attributes), `value` (≈ attributes' values), `years`, `period`, `place`, `rel` (relation), `proposition` are defined as basic types, as follows[4]:

`NilProp` denotes the undefined proposition. I explain these definitions simply by examples. For example, `Around(AD 1907)` means "around 1907", `During((AD 1907),(AD 1920))` "during the period from 1907 to 1920", `At(AD 1907)` "on

[2] Henceforth, by the term *definition* or *define*, I mean a program or programming in Objective Caml.

[3] These subentities include eventualities in the sense of Bach [3] and abstract entities in the sense of Asher [2].

[4] These definitions are not exhaustive here.

Definition 2

```
type dimension = Area|Height|Shape|Color;;

type value = Square_meters of int|Centimeters of int
|Meters of int|Square|Round|Oval|Conical|Cylindrical
|Circle|Rectangular|High|Low|Black|Gray;;

type years = BC of int| AD of int|BC_Century of int
|AD_Century of int;;

type period = Around of years|During of years * years
|At of years|Since of years;;

type place = NilP|Geo of string|Mt of string
|River of string|Province of place * string
|State of place * string|Prefecture of place * string
|North of place|South of place|East of place|West of place
|Southeast of place|Northeast of place
|Lower of place|Upper of place|Along of place|Within of e;;

type rel = Rel of string|IV of string * e
|TV of string * e * e|DTV of string * e * e * e;;

type proposition = Gen of rel| GEN of rel * rel
|And of rel * rel|NilProp
|Past of rel|C of e * rel|Neg of proposition
|Obl of proposition|Bel of e * proposition;;
```

1907", `NilP` the undefined place, `(Geo "Asia")` *the Asia*, `(Mt "Fuji")` *Mt. Fuji*, `Province((Place "Thai"),"Chiang Rai")` *the Chiang Rai Province of Thai*, `Rel "xxx"` *a 0-ary relation called "xxx"*. For the other example, I enumerate their examples, as follows:

- `IV("burn",(Stuff("meat",(K "dog"))))` (a unary relation "meat of dogs burn"),
- `TV("eat",(EG "Akha"),(Stuff("meat",(K "dog"))))` (a binary relation "the Akha people eat meat of dogs"),
- `DTV("sell",(EG "Akha"),(Stuff("meat",(K "dog"))),(EG "Vietnamese"))` (a ternary relation "the Akha people sell meat of dogs to Vietnamese"),
- `Gen(TV("eat",(EG "Akha"),(Stuff("meat",(K "dog")))))` (a generic proposition such that generally the Akha people eat meat of dogs),
- `GEN(And(IV("die", Any("Akha person",1)),(IV("wealthy", Any("Akha person",1)))),IV("be sacrificed",Any("Akha buffalo",2)))` (a

generic proposition such that if a wealthy Akha person has died, a buffalo is sacrificed),

- `Past(TV("flee",(EG "Akha"),(E "Burma"),(E "Thailand")))` (a singular proposition such that the Akha people fled from Burma to Thailand),
- `C((Culture "Akha"),TV("perform",(EG "Akha"),An("ritual of death of the Akha",6)))` (a proposition such that the Akha culture brings it about that the Akha people perform the ritual of a death of the Akha, i.e., in the Akha culture there is a ritual of a death),
- `Obl(Neg(TV("marry",Any((E "man"),1),Any((Of((E "sister"),Any ((E "man"),1))),2))))` (a deontic proposition such that it is obligatorily that any man does not marry any of his sisters),
- `Bel(Any((E "Akha person"),1),IV("exist",(Pl "Akha spirit")))` (a doxastic proposition such that any Akha person believes that there exist spirits).

2.2 Derived Types

Adding to the simple types defined in the previous subsection, using the following type constructors:

- `t1 -> t2` (function type from type `t1` to `t2`)
- `t1 * t2` (product type of `t1` and `t2`)
- `t list` (list type of objects of type `t`)
- `{l1:t1; ... ; l_n:t_n}` (record type which consists of labels `l1, ..., l_n` which are associated with objects of types `t1,...,t_n`; e.g., `{x = "Akha"; y = 7}:{x:string; y:int}`; if `a = {x = "Akha"; y = 7}` then `a.x="Akha"` and `a.y=7`.)

I define the following complex types:

Definition 3

```
type regions = (period * (place list)) list;;
type mereologia = (regions * (rel list)) list;;
type synonymia = (string * rel ) list;;
type taxonomia = e list;;
type extension = {referent:e; index: regions};;
```

These types are explained as follows:

- `regions` is the type of lists of spatio-temporal regions: e.g., `((Around (AD_Century 20)), [Province((Geo "Thai"),"Chiang Rai")])`, which denotes the Chaing Rai Province and the Chiang Mai Province of Thai in the 20th century,

- `mereologia` is the type of spatio-temporal mereology which is modeled by the lists of the pairs of `regions` and `rel`: e.g., `[([((Around(AD_Century 12)),[(South (Geo "Yunnan"))])],IV("subgroup of",(EG "Adzoh Akha")))]`, which means that in the twelfth century the Adoh Akha tribe was a subgroup of the Akha in southern Yunnnan.
- `synomymia` is the type of synonymy: e.g., `[("Hani",IV("called by",(EG "Chinese"))); ("Akha",IV("called by",(EG "Akha")))]`, which means that Akha is called "Hani" by Chinese and Akha's autonym is Akha,
- `taxonomia` is the type of lists of `e` but their elements must be of form of `(K x)`, i.e., the list of kinds, which expresses the information about taxonomy,
- `extension` is the record type whose lables are `referent` (type `e`) and `index` (type `regions`): e.g., {`referent = (EG "Akha"); index = [((Around (AD_ Century 20)), [Province((Geo "Thai"),"Chiang Rai"); Province ((Geo "Thai"),"Chaing Mai")])]`}.

3 Type `qualia`

The basic idea of type `qualia` is a pair of *extension* and *intension*.[5] , which is a structured version of Parsons' [18] *correlates*, i.e., a set of properties which the object has. Furthermore, the correlates in an object of type `qualia` are structured by ideas of Moravcsik's [15,16] *aitiational schemes* and its successor, Pustejovsky's [19] *qualia structure*. In this paper, I define type `qualia` as a record type whose labels consist of the following lables:

- `extent` (type `extension`),
- `taxonomy` (type `taxonomia`),
- `mereology` which almost amounts to Pustejovsky's Constitutive (*the relation between an object and its constituent parts*) and Moravcsik's *constituency* (type `mereologia`),
- `synonymy` (type `synonymia`),
- `activities` (type `proposition list`), which specifies the information about the object's usual activity: e.g., 'running' for 'vehicles', 'breathing' for 'animals', 'livelihood' for 'ethnic groups', and so on,
- `genesis` (type `proposition list`), which almost amounts to Pustejovsky's Agentive,
- `effects` (type `proposition list`), which almost amounts to Moravcsik's *causal power*.
- `results` (type `e list`), which is the list of resuliting states or (by)products,
- `disappearance` (type `proposition list`), which is a list of how to disappear: e.g., 'death' for 'life', 'destruction' for 'building' and so on.
- `quality` (type `(dimension * value) list`), which is a kind of *attribute-value* structure.

Therefore, type `qualia` is defined as follows:

[5] In *Formal Concept Analysis* [5], each concept is formalized as a pair of its extension and its intension.

Definition 4

```
type qualia = {extent:extension;
activities:proposition list;
disappearance:proposition list;
effects:proposition list;
functions:proposition list;
genesis:proposition list;
quality:(dimension * value) list;
results:e list;
mereology:mereologia;
synonymy:synonymia;
taxonomy:taxonomia};;
```

In the following example, *dog* in the sense of the Modern Western cultures and in the sense of the Akha culture (where some specifications are inherited from the other qualia by function `inherit` whose definition is omitted here; the synonymy and quality are omitted by specifying them as the empty list `[]`):

Example 1

```
let dog = {extent = {referent = (K "dog");
   index = ((Since(BC 13000)),[(Geo "Land of the Earth")])};
 mereology = (inherit (mammal.mereology) (K "dog"));
 synonymy = [];
 taxonomy = [(K "mammal"); (K "pet")];
 functions = (inherit (pet.functions) (K "dog"));
 activities = ((inherit (mammal.activities) (K "dog"))
                        @ [Gen(IV("bark",(K "dog")))];
 genesis = (inherit (mammal.genesis) (K "dog"))
          @ [Gen(TV("breed",(E "Western people"),(K "dog")))];
 effects = (inherit (pet.effects) (K "dog"));
 results = (inherit (mammal.results) (K "dog"))
                   @ [(STAGE "puppy")];
 disappearance = (inherit (mammal.disappearance) (K "dog"));
 quality = []}};;
```

Example 2

```
let dog_of_Akha = {extent = {referent = (K "dog of Akha");
   index = akha.extent.index};
 mereology = (inherit (stock.mereology) (K "dog of Akha"));
 synonymy = [("a kui",TV("called by",(EG "Akha")))];
 taxonomy = [(K "stock")];
 functions = (inherit (stock.functions) (K "dog of Akha"));
 activities = ((inherit (stock.activities) (K "dog of Akha"))
           @ [Gen(IV("bark",(K "dog of Akha")))];
 genesis = (inherit (stock.genesis) (K "dog of Akha"));
 effects = (inherit (stock.effects) (K "dog of Akha"));
 results = (inherit (stock.results) (K "dog of Akha"));
 disappearance
  = (inherit (stock.disappearance) (K "dog of Akha"));
 quality = []}};;
```

4 Type `culture`

Although the term *culture* has many definitions[6], Hoult’s [9] articulation of *culture* into the following four elements can be regarded as a subcategorization of aitiational scheme or qualia (structures):

– *value*: ideas about what in life seems important. This amounts to `taxonomy` of qualia or *structure* of aitiational scheme, since both of *value* and `taxonomy` specify the essence or nature of culture (or objects).
– *norms*: expectations of how people will behave in various situations realized as *laws*. This amounts to `functions` of qualia and aitiational scheme, since norms of culture can be regarded as functions such that people belonging to the culture keep or believe ethics, laws, taboos, worldview, philosophy, metaphysics, technology, beliefs, and other knowledge of the culture.
– *institutions*: the structure of a society within which values and norms are transmitted. This amounts to `mereology` of qualia or *constituency* of aitiational scheme.
– *artifacts*: products derived from norms and values, which reflects the material aspects of the culture, i.e., *its material culture*. This amounts to `results` of qualia or *causal power*.

This specification of *culture* is useful to define type `culture` as a record type; this specifications paraphrasing “In culture c, people do something” as “culture c causes the people of c to do something,” and by this paraphrasing we can connect each culture with its specification.

Ethnographies can be regarded as the primal resources of the specifications of each culture. This fact implies that Hoult’s [9] articulation of *culture* is too poor. At least, we need describe the following items which are usually reported by ethnographies and other reports about cultures[7]:

– `c_extent` (type `extension`)
– `c_synonymy` (type `synonymia`)
– `c_taxonomy` (type `taxonomia`)
– `ethnozoology` (type `(string * e) list`): the list of animals constructed from the viewpoint of a culture,
– `ethnobotany` (type `(string * e) list`): the list of plants constructed from the viewpoint of a culture,
– `ethnopharmacology` (type `(string * e) list`): the list of drugs constructed from the viewpoint of a culture,
– `ethnocosmology` (type `(string * e) list`): the list of cosmological entities constructed from the viewpoint of a culture,

[6] According to Kroeber and Kluckhohn [11], *culture* has more than 200 definitions.
[7] By the restriction of the Objective Caml such that each label defined in a record type cannot be used to define the other record type, as for `extent`, `synonymy`, and `taxonomy` are affixed with `c_`.

- products (type (string * e) list): the list of products constructed from the viewpoint of a culture,
- tools (type (string * e) list): the list of tools constructed from the viewpoint of a culture,
- arts (type (string * e) list): the list of arts constructed from the viewpoint of a culture,
- crops (type (string * e) list): the list of crops constructed from the viewpoint of a culture,
- stocks (type (string * e) list): the list of stocks constructed from the viewpoint of a culture,
- foods (type (string * e) list): the list of foods constructed from the viewpoint of a culture,
- gatherables (type (string * e) list): the list of things which people usually gather for living, including fish, fruit, animals, plants, minerals, and other natural resources, constructed from the viewpoint of a culture,
- buildings (type (string * e) list): the list of buildings constructed from the viewpoint of a culture,
- relics (type (string * e) list): the list of relics constructed from the viewpoint of a culture,
- wearings (type (string * e) list): the list of wearings constructed from the viewpoint of a culture,
- social_roles (type (string * e) list): the list of social roles such as leader, head, priest, servant and so on constructed from the viewpoint of a culture,
- kinship (type (string * e) list): the list of kinship terms constructed from the viewpoint of a culture,
- events (type (string * e) list): the list of important events such as ceremonies, festivals, feasts and so on constructed from the viewpoint of a culture,
- ontology (type (string * e) list): the list of basic objects such as food, house, wear, people, person, man, woman, ceremony, and so on constructed from the viewpoint of a culture,
- livelihood (type (string * proposition) list): the list of propositions expressing everyday's activities constructed from the viewpoint of a culture,
- beliefs (type (string * proposition) list): the list of propositions expressing beliefs believed in a culture,
- ethics (type (string * proposition) list): the list of propositions expressing ethics believed in a culture,
- ethnoiatry (type (string * proposition) list): the list of propositions expressing beliefs about medicine or healing believed in a culture,
- ethnotechnology (type (string * proposition) list): the list of propositions expressing technology believed in a culture,
- sociology (type (string * proposition) list): the list of propositions expressing the social system believed in a culture,
- stories (type (string * (proposition list)) list): the list of lists of propositions expressing stories such as legends, myth, rumors believed in a culture,

– ethnotopography: (type (string * place) list): the list of places constructed from the viewpoint of a culture.

For example, the Akha culture is very roughly defined as follows:

Example 3

```
let akha = {c_extent = {referent = (Culture "Akha");
index = [((Around (AD_Century 20)),
[Province((Geo "Thai"),"Chiang Rai")])]};
c_synonymy =[]; c_taxonomy =[(Culture "")];
ethnozoology = [("fish",(K "fish of Akha"));
     ("dog",(K "dog of Akha"))];
ethnobotany = [("dry rice",(K "dry rice of Akha"));
      ("bamboo",(K "bamboo of Akha"))];
ethnopharmacology =[]; ethnocosmology = [];
products = [("dry rice",(K "dry rice of Akha"));
      ("basket",(K "basket of Akha"))];
crops = [("dry rice",(K "dry rice of Akha"))]; social_roles = [];
stocks = [("dog",(K "dog of Akha"))]; sociology = [];
gatherables = [("fish",(K "fish of Akha"))];
stories = []; kinship = []; ethnotechnology = [];
tools = [("basket",(K "basket of Akha"))]; arts = [];
ethnotopography = []; beliefs = []; ethics = [];  ethnoiatry = [];
livelihood =
[C((Culture "Akha"),Gen(TV("weave",Any("Akha woman",1),
                                Any("Akha cloth",2))))];
events = [];
foods = [("dog meat",(STUFF("meat",(E "dog of Akha"))))];
ontology = [("house",(K "house of Akha")); ("food",(K "food of Akha"))];
buildings = [("house",(K "house of Akha"))]; relics = [];
wearings = [("cloth",(K "cloth of Akha")]};;
```

In this example, each entity of type e onlyt plays the role of the ID of the entity and there is no association with specifications of the entity such as qualia. Therefore, I define function from e to qualia roughly, as follows (where nil is the empty qualia):

Definition 5

```
let q x = match x with
   (K "dog") -> dog
  |(K "dog of Akha") -> dog_of_Akha
  | _ -> nil;;
```

Furthermore, I define infix function culture to e, which can use for expressing “dogs of Akha”, as follows (where List.assoc is a predefined functions of Objective Caml such that List.mem x y = z if (x,z) is in y else error):

Definition 6

```
let (//) x y =
(List.assoc y (x.ethnozoology)) || (List.assoc y (x.ethnobotany))
|| (List.assoc y (x.ethnopharmacology))
|| (List.assoc y (x.ethnocosmology))
|| (List.assoc y (x.products)) || (List.assoc y (x.crops))
|| (List.assoc y (x.stocks)) || (List.assoc y (x.gatherables))
|| (List.assoc y (x.tools)) || (List.assoc y (x.arts))
|| (List.assoc y (x.foods)) || (List.assoc y (x.ontology))
|| (List.assoc y (x.buildings)) || (List.assoc y (x.relics))
|| (List.assoc y (x.wearings)) || (List.assoc y (x.kinship))
|| (List.assoc y (x.events));;
```

By function `//`, "dogs of Akha" is represented by `akha // "dog"`, as follows:

Example 4

```
# akha // "dog";;
-: e =  K "dog of Akha"
```

We can also call `qualia` of 'dogs of Akha' by the command `q(akha // "dog");;`.

5 Type `ethnicGroup`

In usual ethnographies, cultures and ethnic groups are not clearly distinguished from each other. This fact means that each ethnic group and its culture share their basic characterisitcs with each other. However, each ethnic group's *extension*, *taxonomy*, *synonymy*, and *mereology* (i.e. subgrouping of each ethnic group) are different from its culture's ones, and furthermore each ethnic groups has *population* and *history* (imigrations, wars, and so on). From this point of view, I define a record type `ethnicGroup` as follows (by the similar reason of definition of `culture`, the correspondent labels are affixed by `e_`):

Definition 7

```
type populations = (years * string * int) list;;
type historia = (years * (proposition list)) list;;
```

Definition 8

```
type ethnicGroup = {e_extent:extension;
e_synonymy: synonymia; e_taxonomy: taxonomia;
population:populations; history: historia;
e_ethnozoology: (string * e) list;
e_ethnobotany: (string * e) list;
e_ethnopharmacology: (string * e) list;
e_ethnocosmology: (string * e) list;
e_products: (string * e) list; e_crops: (string * e) list;
e_social_roles: (string * e) list; e_stocks: (string * e) list;
e_gatherables: (string * e) list; e_kinship: (string * e) list;
e_tools: (string * e) list; e_arts: (string * e) list;
e_foods: (string * e) list; e_ontology: (string * e) list;
e_buildings: (string * e) list; e_relics: (string * e) list;
e_wearings: (string * e) list;
e_events: (string * e) list;
e_beliefs: (string * proposition) list;
e_sociology:(string * proposition) list;
e_stories:(string * (proposition list)) list;
e_ethnotechnology:(string * proposition) list;
e_livelihood:(string * proposition) list;
e_ethics:(string * proposition) list;
e_ethnoiatry:(string * proposition) list;
e_ethnotopography: (string * place) list;};;
```

Example 4

```
let akha_people = {e_extent = {referent = (EG "Akha");
index = [((Around (AD_Century 20)),
[Province((Geo "Thai"),"Chiang Rai")])]};
e_synonymy =[]; e_taxonomy =[(EG "")];
population = [((AD 2000),"total",49903)];
history = [((AD 1903),Past(TV("establish",An("village",1),Geo("Thai"))))];
e_ethnozoology = [("fish",(K "fish of Akha"));
     ("dog",(K "dog of Akha"))];
e_ethnobotany = [("dry rice",(K "dry rice of Akha"));
      ("bamboo",(K "bamboo of Akha"))];
e_ethnopharmacology =[]; ethnocosmology = [];
e_products = [("dry rice",(K "dry rice of Akha"));
      ("basket",(K "basket of Akha"))];
e_crops = [("dry rice",(K "dry rice of Akha"))]; social_roles = [];
e_stocks = [("dog",(K "dog of Akha"))]; sociology = [];
e_gatherables = [("fish",(K "fish of Akha"))];
e_stories = []; e_kinship = []; e_ethnotechnology = [];
e_tools = [("basket",(K "basket of Akha"))]; e_arts = [];
e_ethnotopography = []; e_beliefs = []; e_ethics = [];
 e_ethnoiatry = []; e_events = [];
e_livelihood = [Gen(TV("weave",Any("Akha woman",1),Any("Akha cloth",2)))];
e_foods = [("dog meat",(STUFF("meat",(E "dog of Akha"))))];
e_ontology = [("house",(K "house of Akha")); ("food",(K "food of Akha"))];
e_buildings = [("house",(K "house of Akha"))]; relics = [];
e_wearings = [("cloth",(K "cloth of Akha")]};;
```

6 Generation of culture to ethnicGroup and Generation of ethnicGroup to culture

As we have seen, `culture` and `ethnicGroup` share many labels. Therefore, we can define generation functions from `culture` to `ethnicGroup` (`c2e`) and from `ethnicGroup` to `culture` (`e2c`), as follows:

Definition 8

```
let c2e0 x = match x with (Culture(y)) -> (EG(y)) | _ -> NilE;;

let c2eP x =
match x with (C(Culture(y),p) -> p | _ -> x;;

let c2e(c,x,y,z,w) =
{e_extent={referent = c2e0(c.c_extent.referent);
index = (c.c_extent.index)};
e_synonymy = x; e_taxonomy = [(EG "")];
e_mereology = y; population = z;
history = w;
e_ethnozoology = (c.ethnozoology);
e_ethnobotany = (c.ethnobotany);
e_ethnopharmacology = (c.ethnopharmacology);
e_ethnocosmology = (c.ethnocosmology);
e_products = (c.products); e_crops = (c.crops);
e_social_roles = (c.social_roles); e_stocks = (c.stocks);
e_gatherables = (c.gatherables); e_kinship = (c.kinship);
e_tools = (c.tools);e_arts = (c.arts);
e_foods = (c.foods); e_ontology = (c.ontology);
e_buildings = (c.buildings);
e_relics = (List.map c2eP (c.relics));
e_wearings = (c.wearings);
e_beliefs = (List.map c2eP (c.beliefs));
e_sociology = (List.map c2eP (c.sociology));
e_stories = (c.stories);
e_ethnotechnology = (List.map c2eP (c.ethnotechnology));
e_livelihood = (List.map c2eP (c.livelihood));
e_ethics = (List.map c2eP (c.ethics));
e_ethnoiatry = (List.map c2eP (c.ethnoiatry));
e_ethnotopography = (c.ethnotopography)};;
```

Definition 8

```
let e2c0 x = match x with (EG(x)) -> (Culture(x)) | _ -> NilE;;
let e2cP e p = C((e2c0 (e.e_extent.referent)),p);;

let e2c(e,x,y) = {
c_extent={referent = e2c0(e.e_extent.referent);
index = (e.e_extent.index)};
c_synonymy = x; c_taxonomy = [(Culture "")];
c_mereology = y;
ethnozoology = (c.ethnozoology);
ethnobotany = (c.ethnobotany);
ethnopharmacology = (c.ethnopharmacology);
ethnocosmology = (c.ethnocosmology);
products = (e.e_products); crops = (e.e_crops);
social_roles = (e.e_social_roles); stocks = (e.e_stocks);
gatherables = (e.e_gatherables); kinship = (e.e_kinship);
tools = (e.e_tools); arts = (e.e_arts);
events = (e.e_events);
foods = (e.e_foods); ontology = (e.e_ontology);
buildings = (e.e_buildings); relics = (e.e_relics);
wearings = (e.e_wearings);
beliefs = (List.map (e2cP e) (e.e_beliefs));
sociology = (List.map (e2cP e) (e.e_sociology));
stories = (e.e_stories);
ethnotechnology = (List.map (e2cP e) (e.e_ethnotechnology));
livelihood = (List.map (e2cP e) (e.e_livelihood));
ethics = (List.map (e2cP e) (e.e_ethics));
ethnoiatry = (List.map (e2cP e) (e.e_ethnoiatry));
ethnotopography = (e.e_ethnotopography)};;
```

7 Conclusion

As we have seen, this paper has proposed formal ontologies of cultures and ethnic groups by type theory with record types and subtyping and functional programming, based on the idea of Moravcsik and Pustejovsky's idea of structured lexicons. That is, this paper has proposed a type hierarchy of simple entities for specifying the concepts relating to cultures and ethnic groups and record types called qualia which specify the conceptualization of the entities. Furthermore, record type `culture` and record type `ethnicGroup` have been proposed and generation functions from `culture` to `ethnicGroup` and from `ethnicGroup` to `culture` are also defined by functional programming. Thus, such a proposal of formal ontology of cultures and ethnic groups based on type theory and functional programming not only provides the framework of knowledge sharing of these research fields, but also makes the concepts in these fields computationally manipulatable. As, a result, by using a type-theoretical framework including implementation by functional programming, we can not only specify the conceptualizations relating to each culture and ethnic group for knowledge sharing, but also define effective operations on them and provide a theoretical basis of designing an XML-based *markup language* of the concepts used in cultures or ethnic groups (cf. Objective Caml can treat XML type-theoretically via XDuce [8]). Although I have an idea of an answer to the question of who makes and how we make this type of formal ontology and how we makes, for example, exploiting information extraction from ethnographies and online information such as *Ethnologue* [6] or building a Wiki site, in the next work I will investigate them.

References

1. WordNet: http://wordnet.princeton.edu/
2. Asher, N.: Reference to Abstract Objects in Discourse. Kluwer Academic Publishers, Dordrecht (1993)
3. Bach, E.: The algebra of events. Linguistics and Philosophy 9, 5–15 (1986)
4. Carlson, G.N.: Reference to Kinds in English. PhD thesis, University of Massachusetts, 1977. Published by Garland Publishing, New York (1980)
5. Ganter, B., Wille, R.: Formal Concept Analysis. Springer, Heidelberg (1999)
6. Jr., Gordon, Raymond, G.: Ethnologue: Languages of the world, 15th edn. (2005), http://www.ethnologue.com/
7. Gruber, T.R.: Toward principles for the design of ontologies used for knowledge sharing. International Journal of Human and Computer Science 43(5/6), 907–928 (1995)
8. Hosoya, H., Pierce, B.C.: XDuce: A typed XML processing language. ACM Transactions on Internet Technology 3(2), 117–148 (2003)
9. Hoult, T.F.: Dictionary of modern sociology, Littlefield, Adams, Totowa (1969)
10. Kripke, S.A.: A puzzle about belief. In: Margalit, A. (ed.) Meaning and Use, pp. 239–283. Reidel, Dordrecht (1979)
11. Kroeber, A.L., Kluckhohn, C.: Culture: a critical review of concepts and definitions, Peabody Museum, Cambridge (1952)
12. Leroy, X., Doligez, D., Garrigue, J., Rémy, D., Vouillon, J.: The Objective Caml system, release 3.09: Documentation and user's manual (2005)
13. Link, G.: Algebraic Semantics in Language and Philosophy. CSLI Publication, Stanford (1998)
14. Montague, R.: Formal Philosophy: Selected Papers of Richard Montague (ed. by Richmond Thomason). Yale University Press, New Haven (1974)
15. Moravcsik, J.M.E.: Aitia as generative factor in Aristotile's philisophy. Dialogue 14, 622–636 (1975)
16. Moravcsik, J.M.E.: How do words get their meanings? Journal of Philosophy 78(1), 5–24 (1981)
17. Ohori, A.: A polymorphic record calculus and its compilation. ACM Transactions on Programming Languages and Systems 17(6), 844–895 (1995)
18. Parsons, T.: A prolegomenon to Meinongian semantics. Journal of Philosophy 71(16), 561–580 (1974)
19. Pustejovsky, J.: Generative Lexicon. MIT Press, Cambridge (1995)

Capturing a Conceptual Model for Intercultural Collaboration in Copyright Management
(A Domain Ontology of International Copyright Laws)

Wenhuan Lu and Mitsuru Ikeda

Graduate School of Knowledge Science,
Japan Advanced Institute of Science and Technology,
1-1 Asahidai, Nomi, Ishikawa, 923-1292, Japan
{wenhuan,ikeda}@jaist.ac.jp

Abstract. Getting intercultural collaborators to understand critically about the commonalities and differences between local and foreign copyright legal knowledge is important in order to facilitate cross-cultural interaction, learning and appreciation for worldwide resource-sharing. However it is difficult due to copyright knowledge representation nation-dependent. In this paper, starting from intention theory, the intention behind law is modeled, which can be used to reflect the essential meanings of copyright law articles derived from different countries, as well as the approach in which we can clarify the commonalities and differences of international copyright laws. A conceptualization system has been developed to provide semantic level representation for processing, modeling copyright knowledge and maintaining consistency using ontological technology.

Keywords: law article, intention model, intercultural collaboration.

1 Introduction and Related Work

Resource-sharing worldwide offers a great opportunity for widespread, quick and efficient utilization of creative works. One of the major challenges to realize healthy and meaningful resource-sharing today is the management of copyrighted works, and legal knowledge management regarding different national copyright laws. The management of copyright legal knowledge derived from different countries must begin with a fundamental understanding of the general principles of copyright law. Common understanding of international legal knowledge and a legal knowledge management platform based on such understanding are necessary to build a healthy and legitimate resource-sharing environment.

The difficulties encountered in management and organization of multinational legal knowledge are great. In the field of copyright law, especially concerning the diverse national codification (from law article level), it is necessary to clarify the commonalities and differences of articles, and represent them in an explicit way for semantic searching and mapping. Our research objective is to develop an appropriate methodology and a conceptual model to provide semantic level representation for

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 61–75, 2007.

processing and modeling international copyright law articles using ontological engineering. Such a method and conceptual model can not only deal with the problems of the diversity of national law representations, but can also provide a fine perspective on international law article matching.

Copyright legal knowledge can be shared and retrieved less ambiguously in its explicit form, and this became especially important when machines started to be applied to facilitate legal knowledge management. In [17], knowledge representation can be defined as the application of logic and ontology to the task of construction of computable models of some domain. Motivation for using ontologies can be applied to the domain of law: the inter-relation of the law makes it a natural area for the knowledge sharing; the importance of legal decision argue for a high level of verification [3]. Moreover an ontology can serve as an explicit and less ambiguous description of concepts and relations among them appearing in the target thing [13, 14]. Many organization and researches have tried and are trying to enhance shareability and reusability of legal knowledge and legal information using ontological techniques.

In the last several decades, important efforts in the field of artificial intelligence and law have appeared. McCarty's (1989) Language of Legal Discourse (LLD) supported the following modalities: time, events and actions, and deontic expressions [12]. Hamfelf and Barklund (1990) proposed and implemented a representation of legal knowledge in which Hart's theory was cast in meta-levels of a logic programming formalism [6]. Allen and Saxon's (1991, 1997) Language for Legal Relations (LLR) transformed Hohfeldian legal theory [1, 2]. Valente's (1995, 1999) FOLaw distinguished the various types of knowledge in legal reasoning, including normative knowledge, meta-legal knowledge, world knowledge, responsibility knowledge, reactive knowledge, and creative knowledge [18, 19]. Breuker's (2004) LRI-Core, a core ontology for law that from the thought of a common sense foundational ontology, included five major categories in the top-layer of LRI-Core: physical, mental, abstract, role, and occurrence [4, 5]. Open Digital Rights Language ODRL has been developed to build an open standard for expressing machine-readable licenses for digital materials [16]. IPROnto formalized a semantic web approach, and structured intellectual property right information [7]. These works either try to formalize and systematize legal theory or knowledge, or cope with digital rights management, but all of them could be said at a higher level of abstraction that may serve a number of purposes. However, the ontology we have developed is concerned with the diversity of legal representations of law articles and tries to handle the difficulties of matching law articles among different national copyright laws.

Our study started from intention theory, and a preliminary **i**ntention-oriented Legal Knowledge Model (iLKM) as a *pivotal* model has been proposed [10, 11]. From intention behind the law viewpoint, iLKM manages and models legal knowledge derived from different national copyright law documents. Meanwhile we propose an ontology-aware approach that provides semantic primitive representation of copyright law. International Copyright Law Ontology (ICLOnto) we have developed is used as a fundamental conceptual framework to maintain consistency among diverse legal representations for a certain legal phenomenon.

This paper first gives a narrative overview of international law articles consulting system and its typical functions. After that iLKM and ICLOnto has been elaborated in

Section 3 and Section 4 respectively. Following that, the example was described in Section 5. Finally, the paper closes with conclusion remarks and future perspectives.

2 Narrative Overview of International Law Articles Consulting System

2.1 The Jam of Copyright Management Faced by Intercultural Collaborators

A law is a series of articles represented by text, and each article prohibits activities undesirable in human society, permits reasonable activities or forces people to do desirable activities. Each nation of the world has its own legal system which depends heavily on the nation's language, culture, customs, history and so on. Getting intercultural collaborators to understand critically about the commonalities and differences between their own and others' diverse legal knowledge representation is important in order to facilitate cross-cultural interaction, learning and appreciation.

In general, it is not easy for non experts to comprehend or grasp the essence of law articles derived from foreign legal documents, even if there is no natural language barrier. It falls outside the scope of this research to perform natural language translation. For consistency, the English translation version is considered as the standard text, instead of using each native language representation. We confine ourselves to the issue of diversity of knowledge representation in different legal systems.

For example, the Japanese legal system has a legal subsystem to regulate in detail the educational activities using copyrighted work, while the Chinese one does not have a similar representation of such regulation. In Chinese legal system, a legal subsystem regulates the use of copyrighted work in *public* activities that include educational activities. The two legal systems frame very similar regulations on educational activities because both of them are based on the similar legal intention of "*protect exemptible usage rights*". The differences between the structures of law documents come from the degree of granularity of conceptualization of "public activities."

2.2 An Ontology-Aware International Articles Consulting System

A key requirement for a common conceptual foundation is that it should be less dependent on a nation-specific conceptualization of a legal system. In this research, we take an ontological approach to model legal systems to satisfy the key requirement mentioned. A simple solution to the problem shown in the aforementioned example is to model "educational activity" as "*a-kind-of (is-a)* public activity" and reveal the implicit common legal intention, "*protect exemptible usage rights*", behind the two legal subsystems. To address the above issue, we adopted an ontological approach to legal systems using the intention-oriented Legal Knowledge Model, in which the intention behind law is supposed to be working as a suitable *pivot* among different legal systems.

The main difficulty of building an ontology is to identify and classify the items of a given domain. Since classification criteria depend on purpose and are not universal, we do not seek to build a universal ontology, but merely a special ontology for intercultural collaboration in copyright management. International Copyright Law Ontology, as a domain ontology, is intended to capture the essential conceptual entities and relationships in the knowledge structure about international copyright. With the help of ICLOnto, we want to represent and store models of typical intentions behind copyright law articles, and facilitate users' analysis and understanding of a commonalities or differences among law articles in different countries by referring to the models as the corresponding pivot.

The International Copyright Law Articles Consulting System (under development) is an intelligent article management environment that is now being developed. Its functions are designed based on intentions of law and legal activities defined in the intention-oriented Legal Knowledge Model. There are two law articles in the article browser of Figure1. In this case, how could a cross-cultural collaborator detect the correspondences of these two articles, regarding Japanese and Chinese laws?

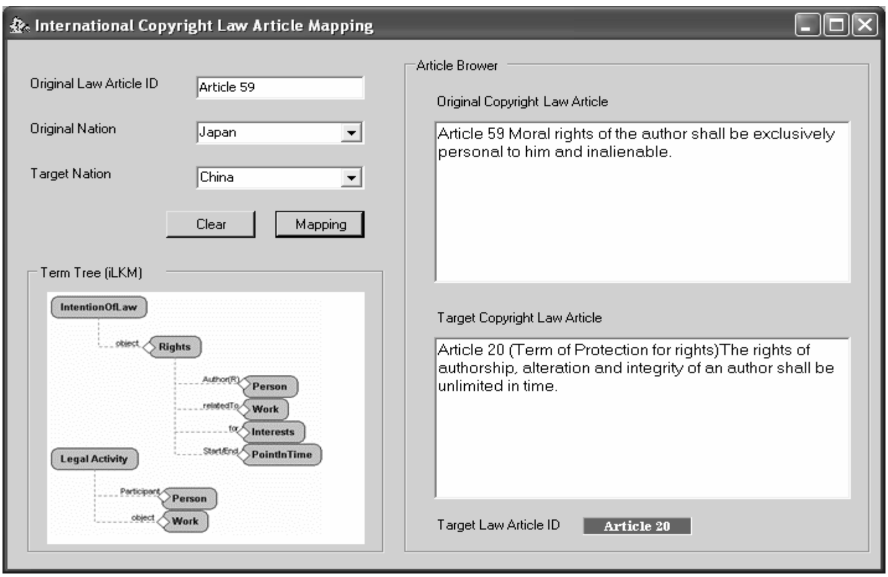


Fig. 1. A cognitive interface for international copyright law article mapping

The typical functions of the system designed are mainly three-fold:

- Browsing services: Services support the users in browsing national copyright law articles of many nations according to their need.
- Matching and retrieval services: Services support the users in matching content benefit if the intentional orientation of a copyright law article is made explicit. Based on matching, the correlative articles can be retrieved. (We should notice that this paper emphasizes this function, as shown in Figure 1. An in-depth analysis of how the system supports matching based on ICLOnto will be elaborated in Section 6.)

❖ Guidance and consulting services: Services help to clarify the users' problems by modeling the user's situation, and suggest a typical procedure to solve those problems by providing copyright law article information.

A cognitive interface of the article consulting system is shown in Figure 1. The upper left part of the cognitive interface is used to input the profile information of the original article, including law ID and nation. The user can designate a target nation for his/her request. The lower left part of the window is the monitor part to provide the user with conceptual index information of the law article modeled in iLKM. If the user clicks "mapping", the detailed article contents will be shown in the right part of window, the article browser, which contains both the original article and the target article. At the bottom left, the appropriate mapped law ID will be indicated to meet the user's request for a target nation copyright law.

3 iLKM Providing Capability of Handling the Diversity of Knowledge Representation

As mentioned above, a copyright law is a series of articles represented by text to regulate the various actions and relationships of copyrights existing in the real world. Each nation of the world has its own copyright legal system which depends heavily on the national culture, customs and history. As shown in Figure 2, there are different legal representations to describe a single law phenomenon, such as Law Article (LA) derived from the United States, China or Japan. We need to find an independent "thing", which works as a desirable pivot amongst different law representations, and which corresponds with this law phenomenon. We call it intention-oriented Legal Knowledge Model (iLKM).

We produced this pivotal model (iLKM) for handling the diversity of legal knowledge representation from two separate models: a law article model and an intention model. The law article model is a model of static profile information of articles. The intention model is a model representing the essential meaning of law per se to capture the commonalities and differences among law articles.

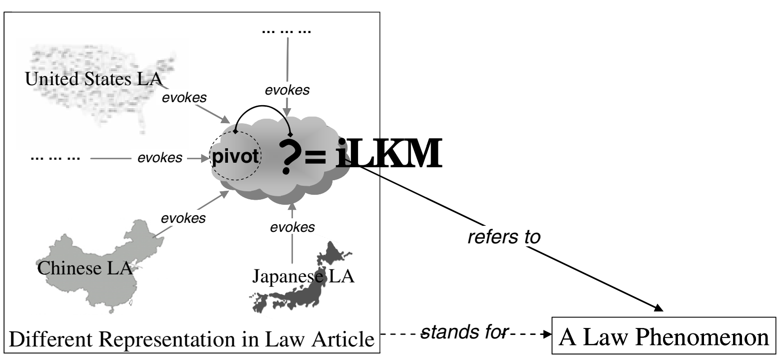


Fig. 2. iLKM serves as a pivot amongst different representations of law articles derived from different countries

3.1 Law Article Model

Even though there is diversity of both legal systems and legal knowledge representations in each nation, national copyright laws always embody some underlying common components from legislative point of view. In general, one copyright article contains 4 aspects of basic information, which can be considered as a blend of 4 elements. They are the law document per se, natural language expression, national derivation, and legal resource identification which is used to point out uniform legal resource locator. We model “Law Article” as a concept that consists of 4 elements in accordance with the above analysis: (1) Country Name, (2) Text, (3) Language, and (4) Law IDs. (as shown in Figure 3-a) These elements are reserved for specifying the basic static information of a law article necessary for international law article mapping.

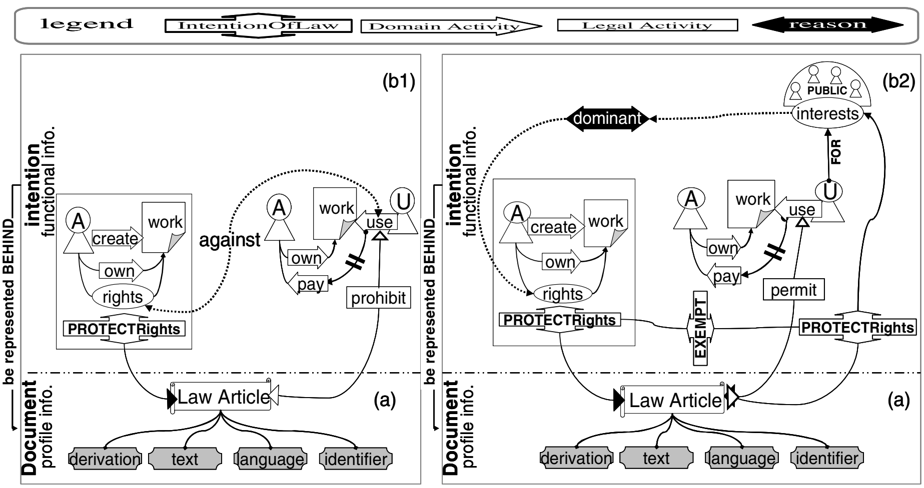


Fig. 3. Law Article Model (a) & Intention Model (for example, (b1)-PROTECTCopyrights and (b2)-EXEMPT)

3.2 Intention Model

In contrast to profile information, the law articles, as a kind of document, should represent its functional information; that is intention of this document. We model the “Intention Of Document” consisting of 2 slots: essential part (i) IntentionOfLaw, and execution part (ii) Legal Activity. Moreover we choose to use the web ontology language (OWL) paradigm according to the W3C recommendation [20] in order to describe reusable module structures, because this seems to have become the new standard for processing information in a web-based environment. For instance, Slot(i) and Slot(ii) can be described in OWL as follows:

```
<owl:Class rdf:ID="Intention_Of_Document">
  <rdfs:label>Intention Of Document</rdfs:label>
  <rdfs:subClassOf>
    <owl:Restriction>
```

```
    <owl:cardinality
    rdf:datatype="http://www.w3.org/2001/XMLSchema#int">
      1
     </owl:cardinality>
     <owl:onProperty rdf:resource="#has_essential_part"/>
   </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty
       rdf:resource="#has_essential_part"/>
       <owl:allValuesFrom rdf:resource="#IntentionOfLaw"/>
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:cardinality
     rdf:datatype="http://www.w3.org/2001/XMLSchema#int">
       1
      </owl:cardinality>
      <owl:onProperty rdf:resource="#has_execution_part"/>
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#has_execution_part"/>
      <owl:allValuesFrom rdf:resource="#Legal_Activity"/>
    </owl:Restriction>
  </rdfs:subClassOf>
</owl:Class>
```

These two slots both constitute core ingredients of “Intention Of Document” for they clarify the essential meaning of law per se, we call it “IntentionOfLaw”; and figure out how to put this intention into practice by legal activity. For instance, the intention “to PROTECT” (author’s economic rights) can be realized by a legal activity “to prohibit” (copy of author’s work for sale without permission or payment).

This research attempts to develop a common conceptual foundation to represent commonalities or differences among different national copyright laws. We propose an ontological approach to systematize and conceptualize copyright legal knowledge with intention-oriented Legal Knowledge Model (iLKM) in which the intention behind the law should work as a suitable pivot among different national legal systems. The first task is to develop iLKM to appropriately reflect the meaning of the legal articles. In order to capture the essential meaning of law articles and facilitate their modeling, IntentionOfLaw is modeled as a *verb* or *verb phrase*, which consists of four typical concepts (PROTECTCopyright, PROTECTExemptibleEnd-userRights, DEFINELegalTerm, and EXEMPT) to clarify four kinds of intention of laws. As argued above, IntentionOfLaw is viewed as an essential part, which will be implemented by various legal activities (as execution part clarified in Section 4.3) in different countries.

- Concept of "PROTECTCopyright" (Figure3-b1) and concept of "PROTECTExemptibleEnd-userRights" both are kinds of single intention, which specify an article has only one intention, that is, to protect. The object of single intention is something that appears as a focus of intentional behavior.
- Concept of "DEFINELegalTerm" is a kind of single intention that assigns a meaning to a legal term or its variant form used in a law article. The object of "DEFINELegalTerm" is a Class.
- In Contrast to the above two concepts, "EXEMPT" (Figure3-b2) is a compound intention that specifies at least two single intentions involved in one article, and interaction takes place between these intentions. The interaction will be considered as a domination relation between "Rights" or "Activities" in order to account for intention of exempt.

4 Description of the Domain Ontology of International Copyright Laws

To reach the aforementioned research objective, we need to design and develop an ontology that conceptualizes different national copyright law articles from an intentional perspective. The concept defined in this ontology stands for a particular intentional role a legal article, for instance the "PublicInterest", can be protected by performing different legal activities in different countries. In the following, we will describe the classes, properties and relationships in the domain ontology (ICLOnto).

4.1 Concept of "Rights"

Legal rights are, clearly, rights which exist under the rules of legal systems. From the copyright law point of view, in International Copyright Law Ontology (ICLOnto) Rights is a key and complex concept consists of participants involved, related work, and interests affected by rights. At the topmost of the definition hierarchy, the concept of "Rights" is classified into two sub-concepts, which are, Copyright and End-userRights. The former has mainly followed the World Intellectual Property Organization (WIPO) recommendations [21]. Figure 4 shows the included rights hierarchy starting from Rights. There are "Economic Rights" and "Moral Rights", as promoted by the WIPO, and "Neighboring Rights" (or named "Related Rights"). The latter one we defined, actually that is, a kind of rights to consume copyrighted work. It has been further distinguished into two types: rights for general usage and rights for exemptible usage. As for GeneralEnd-userRigths, A1) the participant involved is a person who represents their own PrivateInterests, and A2) effect brought by Rights is a certain PrivateInterests (e.g. general user's interests who utilizes the copyrighted work for profitable business). The relationship between these two PrivateInterests in a1 and a2 is either selfsame or is-a kind of relation depending on case. On the other hand, as for ExemptibleEnd-userRights, B1) the participant involved is a person who represents certain PublicInterest, and B2) effect brought by these Rights is also a kind of PublicInterest (e.g. teacher's interest for non-profit education).

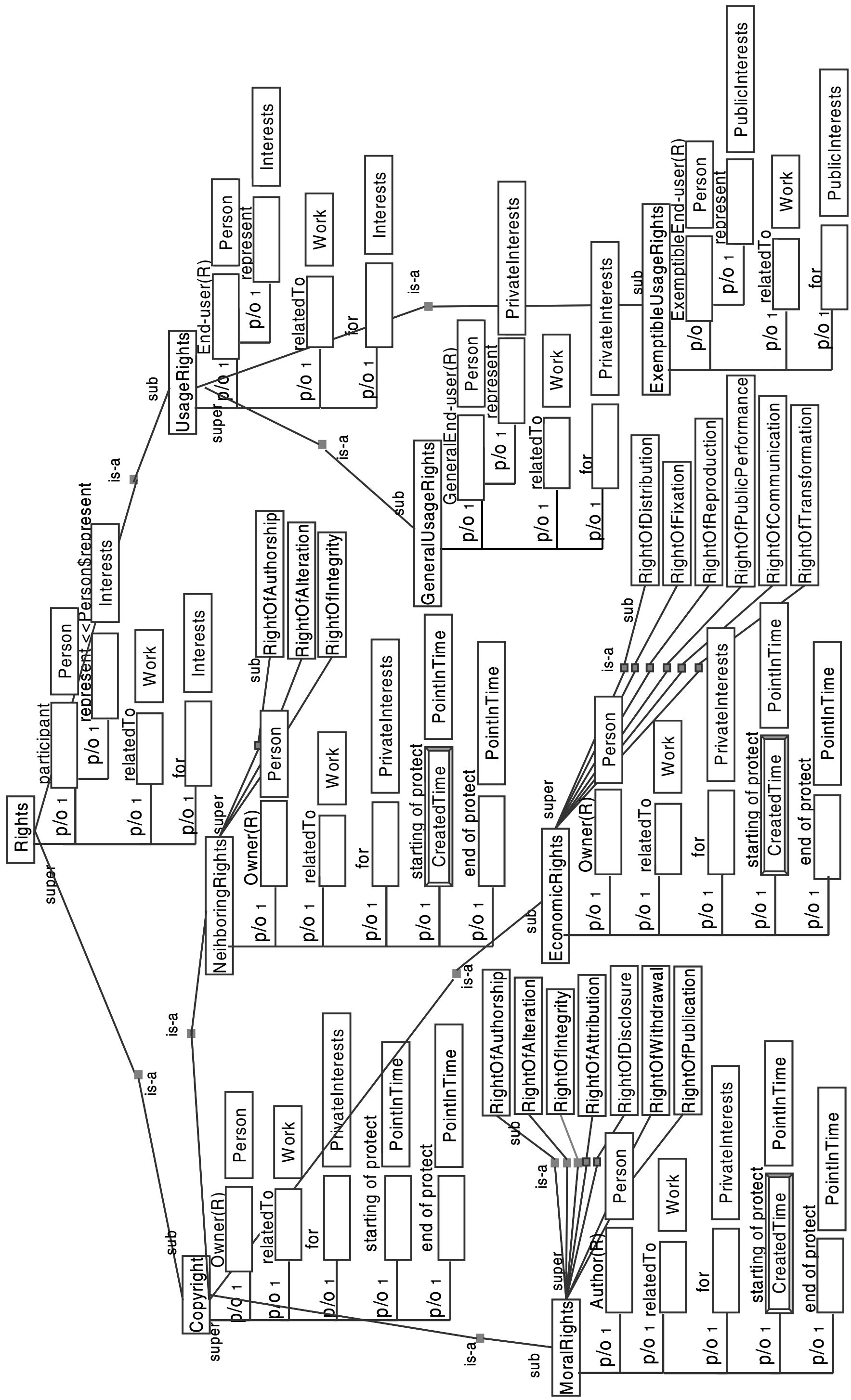


Fig. 4. A definition hierarchy of ‘Rights’

4.2 Basic Concept of "Person" and Role Concept of "Participant"

In terms of legal dictionary, "person" is defined as a) a human being (we call it "Natural Person" in our ontology); and b) a corporation treated as having the rights and obligations of a person (we call it "Corporate Person" in our ontology). Meanwhile the basic concept of "Person" (both Natural Person and Corporate Person) is defined with the objective of reflecting participant involved (1) static information (profiles such as name, gender, nationality, ID, age and affiliation for Natural Person; name, nationality, location, builtTime for Corporate Person) and (2) dynamic information including certain rights held and certain interests represented.

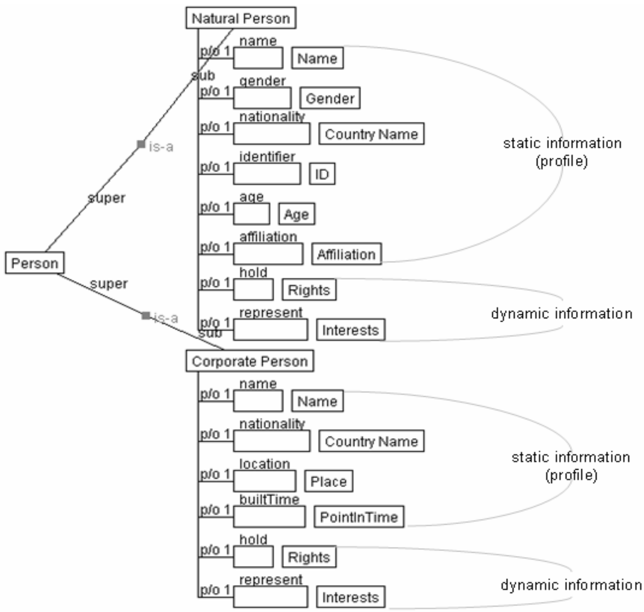


Fig. 5. Basic concept of "Person"

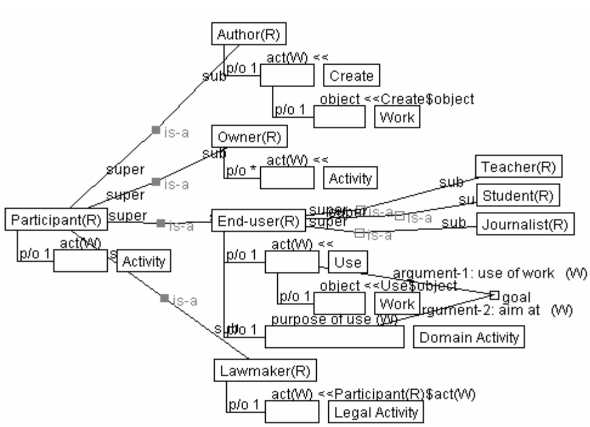


Fig. 6. Role concept of "Participant"

The Participant as role concept is defined here, and represents a role which a person plays in a specific legal context, and it is defined with other concepts (e.g. Activity). These leaf nodes contain incomplete lists, for example, there can be other types of role concept of "End-user(R)" in addition to "Teacher(R)", "Students(R)" and "Journalist(R)".

4.3 Legal Activity and Domain Activity

As argued above, an intention of law can be put into practice by a legal activity. It should be noted that it is a special characteristic of iLKM to distinguish the concept of "Legal Activity" from the concept of "Domain Activity"; this provides a pragmatic approach to clarify the commonalities and differences based on IntentionOfLaw.

As we can detect from this frame structure, the object of "Legal Activity" is restricted by sets of Activity (it might be either Legal Activity or Domain Activity) or Rights; while for the object of "Domain Activity", it would be "entity", "substance" (e.g. copyrighted work, human being, and so forth) or Rights.

In essence, every legal or illegal act has a specifically juristic meaning. "Legal norms themselves are present as the meaning of certain 'acts of will': law-creating acts of will…… 'Norm' is the meaning of an act by which a certain behavior is commanded, permitted, or authorized." [8, 9] The subject of Legal Activity is Law per se, which represents the will of legislators. Therefore the sub-concept of Legal Activity is capable of interpreting certain "acts of will" from a legal point of view (e.g. "permit", "forbid", "prohibit", "oblige", "grant", "authorize", "enable" and so forth). The pragmatic initiative of iLKM is to semantically match diverse representations of Legal Activity among national copyright laws by detecting the same or similar IntentionOfLaw.

4.4 The Role of Relational Concepts

A relation is a modeling construct which links (sets of) objects to each other, which express interdependencies among phenomena. A relation may state the consequences of some event, or impose new roles on existing objects. ICLOnto is composed of concepts necessary to explain the target world (copyright legal knowledge) and the relationships between these concepts. The most basic relationship is the general-specific (is-a) relationship, and description using concept labels and hierarchy with "is-a" relationship is the most primitive ontology.

Moreover, in order to cater to the specific requirements for building models, definition (constraints) and description of relationships (axiomatic description) can be added in our ontology, because most things are composed of parts and those parts are connected by a specific relation to form the whole. In contrast to the wholeness concept, in which a concept of a thing is considered as a whole and is composed of multiple concepts, the relational concept is conceptualized relationship between multiple concepts [15]. The number of arguments in a relation is not restricted. Currently in ICLOnto, binary relation is viewed as the main type of conceptualized relationship. The relational concept can be used to link any types of concepts to each other, including basic concepts (e.g. IntentionOfLaw and Legal Activity), role concepts (e.g. between Participants), and role holders (e.g. between the instances of Participants). In the Table 1 below we give some examples of relational concepts defined in ICLOnto.

Table 1. Instantiate the conceptual structure of relational concept

Relational Concept Name	Arguments in relationship (binary)	Definition (constraints)	Axiomatic description	Example
RC: realization	*arg-1* intention	IntentionOfLaw	IntentionOfLaw will be realized by Legal Activity.	The intention "to PROTECT" (author's economic rights) can be realized by a legal activity "to prohibit" (copy of author's work for sale without permission and payment).
	arg-2 ways to realize intentio n	Legal Activity		
RC: domination	*arg-1* superior	Rights/Legal Activity	A certain Right/ Legal Activity can dominate another one.	For educational usage of copyrighted work, the PublicInterest dominates PrivateInterests.
	arg-2 inferior	Rights/Legal Activity		
RC: goal	*arg-1* use of work	Use	To use work aims at the certain domain activity.	A librarian copies a copyrighted book in order to archive.
	arg-2 aim at	Domain Activity		

5 Example

In this paper, we have opted for use of the current Chinese Copyright Law and the current Japanese Copyright Law to serve as examples. Here we revisit the class and properties defined in ICLOnto to illustrate the two articles mentioned in Section 2.2.

(a) "Moral rights of the author shall be exclusively personal to him and inalienable."

[in JP Law]

(b) "Term of protection for rights: The rights of authorship, alteration and integrity of an author shall be unlimited in time."

[in CN Law]

Let assume that here two relevant queries are how the Chinese law and the Japanese law treat the moral rights of an author and whether they are same or not. Only using text processing, it is hard to find the semantic similarity between the above two articles. There are just clusters of unstructured words and phrases, i.e. "moral rights of author", "exclusively personal", "inalienable" (in JP (a)) and "rights of authorship", "rights of alternation", "rights of integrity", "unlimited", "in time" (in CN (b)).

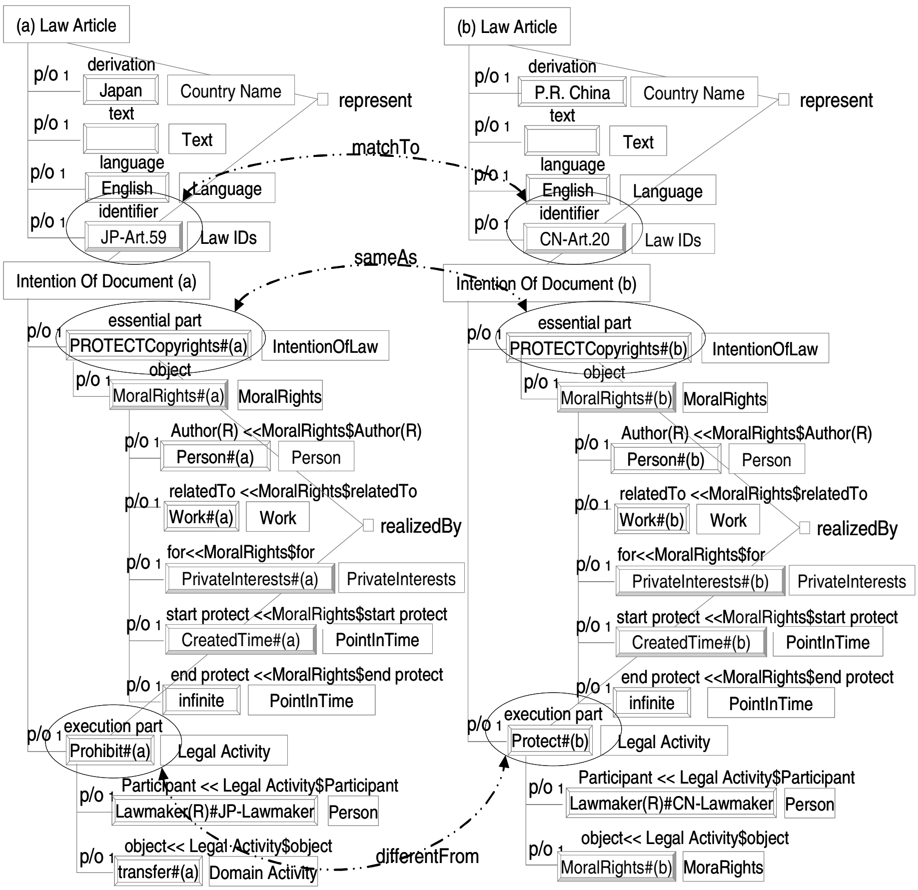


Fig. 7. Illustration of iLKM with real national copyright law articles

JP (a) and CN (b) have the same intention; both are to PROTECTCopyrights (moral rights of author is <u>*a-kind-of*</u> copyright). However the legal activities for realization of such intentions are different, and in this case entirely opposite. The legal activity in JP is "to prohibit", while in CN "to protect".

Figure 7 makes clear both the structural similarity and relationship of reflection at the semantic level. Such structural similarity reflects the semantic similarity that provides the sharable common conceptual foundation between JP article and CN article. It can be expected that such structural similarity can derive right answer to the questions above.

Meanwhile there is no legal content to regulate if moral rights can be transferred (a-kind-of domain activity) in Chinese law system. However these two articles about moral rights have been augmented with iLKM. Based on the same intention (PROTECTCopyrights), the implicit legal knowledge is revealed, that is, law prohibits moral rights of author from being transferred in Chinese legal systems too.

6 Summary and Future Work

This paper gives a description of the effort to build a domain ontology concerning multinational copyright law -- International Copyright Law Ontology with intention-oriented Legal Knowledge Model that is a suitable pivotal model coping with diversity of legal representation. The intention behind law is modeled as the core of this conceptualized framework that may be used to reflect the essential meanings of law documents derived from different countries, as well as the approach in which we can clarify the commonalities and differences of intentional copyright laws. We have given an example in order to show how elements of the intention of law can be explicated and serve as bridges for semantic mapping in terms of the model. In this study, semantic mapping of national copyright law, by systematizing and conceptualizing legal knowledge, is typically achieved in three affinitive layers: ontology layer, law model layer, and case model layer. iLKM is model of law, a kind of abstract modeling, which models legal knowledge including intention of law, activity, legal claims, and relations. The case model is a more concrete instance modeling. The near future work will be to take into account the case model in order to verify case-specific conceptualization as well to expand the domain ontology.

References

1. Allen, L.: The Language of Legal Relations, RLL: Useful in a Legal Ontology Toolkit? Legal Ontologies. In: Proceedings of the ICAIL-97 workshop, pp. 47–60 (1997)
2. Allen, L., Saxon, C.: More IA Needed in AI: Interpretation Assistance for Coping with the Problem of Multiple Structural Interpretations. In: Proceedings of the Third International Conference on AI and Law, pp. 53–61 (1991)
3. Bench-Capon, T.J.M., Visser, P.R.S.: Ontologies in legal information systems: the need for explicit specifications of domain conceptualizations. In: Proc. of the 6th international conference on artificial intelligence and law, Australia (1997)
4. Breuker, J.: Constructing a legal core ontology: LRI-Core. In: Proceedings WONTO-2004, Workshop on ontologies and their applications, Porto Alegre, BR, pp. 115–126 (2004a)
5. Breuker, J., Hoekstra, R.: Epistemology and ontology in core ontologies: FOLaw and LRI-Core, two core ontologies for law. In: Proceedings of EKAW Workshop on Core ontologies2004 (2004b)
6. Hamfelt, A., Barklund, J.: Metaprogramming for Representation of Legal Principles. In: Proceedings of Meta' 90, pp. 105–122 (1990)
7. IPROnto: available online at: http://dmag.upf.es/ontologies/ipronto/index.html

8. Kelsen, H.: Pure Theory of Law, 2nd edn. University of California Press, Berkeley ans Los Angeles (1967)
9. Kelsen, H.: General Theory of Norms. Clarendon Press, Oxford (1991)
10. Lu, W., Ikeda, M.: An Intention-oriented Model of Copyright Law for e-Learning: International Semantic Mapping of Copyright Laws Based on A Copyright Ontology. In: Proceedings of the 13th International Conference on Computers in Education 2005, Singapore, pp. 753–756. IOS Press, Amsterdam (2005a)
11. Lu, W., Ikeda, M.: International Semantic Mapping of Copyright Law: A Conceptualization Supporting System with Capability of "Intention" and Ontology. In: Proc. of IFSR2005, Japan. CDROM (2005b)
12. McCarty, T.: A Language for Legal Discourse I. Basic Structures. In: Proceedings of the Second International Conference on AI and Law, Vancouver, pp. 180–189 (1989)
13. Mizoguchi, R., Ikeda, M.: Towards ontology engineering. In: Proc. of the Joint Pacific Asian Conference on Expert Systems. Singapore International Conference on Intelligent Systems, pp. 259–266 (1997)
14. Mizoguchi, R., Bourdeau, J.: Using Ontological Engineering to Overcome AI-ED Problems. International Journal of Artificial Intelligence in Education 11(2), 107–121 (2000)
15. Mizoguchi, R.: Ontology Engineering Environments. In: Staab, S., Studer, R. (eds.) Handbook on Ontologies, pp. 275–295. Springer, Heidelberg (2003)
16. ODRL: available online at: http://odrl.net
17. Sowa, J.F.: Knowledge Representation: Logical, Philosophical, and Computational Foundations. Brooks Cole Publishing Co. Pacific Grove, CA (2000)
18. Valente, A.: Legal Knowledge Engineering: A Modelling Approach. IOS Press, Amsterdam, Netherlands (1995)
19. Valente, A., Breuker, J., Brouwer, P.: Legal Modelling and automated reasoning with ON-LINE. International Journal of Human Computer Studies 51, 1079–1126 (1999)
20. Web Ontology Language (OWL) W3C recommendation (2004), available online at: http://www.w3.org/2004/OWL
21. WIPO: available online at: http://www.wipo.int/copyright/en/

The Lextype DB: A Web-Based Framework for Supporting Collaborative Multilingual Grammar and Treebank Development

Chikara Hashimoto[1], Francis Bond[2,⋆], and Dan Flickinger[3]

[1] Graduate School of Informatics, Kyoto University
Yoshida-Honmachi, Sakyo-ku, Kyoto, Japan 606-8501
ch@yz.yamagata-u.ac.jp

[2] Natural Language Research Group, NTT Communication Science Laboratories
2-4 Hikaridai, Seika-cho, Soraku-gun, Kyoto, Japan 619-0237
bond@nict.go.jp

[3] CSLI, Stanford University
Stanford CA 94305-2150 USA
danf@csli.stanford.edu

Abstract. We have constructed a web-based framework for collaborative multilingual grammar and treebank development in which developers are distributed around the world. It is important for developers of the world-wide collaboration to **i)** grasp and share the big picture of the grammar and treebank of each language and **ii)** understand commonalities of languages. Our framework, the Lextype DB, describes lexical types of the grammar and treebank. Lexical types can be seen as detailed parts-of-speech and are the essence for the two important points just mentioned. Information about a lexical type that the Lextype DB provides includes its linguistic characteristics; examples of usage from a treebank; the way it is implemented in a grammar; and correspondences to major computational dictionaries. It consists of a database management system and a web-based interface, and is constructed semi-automatically. Currently, we have applied the Lextype DB to grammars and treebanks of Japanese and English.

Keywords: Multilingual Grammar, Multilingual Treebank, Collaborative Development, Documentation, Web-based Technology.

1 Introduction

Treebanks constructed with detailed linguistic information play an important role in various aspects of natural language processing; for example, grammatical knowledge acquisition; world knowledge acquisition [1]; and statistical language model induction [2,3]. Such treebanks are typically semi-automatically constructed by a linguistically rich computational grammar. A detailed grammar

⋆ Current Affiliation: NICT Computational Linguistics Group.

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 76–90, 2007.

in turn is a fundamental component for **precise** natural language processing. It provides not only detailed syntactic and morphological information on linguistic expressions but also precise structural semantics, which can be used in, for example, machine translation [4].

The Deep Linguistic Processing with HPSG Initiative (DELPH-IN)[1] has been constructing open-source linguistically precise grammars and treebanks for several languages, including English [5], Korean [6], German [7], Spanish, French, Norwegian and Japanese [8]. DELPH-IN grammars are compatible in that they are all based on the same formalism, which is Head-driven Phase Structure Grammar (HPSG) [9,10], and can be used by the same processors. The semantics are based on Minimal Recursion Semantics [11], a shallow semantic representation that allows for underspecification of scope.

Developers of DELPH-IN are distributed all over the world and are contributing their expertise in linguistics to the DELPH-IN grammar and treebank construction via the Internet. Most grammars are developed along with one or more treebanks of examples. The grammars and treebanks are then available for download, either as snapshots or through CVS.

One of the purposes of DELPH-IN is to capture commonalities of human languages in the course of the development. Capturing commonalities (or universality) of human languages is not only of interest for theoretical linguistics but also an aid to multilingual grammar development since it makes existing grammars more compatible and developing a new grammar much easier. The universal core of the DELPH-IN grammars is codified in the Matrix [12], a bottom-up approach to building a universal grammar.

However, grammars and treebanks are getting more complicated and hard to maintain in the course of development. This is brought about by two factors; one is the linguistically sophisticated nature of DELPH-IN grammars and treebanks, and the other is involved with difficulties in communication during the collaboration over the Internet. The former comes about as a natural result of hand-crafting large-scale HPSG grammars and treebanks for practical NLP purposes, and the latter is a natural consequence of a collaboration in which participants are located away from each other, speak different languages and have different backgrounds and interests.

At this point, the Lextype DB comes on stage. It plays two roles; one is to automatically summarize a grammar and a treebank for each language in terms of lexical types, and the other is to show the summary to developers around the world through the Web. With the Lextype DB, a developer can grasp the big picture of the grammar and the treebank no matter how large they are, and developers distributed over the world can share the big picture.

In addition to alleviating the complication of a large-scale grammar and treebank, the Lextype DB helps to facilitate the understanding of commonalities of human languages. This is because the Lextype DB reveals the linguistic essence of a grammar in terms of lexical types for whatever language it deals with. If developers who are in charge of a particular language show the linguistic essence

[1] http://www.delph-in.net/

of the grammar by means of the Lextype DB, other developers can compare it with their own grammars easily through the Web.

Accordingly, the relations between the Lextype DB and the intercultural collaboration are twofold:

1. The Lextype DB facilitates the DELPH-IN collaboration, members of which come from many cultural (or linguistic) backgrounds. Thus, it facilitates the intercultural collaboration.
2. Also, it facilitates the understanding of commonalities of human languages, which is indispensable for intercultural understanding. This is another way of facilitating the intercultural collaboration.

The next section illustrates what a lexical type is in detail and how helpful it is to understand lexical types for collaborative grammar and treebank development. In the Section 3, we describe what contents the Lextype DB provides and how it is used in the grammar and treebank development collaboration. In Section 4, we discuss how it can be useful as a general lexical resource. After describing related and future work in Sections 5 and 6, we conclude the paper in the final section.

2 A Lexical Type in Grammar and Treebank Development

A lexical type in a DELPH-IN context refers to a group of lexical items whose linguistic behavior is the same as each other. Thus, you can think of a lexical type as a kind of a part-of-speech, but a lexical type is usually a finer-grained notion.

For example, the DELPH-IN grammar of Japanese distinguishes several usages of the Japanese dative marker *ni*. The Japanese sentence (1) can represent the two meanings described in (1a) and (1b). Lexical type names for each usage of *ni* are written in `typewriter` font.[2]

(1) hanasiai-wa sinya-**ni** itaru
discussion-TOP midnight-DAT reach

a. "The discussion comes (to a conclusion) at midnight."
ni as `adv-p-lex-1`

b. "The discussion continues until midnight."
ni as `ga-wo-ni-p-lex`

The dative phrase, *sinya-ni* (midnight-DAT), can act as either an adjunct (1a)[3] or an object of *itaru* "reach" (1b). Clearly, these two usages of *ni* show differences in both syntax and semantics. Below is an example showing other usages of *ni*.

[2] These are actual names of the lexical types implemented in the Japanese grammar and might not be understandable to people in general.

[3] The object, *a conclusion*, is unexpressed, the so-called "pro-drop" phenomenon.

(2) Ken-wa yuka-o kirei-**ni** migaku
-TOP floor-ACC clean-DAT polish

a. "Ken polishes a floor clean."
(The floor is clean.)
ni as `naadj2adv-end-lex`

b. "Ken cleanly polishes a floor."
(His way of polishing the floor is clean.)
ni as `adv-p-lex-6`

The dative phrase, *kirei-ni* (clean-DAT), is used as an adjunct in both (2a) and (2b), but their usages and meanings are different. The usage in (2b) is an ordinary adverb that describes the manner of Ken's polishing the floor as clean, while in (2a) the dative phrase describes the resulting situation of the floor after polishing as clean. In addition, the *ni*s in (1) and (2) are different in that the former takes nouns as its complement while the latter takes adjectives. Thus, the four usages in (1a), (1b), (2a) and (2b) must be distinguished so that we can obtain correct syntactic structures and semantic representations. In our terms, these *ni*s are said to belong to different lexical types.[4] Similarly, the Japanese grammar distinguishes usages of other words, notably functional ones.

Although the details of the lexical types are language specific, the differences in the semantics (argument versus modifier in (2) and resultative versus modifier in (3)) appear in all languages. Even the syntactic similarities will be similar in closely related languages, such as Korean.

There are several kinds of linguistic notion implemented in DELPH-IN grammars and treebanks other than lexical types: rules, principles, and root nodes, among others. However, we chose the lexical type as a first step toward a full support of the world-wide collaboration of grammar and treebank development for several reasons.

Firstly, HPSG is a lexical grammar, where lexical elements and phrases are associated with categories that have considerable internal structure. A typical grammar will have many more lexical types than it has construction specific rules or rule schemata.

Secondly, in treebank annotation, the most frequent operation is judging the correct lexical type for each lexical item in each sentence in the treebank. Thus, treebank annotators must be familiar with all the lexical types so that they would know which word usage is correct for a lexical item at hand.

Thirdly, in grammar development, the most frequent operation is adding lexical items. Then, developers have to see which lexical type each lexical item should belong to, and they must be familiar with all the lexical types implemented in a grammar. Otherwise, they might add a new lexical type that has an overlapping functionality with an existing lexical type. This causes spurious ambiguity. And the grammar will be unnecessarily bloated, and the treebank will be easily inconsistent.

[4] Usages of the Japanese dative marker, *ni*, are extensively discussed in, for example, [13].

Finally, lexical types are not so theory specific, so can be used to link to external linguistic resources.

As for human language commonality, other linguistic notions like rules and principles show more commonalities than lexical types in general. However, lexical types reveal commonalities of a language that are much easier to understand for developers who are not familiar with the language. Usually, commonalities related to rules and principles are harder to understand.

In summary, a lexical type, which is dealt with most frequently in the grammar and treebank development, is usually such a complicated notion that developers around the world can hardly grasp and share the big picture of a grammar.

3 Architecture of the Lextype DB

The Lextype DB is required to alleviate the complication of lexical types implemented in a grammar so that developers around the world can access the information. Detailed information about lexical types is, of course, available in the grammar development environment (the LKB[14]). However, to see this, the grammar must be loaded. This makes it difficult to look at, for example, Korean and English grammars while developing Japanese. Further, the grammar development environment is not integrated with the treebank, so it is not easy to go from the types to their usage examples.

3.1 Content of the Database

First of all, what information should be included in such a database to help treebank annotators and grammar developers to work consistently? Obviously, once we construct an electronic lexicon, whatever information it includes, we can easily see what lexical types are assumed in the grammar and treebank. But we have to carefully consider what to include in the database to make it clear how each of the lexical types are used and distinguished.

We include five kinds of information:

(3) Contents of the Database
- a. Linguistic discussion
  - i Name
  - ii Definition and linguistic discussion
  - iii Criteria to judge a word as belonging to a given lexical type
- b. Exemplification
  - i Words that appear in a treebank
  - ii Sentences in a treebank that contain the words
- c. Implementation
  - i The portion of grammar source file that corresponds to the usage
  - ii Comments related to the portion
- d. Links to other lexical resources

That is, we describe each lexical type in depth (3a–3c) and show correspondences to other computational dictionaries (3d).

Linguistic Discussion. To understand lexical types precisely, linguistic observations and analyses are a basic source of information.

Firstly, the requirements for naming lexical-types in a computational system (3ai) are that they be short (so that they can be displayed in large trees) and easily distinguishable. Type names are not necessarily understandable for anyone but the developers, so it is useful to link them to more conventional names. For example `ga-wo-ni-p-lex` is a *Case Particle*.

Next, the definition field (3aii) contains a widely accepted definition statement of the lexical type. For example, `ga-wo-ni-p-lex` (1b) can be defined as "a particle that indicates that a noun it attaches to functions as an argument of a predicate. Individual particles are distinguished by their PFORM." Users can grasp the main characteristics from this. Links to representative papers or books dealing with the lexical type can also be added here. This allows the grammar developers to quickly check against existing analyses, and allows users to find more information.

Thirdly, the criteria field (3aiii) provides users with means of investigating whether a given word belongs to the class. That is, it provides positive and negative usage examples. By such usage examples, developers can easily find differences among lexical types. For example, `adv-p-lex-1` (1a) subcategorizes for nouns, while `adv-p-lex-6` (2b) subcategorizes for adjectives. Sentences like (1a) and (2b) that fit such criteria should also be treebanked so that they can be used to test that the grammar covers what it claims. This is especially important for regression testing after new development.

Exemplification. Examples help users understand lexical types concretely. As we have constructed a treebank that is annotated with linguistic information, we can automatically extract relevant examples exhaustively. We give the database two kinds of examples: words, that are instances of the lexical types (3bi), and sentences, treebanked examples that contain the words (3bii). This link to the linguistically annotated corpus examples helps treebankers to check for consistency, and grammar developers to check that the lexical types are grounded in the corpus data.

For the Japanese grammar we link to the Hinoki treebank[15], and for the English grammar to the Redwoods treebank[3]. DELPH-IN grammars are moving towards including treebanks with the grammars, updated with each release. Therefore, there will always be some usage examples available to the Lextype DB.

Implementation. Grammar developers need to know the actual implementation of lexical types (3ci). Figure 1 shows the implementation of `naadj2adv-end-lex`.

In addition to the actual implementation shown above, we currently show its parent type or types, category of the head, which is, "SYNSEM|LOCAL|CAT|HEAD" in HPSG terms, valence information, which is "SYNSEM|LOCAL|CAT|VAL," and the semantic type, which is "SYNSEM|LOCAL|CONT." This is not always visible in the type's definition, as it may be inherited from a supertype.

```
naadj2adv-end-lex := lexical_sign-word &
 [SYNSEM j-synsem &
       [LOCAL[CAT[HEAD case-adv_head,
                 VAL obj-arg &
                     [COMPS #comps &
                          [FIRST[OPT - ,
                                LOCAL[CAT[HEAD na-adj_head,
                                          VAL.UNSAT +],
                                      CONT[RELS <!#key & [ARG1 #xarg]!>,
                                           HOOK[LTOP #tophand,
                                               INDEX #ind]]]]]]],
             ARG-S #comps,
             CONT[HOOK[LTOP #tophand,
                      XARG #xarg,
                      INDEX #ind],
                  RELS <! !>,
                  HCONS <! !>]],
        LKEYS.KEYREL #key & [ LBL #tophand],
        NON-LOCAL[QUE <! !>,
                 AFFIX <! !>]],
  INFLECTED +].
```

Fig. 1. Actual implementation of `naadj2adv-p-end-lex`

Comments about the implementation (3cii) are also helpful to ascertain the current status. Although this section is necessarily framework-dependent information, all project groups that are constructing detailed linguistic treebanks need to document this kind of information. We take our examples from JACY [16], a large grammar of Japanese built in the HPSG framework. As actual implementations are generally incomplete, we use this resource to store notes about what remains to be done.

Links to Other Lexical Resources. This information helps us to compare our grammar's treatment with that of other dictionaries. This comparison would then facilitate understanding of lexical types and extension of the lexicon.

We currently link the Japanese grammar to those of ChaSen [17], Juman [18], ALT-J/E [19] and EDICT [20]. For example, `ga-wo-ni-p-lex` is linked to ChaSen's particle-case_particle-general, Juman's case_particle, and ALT-J/E's adjunct-case_particle-noun/particle_suffix [21]. In general, Jacy makes finer distinctions than ChaSen, Juman or EDICT, and has roughly the same level of granularity as ALT-J/E.

We link the English grammar's lexical types to COMLEX [22].

In addition to these language specific resources, we link the lexical types to GOLD's linguistic ontology.[5] GOLD is not a lexicon, rather it is an upper ontology for descriptive linguistics, providing a set of (possibly universal) linguistic notions. Hence, linking to GOLD is an aid to understand a implemented linguistics grammar and a treebank from a universal grammatical point of view. For example, `ga-wo-ni-p-lex` is linked to GOLD's Postposition.

[5] `http://www.linguistics-ontology.org/gold.html`

3.2 Method of Database Construction

The next question is how to construct such a database. Needless to say, fully manual construction of the database is not realistic, since there are hundreds of lexical types and they change as the grammar is developed. In addition, we assume that we will refer to the database each time we annotate parser outputs to build the treebank and that we develop the grammar based on the treebanking result. Thus the database construction process must be quick enough not to delay the treebanking and grammar development cycles.

To meet the requirement, our method of construction for the lexical type database is semi-automatic; most of the database content is constructed automatically, while the rest must be entered manually. This is depicted in Figure 2.

- Content that is constructed automatically
  - Lexical Type ID (Grammar DB)
  - Exemplification (3b) (Treebank DB)
  - Implementation (3ci,ii) (Grammar DB)
  - Link to Other Lexicons (3e) (OtherLex DB)
- Content that is constructed manually
  - Linguistic discussion (3a)

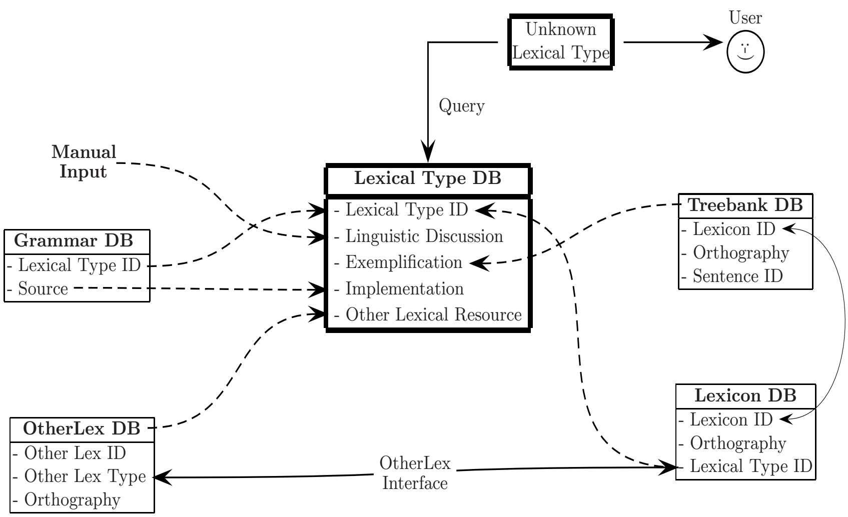


Fig. 2. The lexical type database construction

Component Databases. To understand the construction process, description of the four databases that feed the lexical type database is in order. These are the grammar database, the treebank database, the lexicon database, and the OtherLex database.

– The grammar database contains the actual implementation of the grammar, written as typed feature structures using $\mathcal{TDL}$ [23]. Although it contains the whole implementation (lexical types, phrasal types, types for principles and so on), only lexical types are relevant to our task.
– The lexicon database gives us mappings between words in the grammar, their orthography, and their lexical types. Thus we can see what words belong to a given lexical type. The data could be stored as $\mathcal{TDL}$, but we use the Postgresql lexdb [24], which simplifies access.
– The treebank database stores all treebank information, including syntactic derivations, words, and the lexical type for each word. The main treebank is stored as structured text using the [incr tsdb()] competence and performance profiler [25]. We have also exported the derivation trees for the treebanked sentences into an SQL database for easy access. The leaves of the parse data consist of words, and their lexicon IDs, stored with the ID of the sentence in which the word appears.
– We also use databases from other sources, such as ChaSen, Juman, EDICT and GOLD.

These databases provide a snapshot of the current state of the grammar. They are constructed automatically from the grammar development environment and a series of perl scripts.

Linguistic discussion (3a) has to be entered manually. Linguistic discussion is especially difficult to collect exhaustively since the task requires an extensive background in linguistics. We have several linguists in our group, and our achievements in this task owe much to them. We plan to make the web interface freely accessible, as well as releasing the data and tools as open source to encourage the participation of anyone interested in the task.

The on-line documentation is designed to complement the full grammar documentation [26]. The grammar documentation gives a top down view of the grammar, giving the overall motivation for the analyses. The lexical-type documentation gives bottom up documentation. It can easily be updated along with the grammar.

Each time the grammar is revised based on treebank annotation feedback, grammar developers consult the database to see the current status of the grammar. After finishing the revision, the grammar and lexicon DBs are updated, as are the corresponding fields of the lexical type database. Each time the treebank is annotated, annotators can consult the database to make sure the chosen parse is correct. Following annotation, the treebank DB is updated, and so is the lexical type database. In parallel to this, collaborators who are familiar with linguistics continue to enter relevant linguistic discussions via the WWW.

As much as possible, we are integrating the Lextype DB infrastructure into the existing grammar development environment, so that any grammar developer can take advantage of it with almost no additional effort. To this end, we are now simplifying the software. In particular, we are trying to reduce the number of external dependencies: for example, moving the database from java to perl, which we need for the scripts anyway; and adding routines to dump any information needed directly from the grammar development environment and treebanking tools.

4 Lexical Type Database as a General Linguistic Resource

In this section, we discuss some of the ways the database can enhance collaboration between treebank annotators, grammar developers and other interested parties.

One way is by serving as a link to other lexical resources. As mentioned in the previous section, Lextype DBs are linked together thought the GOLD upper ontology. Each language also includes links to other resources such as EDICT and COMLEX. Many lexical resources have been developed, but their intercorrespondences are not always clear. These lexical resources often play complementary roles, so synthesizing them seamlessly will make a multilingual lexicon with the widest and deepest knowledge ever.

In particular the linguistic notions assumed in GOLD that has been briefly mentioned in §3 act as a "hub" that connects already-connected lexical resources of several languages (Figure 3). Although GOLD's linguistic notions are a bit

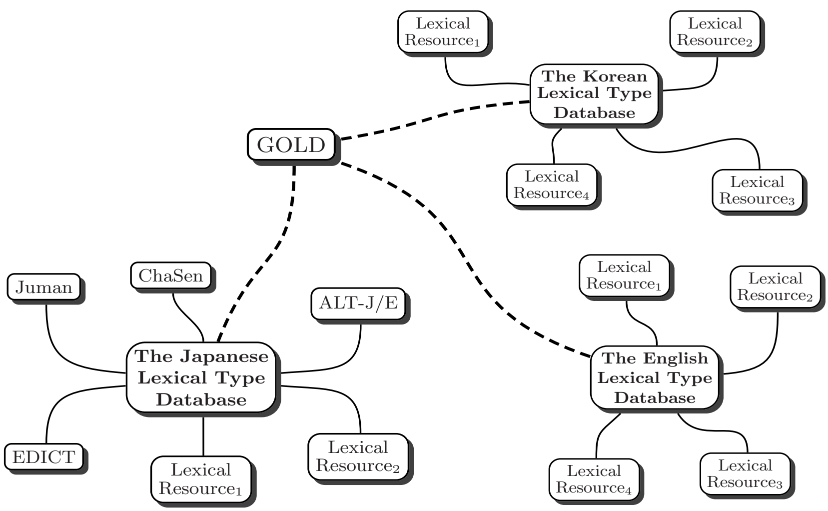


Fig. 3. Synthesis of universal lexical resources

coarse, they would tell us inter-lingual correspondences between lexical types of different languages. Further, the linking of precise lexicons in many languages to GOLD will allow it to be developed further into more specific types.

Apart from NLP, how can the database be used? In the short term our database is intended to provide annotators and grammar developers with a clear picture of the current status of the treebank and the grammar. In the long term, we expect to create successively better approximations of the languages described, as long as the deep linguistic broad coverage grammar describes its language's syntax and semantics precisely. Consequently, the database would be of use to anyone who needs an accurate description of a language. Language teachers can use its detailed descriptions of word usages, the links to other words, and the real examples from the treebank to show students subtle differences among words that look the same but are grammatically different. Lexicographers can take advantage of its comprehensiveness and the real examples to compile a dictionary that contains full linguistic explanations.

The confidence in the linguistic descriptions is based on the combination of the precise grammar linked to the detailed treebank. Each improves the other through the treebank annotation and grammar development cycle, and they are further enhanced by the incorporation of linguistic description of the phenomena.

5 Related Work

The ParGram consortium [27] are producing a family of precise grammars based on Lexical-Functional Grammar. Because the grammars are propriety, there is no attempt to make the documentation freely available. On the other hand the collaboration is close, and the grammars are kept harmonized at biannual meetings.

Earlier work with collaborative development of HPSG grammars has focused on integrating testing (regression testing and treebanking) into the grammar development cycle [28,29]. We are extending this work by (a) linking the grammars to external resources and (b) making the information more accessible. For example, we take the test sets, proposed originally for regression testing, then used for treebanking, and make them available as a corpus, indexed by word and lexical type.

Hypertextual Grammar development [30] attempts to support grammar collaboration by documenting grammars. They suggested creating the documentation in the same file along with the grammar, in the style of *literate programming*[31]. This is an attractive approach, especially for grammars that change constantly. However, we prefer the flexibility of combining different knowledge sources (the grammar, treebank and linguistic description, in addition to external resources).

The Montage project [32] aims to develop a suite of software whose primary audience is field linguists working on underdocumented languages. Among their tasks is to facilitate traditional grammatical description from annotated texts by means of one of their products, the Grammar export tool. Although in the paper there is little explicit detail about what the "traditional grammatical description" is, they seem to share a similar goal with us: in the case of Montage, making grammatical knowledge assumed in underdocumented languages explicit, while in our case making lexical types assumed in the treebank and the computational grammar understandable to humans. Also, some tools they use are used in our project as well. Consequently, their process of grammatical description and documentation looks quite similar to ours. The difference is that their target is underdocumented languages whose grammatical knowledge has so far not been made clear enough, while we target any language whose computational implementation is so large and complex as to be difficult to fully comprehend.

The Language Grid is a new initiative to increase the accessibility and usability of online language services[33]. The Lextype DBs make linguistic information *accessible*, as they are based on the open-source DELPH-IN grammars, and *usable* by abstracting the information in each language's grammar and treebanks and presenting it with a common interface, linked by a standard upper ontology. We hope to take advantage of the Language Grid's language service layer to automatically link the Lextype DBs to existing lexical resources in each language. Having a common API will make it possible to do this automatically each time we create a snapshot. We also hope that the Lextype DBs themselves can serve as useful resources within the grid.

Regarding inter-lexicon connection described in §4, [34] discuss the development of Global Grid, where WordNets of various languages are interconnected, and its challenges. They are more ambitious in that they try to establish mappings between words at the synset level and to understand cultural commonalities and differences that are encoded in languages.

6 Future Work

We are currently experimenting with moving some of the information (in particular the type name and criteria) into the actual grammar files, in the same way as [30]. This would make it easier to keep the information in sync with the actual grammar.

We have created lexical type databases for Japanese and English. However, the importance of such a database certainly holds for any large scale deep grammar. We are now integrating our software more directly into DELPH-IN tools in order to make the Lextype DBs available for groups working with other languages.

Another way we would like to improve the system is to add examples from other corpora by exploiting the links to other resources. For example, if the lexical type link is close enough, we can get examples using the Juman parts of speech, which are exemplified in the Kyoto Corpus [18].

7 Conclusion

We have constructed a web-based framework for supporting collaborative multilingual grammar and treebank development in which developers are distributed around the world. Our framework, which we call the Lextype DB, tells developers around the world about lexical types of the grammar and treebank they are developing. Lexical types can be seen as very detailed parts-of-speech and are the essence for the two important points just mentioned.

We have applied the Lextype DB to grammars and treebanks of Japanese and English, in the near future we plan to make the framework available to other grammar developers, and allow them to create their own Lextype DBs.

References

1. Bond, F., Nichols, E., Fujita, S., Tanaka, T.: Acquiring an Ontology for a Fundamental Vocabulary. In: 20th International Conference on Computational Linguistics (COLING-2004), Geneva, pp. 1319–1325 (2004)
2. Baldridge, J., Osborne, M.: Active learning for HPSG parse selection. In: Proceedings of the 7th Conference on Natural Language Learning (2003)
3. Toutanova, K., Manning, C.D., Flickinger, D., Oepen, S.: Stochastic hpsg parse disambiguation using the redwoods corpus. Research on Language and Computation 3(1), 83–105 (2005)
4. Bond, F., Copestake, A., Flickinger, D., Oepen, S., Siegel, M.: Open source machine translation with delph-in. In: Proceedings of The 10th Machine Translation Summit (2005)
5. Flickinger, D.: On building a more efficient grammar by exploiting types. Natural Language Engineering 6(1), 15–28 (2000)
6. Kim, J.B., Yang, J.: Parsing korean case phenomena in a type-feature structure grammar. In: Gelbukh, A. (ed.) CICLing 2005. LNCS, vol. 3406, pp. 60–72. Springer, Heidelberg (2005)
7. Crysmann, B.: Relative clause extraposition in German: An efficient and portable implementation. Research on Language and Computation 3(1), 61–82 (2005)
8. Siegel, M.: HPSG analysis of Japanese. In: Wahlster, W. (ed.) Verbmobil: Foundations of Speech-to-Speech Translation, pp. 265–280. Springer, Heidelberg (2000)
9. Pollard, C.J., Sag, I.A.: Head-Driven Phrase Structure Grammar. University of Chicago Press, Chicago (1994)
10. Sag, I.A., Wasow, T., Bender, E.M.: Syntactic Theory: A Formal Introduction, 2nd edn. CSLI Publications, Stanford (2003)
11. Copestake, A., Flickinger, D., Pollard, C., Sag, I.A.: Minimal Recursion Semantics. An introduction. Research on Language and Computation 3(4), 281–332 (2005)
12. Bender, E.M., Flickinger, D., Oepen, S.: The grammar matrix: An open-source starter-kit for the rapid development of cross-linguistically consistent broad-coverage precision grammars. In: Proceedings of the Workshop on Grammar Engineering and Evaluation at the 19th International Conference on Computational Linguistics, Taipei, Taiwan, pp. 8–14 (2002)
13. Sadakane, K., Koizumi, M.: On the nature of the "dative" particle *ni* in Japanese. Linguistics 33, 5–33 (1995)
14. Copestake, A.: Implementing Typed Feature Structure Grammars. CSLI Publications, Stanford (2002)
15. Bond, F., Fujita, S., Hashimoto, C., Nariyama, S., Nichols, E., Ohtani, A., Tanaka, T., Amano, S.: The Hinoki Treebank — A Treebank for Text Understanding. In: Proceedings of the First International Joint Conference of Natural Language Processing, pp. 554–559 (2004)
16. Siegel, M., Bender, E.M.: Efficient Deep Processing of Japanese. In: Proceedings of the 3rd Workshop on Asian Language Resources and International Standardization, Taipei, Taiwan (2002)

17. Matsumoto, Y., Kitauchi, A., Yamashita, T., Hirano, Y., Matsuda, H., Takaoka, K., Asahara, M.: Morphological Analysis System ChaSen version 2.2.1 Manual. Nara Institute of Science and Technology (2000)
18. Kurohashi, S., Nagao, M.: Building a Japanese parsed corpus — while improving the parsing system. In: Abeillé, A. (ed.) Treebanks: Building and Using Parsed Corpora, pp. 249–260. Kluwer Academic Publishers, Dordrecht (2003)
19. Ikehara, S., Shirai, S., Yokoo, A., Nakaiwa, H.: Toward an MT system without pre-editing – effects of new methods in ALT-J/E–. In: Third Machine Translation Summit: MT Summit III, Washington DC, pp. 101–106 (1991), http://xxx.lanl.gov/abs/cmp-lg/9510008
20. Breen, J.: Jmdict: a japanese-multilingual dictionary. In: Coling 2004 Workshop on Multilingual Linguistic Resources (2004)
21. Miyazaki, M., Shirai, S., Ikehara, S.: gengo katēsetsu-ni motozuku nihongo hinshi-no taikēka-to sono kōyō [a Japanese syntactic category system based on the constructive process theory and its use]. Journal of Natural Language Processing 2(3), 3–25, 1995 (in Japanese)
22. Grishman, R., Macleod, C., Myers, A.: COMLEX syntax: Building a computational lexicon. In: 15th International Conference on Computational Linguistics: COLING-94, Kyoto, Japan, vol. 1, pp. 268–272 (1994)
23. Krieger, H.U., Schafer, U.: $\mathcal{TDL}$ — a type description language for constraint-based grammars. In: Proceedings of the 15th International Conference on Computational Linguistics (1994)
24. Copestake, A., Lambeau, F., Waldron, B., Bond, F., Flickinger, D., Oepen, S.: A lexicon module for a grammar development environment. In: 4th International Conference on Language Resources and Evaluation (LREC 2004), Lisbon, vol. IV, pp. 1111–1114 (2004)
25. Oepen, S., Flickinger, D., Toutanova, K., Manning, C.D.: LinGO Redwoods: A Rich and Dynamic Treebank for HPSG. In: Proceedings of The First Workshop on Treebanks and Linguistic Theories, Sozopol, Bulgaria, pp. 139–149 (2002)
26. Siegel, M.: JACY, A Grammar for Annotating Syntax, Semantics and Pragmatics of Written and Spoken Japanese for NLP Application Purposes. ms (2006)
27. Butt, M., King, T.H., no, M.E.N, Segond, F.: A Grammar Writer's Cookbook. CSLI publications, Stanford (1999)
28. Oepen, S., Bender, E.M., Callmeier, U., Flickinger, D., Siegel, M.: Parallel distributed grammar engineering for practical applications. In: Proceedings of COLING 2002 Workshop on Grammar Engineering and Evaluation, Taipei, Taiwan, pp. 15–21 (2002)
29. Oepen, S., Flickinger, D., Bond, F.: Towards Holistic Grammar Engineering and Testing. Grafting Treebank Maintenance into the Grammar Revision Cycle. In: Su, K.-Y., Tsujii, J., Lee, J.-H., Kwong, O.Y. (eds.) IJCNLP 2004. LNCS (LNAI), vol. 3248, Springer, Heidelberg (2005)
30. Dini, L., Mazzini, G.: Hypertextual Grammar Development. In: Computational Environments for Grammar Development and Linguistic Engineering, Madrid, ACL, pp. 24–29 (1997)
31. Knuth, D.E.: Literate Programming. CSLI Publications, Stanford, CA (1992)
32. Bender, E.M., Flickinger, D., Good, J., Sag, I.A.: Montage: Leveraging Advances in Grammar Engineering, Linguistic Ontologies, and Mark-up for the Documentation of Underdescribed Languages. In: Proceedings of the Workshop on First Steps for the Documentation of Minority Languages: Computational Linguistic Tools for Morphology, Lexicon and Corpus Compilation, LREC2004, Lisbon, Portugal (2004)

33. Ishida, T.: Language grid: An infrastructure for intercultural collaboration. In: IEEE/IPSJ Symposium on Applications and the Internet (SAINT-06), pp. 96–100, 2006(keynote address)
34. Fellbaum, C., Vossen, P.T.: Connecting the universal to the specific: Towards the global grid. In: Ishida, T., Fussell, S., Vossen, P.T. (eds.) IWIC 2007. LNCS, vol. 4568, Springer, Heidelberg (2007)

Collaborative Platform for Multilingual Resource Development and Intercultural Communication

Virach Sornlertlamvanich[1], Thatsanee Charoenporn[1], Kergrit Robkop[1], and Hitoshi Isahara[2]

[1] Thai Computational Linguistics Laboratory, NICT Asia Research Center, Thailand
{virach,thatsanee,kergrit}@tcllab.org
[2] National Institute for Communications Technology, Japan
isahara@nict.go.jp

Abstract. In the present borderless information society, we need a lot of fundamental linguistic tools as well as the standard reference resources to facilitate our daily communications across the languages and cultures for better understanding or smoothing the communications. Online collaborative works are efficiently conducted among expert groups via many existing services such as Sourceforge, Wiki or Weblog. However, in the process of multilingual resource development and intercultural communication we still need to fulfill the requirements in well-structured design of the database, and communication tools that provide necessary linkages between records of intention to particular assertions, and functions to realize selectional preference in case that there are more than one assertion. In this paper, we propose a new platform, called Knowledge Unifying Initiator (KUI). We conducted a study on multilingual medical text collaborative translation and the initiative in Asian WordNet development to evaluate our proposed platform.

Keywords: multilingual resource, intercultural communication, collaborative translation, WordNet.

1 Introduction

Nowadays, the Internet plays a important role in our daily life communication. It can substitute almost all of our typical communication method i.e. letter, fax, and sometimes telephone in the case that we want to avoid direct speech which may interrupt our partner's inconvenient time. This handy communication environment can also be extended to create a web community where participants can efficiently make use of their available time to share their experiences. In the web community, we are aware of its global scale of communication though the English language occupies the majority of the online population. In this paper, we proposed a web based communication tool for building a web community. The community members should be able to work collaboratively to create community knowledge disregarding the difference of the languages and be able to get support from others in order to make a community decision. Knowledge Unifying Initiator (KUI) [2], [7] was developed to maintain the process of knowledge creation by providing four possible

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 91–102, 2007.

stages of participation, namely (1) new task, to allow a participant to initiate a task, (2) opinion, to allow a participant to post his own opinion, (3) localization, to allow a participant to bring in a new knowledge into the community by translation, and (4) public-hearing, to allow a participant to post a draft of concept for conceptualizing the knowledge. A community decision on an opinion in any stages can be made by poll taking from the participants. Moreover, an online chat was also provided to keep record of discussion for each topic. These features of KUI were preliminarily integrated and evaluated in a task of Intercultural Collaboration Experiment (ICE) 2003 [1], [6] before extensive development in the current version of KUI. In the experiment, we selected a task of translation of a medical text, i.e. English emergency diagnosis phrases from MedSLT [4], for forming a web community to collaboratively translate the original English text into their native languages. The participants discussed to share their understanding of the original text to conduct a better translation by using the provided chat function. In the same time, they could refer to other translation which could help in improving their own translation. As a result, the multi-lingual translated medical text was produced and the behavior of participants in the web community was studied by analyzing the chat log. We are also providing a new task for Asian WordNet development by preparing records of English WordNet [5], [9] entries together with the word information to help discriminating the meaning of the head word.

The paper is organized in the following way. Section 2 explains the design of KUI for collaborative multilingual resource development. Section 3 describes the experiment of using KUI for medical text translation task. Section 4 discusses the nature of communication by analyzing the chat log. Section 5 describes the potential in adopting KUI for Asian WordNet development.

2 Collaborative Platform for Multilingual Resource Development

We developed KUI (Knowledge Unifying Initiator) for a knowledge development supporting tool of a web community. In this paper, we implemented KUI to be a Knowledge User Interface for a collaborative translation task of a medical text, MedSLT. The evaluation of KUI was also conducted by analyzing the result of communication in the chat log. Actually, KUI is a platform to unify the various thoughts created by following the process of thinking, i.e. (1) new task, to allow a participant to initiate a task, (2) opinion, to allow a participant to post his own opinion, (3) localization, to allow a participant to bring in a new knowledge into the community by translation, and (4) public-hearing, to allow a participant to post a draft of concept for conceptualizing the knowledge. The process of thinking is done under the selectional preference simulated by voting mechanism in the case that many alternatives occur.

In this section, we describe the concept behind KUI, the knowledge development process, various features in KUI, and the implementation of KUI for the collaborative medical text translation.

2.1 What is KUI?

Knowledge Unifying Initiator or KUI is a GUI for knowledge engineering, in other words Knowledge User Interface (KUI) [2], [7]. It provides a web interface accessible for pre-registered members only for the accountability reason. An online registration is offered to manage the account by profiling the login participant in making contribution to the community. A contributor can comfortably move around in the virtual space from desk to desk to participate in a particular task. A login member will be assigned to a desk when a participation task is defined. Members can then participate in the chat group of the same desk. A desk can be a meeting place for collaborative work that needs discussion through the chat function, or allow a contributor to work individually by using the message slot to record each own opinion. The working space can be expanded by closing the unnecessary frames so that the contributor can concentrate on a particular task. All working topics can also be statistically viewed through the provided tabs. These tabs help contributors to understand KUI in the aspects of the current status of contribution and the available tasks. A web community can be formed to create a domain specific knowledge efficiently through the features provided by KUI. These KUI features fulfill the process of human thought to record the knowledge.

In addition, KUI also provides a KUI look up function for viewing the composed knowledge. It is equipped with a powerful search and statistical browse in many aspects. Moreover, the chat log is provided to learn about the intention of the knowledge composers. We frequently want to know about the background of the solution for better understanding or to remind us about the decision, but we cannot find one. To avoid the repetition of a mistake, we systematically provide the chat log to keep the trace of discussion or the comments to show the intention of knowledge composers.

2.2 Knowledge Development in KUI

Adopting the concept of Open Source software development, we will be possibly able to develop a framework for domain specific knowledge development under the web community environment. Sharing and collaboration are the considerable features of the framework. The knowledge will be finally shared among the communities by receiving the consensus from the participants in each step. To facilitate the knowledge development, the process is deliberated into four steps (Sornlertlamvanich, 2006).

- *New Task*

A new task can be posted to draw intention from participants. The only selected tasks by a major vote will then be proceeded for further discussion in the requested type of task i.e. Opinion Poll or Localization or Public-Hearing.

- *Opinion Poll*

The selected task is posted to call for opinions from the participants in this step. Opinion poll is conducted to get the population of each opinion. The result of the opinion poll provides the variety of opinions that reflects the current thought of the communities together with the consensus to the opinions.

- *Localization*

Translation is a straightforward implementation of the localization. Collaborative translation helps producing the knowledge in multiple languages in the most efficient way. Multi-lingual texts are generated in this type of task.

- *Public-Hearing*

The result of discussion will be revised and confirmed by gathering the opinions to develop the final draft of the proposal. Suggestions for revision are ranked according to the vote. The author may consider the weight of suggestion to make decision on the final revision.

The developed knowledge is started from posting 'New Task', participants express their supports by casting a vote. Upon a threshold the ' New Task ' is selected for conducting a poll on 'Opinion', or introducing to the community by 'Localization', or posting a draft for 'Public-Hearing' to gather feedbacks from the community. The transition from 'Opinion' to either 'Localization' or 'Public-Hearing' occurs when the 'Opinion' has a concrete view for implementation. The discussion in 'Localization' and 'Public-Hearing' is however interchangeable due to purpose of implementation whether to adopt the knowledge to the local community or to get feedbacks from the community.

The knowledge creating is managed in 4 different categories corresponding to the stage of knowledge. Each individual in the community casts a vote to rank the appropriateness of solutions at each category. The community can then form the community knowledge under the 'Selectional Preference' background.

2.3 Features in KUI

Poll-Based Opinion or Public-Hearing. A contributor may choose to work individually by posting an opinion e.g. localization, suggestion etc., or join a discussion desk to conduct 'Public-Hearing' with others on the selected topic. The discussion can be conducted via the provided 'Chat' frame before concluding an opinion. Any opinions or suggestions are committed to voting. Opinions can be different but majority votes will cast the belief of the community. These features naturally realize the online collaborative works to create the knowledge.

Individual or Group Work. Thought may be formed individually or though a concentrated discussion. KUI facilitates a window for submitting an opinion and another window for submitting a chat message. Each suggestion can be cast through the 'Opinion' window marked with a degree of its confidence. By working individually, comments to a suggestion can be posted to mark its background to make it better understanding. On the other hand, when working as a group, discussions among the group participants will be recorded. The discussion can be resumed at any points to avoid the iterating words.

Record of Intention. The intention of each opinion can be reminded by the recorded comments or the trace of discussions. Frequently, we have to discuss again and again on the result that we have already agreed. Misinterpretation of the previous decision is also frequently faced when we do not record the background of decision. Record of intention is therefore necessary in the process of knowledge creation. The knowledge interpretation also refers to the record of intention to obtain a better understanding.

Selectional Preference. Opinions can be differed from person to person depending on the aspects of the problem. It is not always necessary to say what is right or what is wrong. Each opinion should be treated as a result of intelligent activity. However, the majority accepted opinions are preferred at the moment. Experiences could tell the preference via vote casting. The dynamically vote ranking will tell the selectional preference of the community at each moment.

2.4 KUI for Collaborative Translation Task

In this collaborative text translation, participants of different mother language work online as a virtual group by using KUI. After registering the system, KUI automatically provides a group of discussion for each task. The group consists of participant from different languages. Multi groups operate in parallel. Before translating, they are encouraged to discuss by the provided chat function about the topic in question, system, personal information and so on.

3 Experiment

According to the purpose of developing a multi-lingual medical text, we set an experiment for online collaborative translation task in the ICE experiment [1], [6]. The source text used in the experiment is a collection of English emergency diagnosis phrases from MedSLT, an Open Source project for developing a medical speech translation system [4]. The translation task is done collaboratively online via KUI interface. The volunteer translators join the discussion group to translate the source text (English) into their native language sentence by sentence. Each group can be participated by translators of more than one language. During the translation task, they all are encouraged by the group communication to build their own community.

As a result, we obtain both the translated medical text and "chat log" which is considered to be the background intention of translation. This chat log will later be a resource for further analysis on cross language communication.

3.1 Medical Text

As described in the previous section, the source text for the translation is the English emergency diagnosis phrases from MedSLT which doctors suspect that patients may be suffering from. These phrases includes the range of utterances of standard examination question about chest pain and factors that increase or decrease such pain, which can be accomplishedly communicated by one or two word responses or gestures [8]. The patterns of the question utterances are grammatically enough to ask about most domain concept in a natural way.

In terms of sentence pattern, the questions are usually limited to be the basis form of the followings. Table 1 shows the type of utterance.

Do you … ?, Have you… ?
How long … ? How (usually, often, ever, …) … ? When … ?
or
Does (it) cause (a symptom)?

Table 1. Type of utterance

Type of Utterance	Sample
Sentence	Have you had pain for weeks? How long do your headaches usually last? Do you ever have chest pain in the morning? Do your chest pains appear at night is the pain gradual?
Phrase	muscle aches after meals high blood pressure heat cheese

Currently, there are 915 utterances in the experiment including noun phrases, verb phrases and simple sentences.

3.2 Participation

The volunteer participants are to translate the English medical text into their native languages, by using KUI. They act as a virtual group and participate in the translation

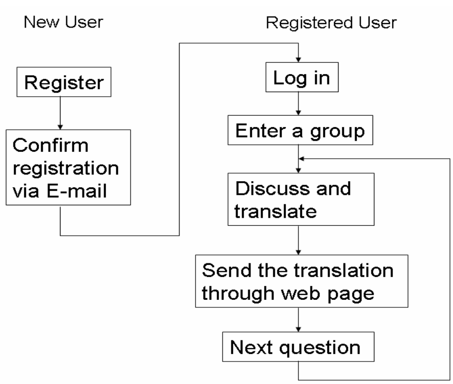


Fig. 1. Participation work flow of collaborative translation task

via this web interface. With different backgrounds and degrees of translation abilities, they, therefore, are encouraged to chat, discuss or exchange their opinion while translating each utterance. The communication is not only for getting to know each other, but also for better understanding of the utterance before translation. Figure 1 shows the participation work flow.

There were 3 tracks in the experiment. Each consisted of participants from two different languages as shown in Table 2. Each group was assigned automatically to translate different questions into their languages.

Table 2. Participation in the ICE experiment

Country pair	Language pair	No. of Group	No. of Participant
Thailand – China	Thai – Chinese	2	4
Thailand – Japan	Thai – Japanese	5	10
Thailand – Korea	Thai – Korean	3	6
Total		10	20

And after two days of experiment, 539 utterances were translated from these 20 participants.

4 Analysis of Chat Log

From the chat log, we can divide the topics on which participants talk together during the experiment into two groups. One is the group of topics related to the utterance in question and another is the group of topics related to personal data enquiry.

It is noted that the conversation texts are not so grammatical. One possible reason is that the conversations were done promptly and the addressers did not pay so much attention on the texts but rather on the quick responding time. This also shows that non-native English speakers are not quite familiar with the expression and they choose to communicate in short using a common text straight to the meaning.

Lacking of revision time in writing the conversation text, we can find the influence of the culture and the mother tongue of the addressers. Some particles are seen added to the sentence. For example, a particle "na" at the end of sentences such as "but i still in the medicine na" in case of Thai addressers. Copula such as verb to be and articles such as "the" are frequently omitted because there is no such kind of word used in the expression. Noun-verb agreement for example in "that because it take time to take you messages from Server" is often neglected. These can be seen in Thai, Japanese and Chinese addressing texts. Except for the common usage mistakes in such as misspelling, capitalization, omission etc., Table 3 shows the analysis result of the influence of culture and language background of addressers.

Being interested in the culture and language background, we decided to show the original text without any modification for thorough understanding.

Table 3. Type of grammatical errors influenced by the culture and language background

Type of grammatical error	Example
Omission	
- Punctuation mark	what do you think about better how about your name
- Article	sometime we translate same sentence
Misuse	
- Copula be	system is delay i'm write in oppose meanging...yeap gogo what's your name mean..in thai?
- Subject-verb agreement	better mean... worse?? headache decrease It seem my computer don't reset
- Word class	Type so quick but typing quick doens't make time follow quick
- Tense	now it came back again.
Addition	
- Preposition	the ideal thing is, that we help to each other
- Particle	go ne? but I still in medicine na
- Mixture of mother tongue	hello p'som chai mai

Topics related to the utterance in question. The topics related to the sentence in question can be classified into four sub-topics as followings.

- Requiring more information about the system, KUI

There are some questions asking about how to work with KUI, components and features of KUI. For example:

English text: Does it hurt in the left chest?
Thai1: what does desk 1, desk 4 mean? **Thai2**: It means you do different topics **Thai1**: but I still in medicine na **Thai2**: yeah, medicine has a lot of topics **Thai1**: but i still in the medicine na **Thai1**: all right **Thai2**: if you stay in the same group so you do in the same item in that topic **Thai2**: sorry same 'Desk' not same group **Thai1**: can i use Thai here? **Thai2**: of course, yes **Thai1**: why I don't see my message? **Thai2**: every things in here show in Unicode, you can type every languages **Thai2**: that because it take time to take you messages from Server

- Discussion on the question for detail understanding

The participants are encouraged to discuss to understand the question before translating. This is because the original utterances are defined for a given English or Western context. For example:

English text: Can exertion sometimes cause chest pains?
Japanese1: ok **Japanese2**: if you have same words like 'some times' do you use the translation you did before? **Japanese1**: i am trying to mention the possibility
English text: Can exertion sometimes give you chest pains?
Japanese1: because sometimes means there are occasion that symptom doesn't occur **Japanese2**: that's right.. and also I am not sure the frequency of 'sometimes'
English text: Can your headache be caused by exertion?
Japanese2: for example two or three times par day... or something like that..but 2 or 3 par day is a lot ! :)
English text: Can exertion cause abdominal pains?
Japanese1: even in the Japanese language definition of 'sometimes' in the Japanese language definition of 'sometimes' = tokidoki difficult **Japanese1**: excuse me. my last sentence was difficult to read **Japanese2**: and also we should take care not to translate into particular symptoms.. **Japanese2**: abdominal pain is general I think **Japanese1**: yes I think it's correct **Thai**: maybe **Japanese1**: it doesn't immediately mean 'cancer' or some particular diseases ..

- Problems found during translation relating to the system. For example:

English text: Does it hurt in the lower abdomen?
Thai: it's hard to use the tools. cursor jumps all the time :(**Thai**: yes **Thai**: when i write chat box, it often jumps to translation box. **Thai**: i have to wait until start icon activates, right? **Thai**: Anybody there? can u read me? **Chinese**: yes, i ok now, go ahead **Chinese**: yes, i ok now, go ahead

- The comment on the utterance

During the translation, some participants expressed their comment concerning to the utterance such as linguistic knowledge, ambiguity of the utterance, etc. For example:

English text: a stabbing pain
Thai1: quite difficult... **Thai1:** Are U OK Ou?
English text: after meals
Thai2: Do you think it's difficult? **Chinese:** has login

Topic related to the personal data enquiry. This kind of topic includes general conversations which are not related to the translation task, such as greeting, persuading to join the experiment next time, background of the participants (nationality, country, address, age and so on). For example:

English text: Is there family history of heart disease?
Thai1: hello guys! **Japanese2**: hello! **Thai1**: you are so on time! **Thai1**: it's just 3.00 p.m. in Thai **Japanese1**: hello! **Japanese2**: yes, here in France, it's 10 in the morning **Thai1**: oh i c **Thai2**: how are u? **Japanese2**: so we are all three or more? **Thai1**: did your guys type anything? **Japanese2**: type? where? here? or translation box? **Thai1**: i see nothing in the translation tab **Japanese2**: ah, translated already? **Thai1**: yes, i just done it **Thai1**: and yes for the upper question, we have 3 person in our group **Japanese2**: hello, yes, I came here to know about the situation

English text: Is your headache caused by bright light?
Thai: have you ever been out of your country? **Japanese**: what do you like? **Thai**: i like movies, traveling, reading, swimming, hanging out with friends **Japanese**: I never go to out of japan **Thai**: i also like the beaches **Japanese**: it is nice **Thai**: yeap here has many nice beaches

5 Initiative in Asian WordNet Development

Adopting the same translation framework, we can make use of KUI for making a translation version of WordNet [5], [9] and also establish a link between the translated WordNet's via the synset ID. We convert the Princeton's WordNet into KUI's internal database. In the Princeton's WordNet 2.1 [9], there are 145,103 nouns, 24,890 verbs, 31,302 adjectives, and 5,720 adverbs differentiated by the synset ID. Since there are a large number of records in noun, we group them into a group of frequent noun and the rest. As a result, we have a group of frequent noun of 28,421 nouns according to the frequency count of noun in Penn Treebank WSJ corpus [3]. The group of frequent noun may have a priority in translation.

In each record, only necessary information for understanding the meaning of the head word will be depicted. Table 4 shows a record of WordNet displayed for translation.

Table 4. A record of WordNet

Car
[Options] **POS**: NOUN **Synset**: auto, automobile, machine, motorcar **Gloss**: a motor vehicle with four wheels; usually propelled by an internal combustion engine

Each record is assigned a synset ID. Translated word will be attached to each synset ID. As a result, a multi-lingual WordNet will be aligned according to the synset ID, and each language WordNet can be generated by substitution of the words in the synset.

Translation of WordNet is an appropriate task for open collaborative work and can yield an effective result. Word by word translation is simple enough for open participation. The provided synset is also a good information source for meaning differentiation and measurement to find the relationship between a particular concept pair.

6 Conclusion

We proposed an efficient online collaborative framework in producing and maintaining the multi-lingual resources. KUI was designed to support an open web community by introducing a voting system and a mechanism to realize the function of selectional preference. It was efficiently introduced to encourage the communication among participants from different language background by providing a task of translating a list of medical text. We collected and analyzed their communication resulting in a set of common segment of the online communication. This will lead to an efficient retrieval system of the response to either the request of knowledge about the system or the topic in question. KUI was also proved to support the collaborative work in producing the multilingual medical text. The translated text will be voluntarily maintained by the online participants under the selectional preference based on the voting function. In future, we plan to provide a task for Asian WordNet development. Each word entry is given the information of part-of-speech, synset and gloss to determine the meaning. Translation of the word entry will establish a link between the synset ID and the translated word. This will result the aligned word entries between languages and each language WordNet can be generated by replacing the translated words in the synset.

References

1. ICE (2006), http://character.kuis.kyoto-u.ac.jp/ICE/
2. KUI (2006), http://www.tcllab.org/kui/
3. Marcus, M.P., Santorini, B., Marcinkiewicz, M.A.: Building a Large Annotated Corpus of English: The Penn Treebank. Computational Linguistics 19(2), 313–330 (1993)
4. MedSLT (2006), http://sourceforge.net/projects/medslt/
5. Miller, G.A.: WordNet: A Lexical Databases for English. Communications of the ACM, 39–41 (November 1995)
6. Nomura, S., Ishida, T., Yasuoka, M.: Open Source Software Development with Your Mother Language: Intercultural Collaboration Experiment 2002. In: HCI International (2003)
7. Sornlertlamvanich, V.: KUI: The OSS-Styled Knowledge Development System. In: Handbook of The 7th AOSS Symposium, Kuala Lumpur, Malaysia (2006)
8. Starlander, M., Bouillon, P., Rayner, M., Chatzichrisafis, N., Hockey, B.A., Isahara, H., Kanzaki, K., Nakao, Y., Santaholma, M.: Breaking the Language Barrier: Machine Assisted Diagnosis using the Medical Speech Translator. In: Proceedings of the XIX Internation Congress of the European Federation for Medical Informatics MIE, Geneva, Switzerland (2005)
9. WordNet (2006), http://wordnet.princeton.edu/

Exploiting Aligned Parallel Corpora in Multilingual Studies and Applications

Dan Tufiş

Research Institute for Artificial Intelligence, Romanian Academy, 13, "13 Septembrie", 050711, Bucharest, Romania
tufis@racai.ro

Abstract. Parallel corpora encode extremely valuable linguistic knowledge, the revealing of which is facilitated by the recent advances in multilingual corpus linguistics. The linguistic decisions made by the human translators in order to faithfully convey the meaning of the source text can be traced and used as evidence on linguistic facts which, in a monolingual context, might be unavailable to (or overlooked by) a computer program. Multilingual technologies, which to a large extent are language independent, provide a powerful support for systematic and consistent cross-lingual studies and allow for easier building of annotated linguistic resources for languages where such resources are scarce or missing. In this paper we will briefly present some underlying multilingual technologies and methodologies we developed for exploiting parallel corpora and we will discuss their relevance for cross-linguistic studies and applications.

Keywords: alignment, annotations, collocations, cross-language studies, disambiguation (POS, WSD), encoding, parallel corpora, multilingual technologies, tagging, wordnets.

1 Introduction

The "world of knowledge", as the virtual space of the internet has rightfully been called, is the conceptual framework where the notion of "digital-divide" has been coined. This phenomenon results from the unequal application of, and access to, information and communication technologies. Narrowing the knowledge gaps between different communities of the world has been, and continues to be a top priority not only for local authorities but for major international organizations as well. For instance, in its 32nd Session (30 September-17 October 2003) the UNESCO General Conference adopted a highly relevant document "RECOMMENDATION ON THE PROMOTION AND USE OF MULTILINGUALISM AND UNIVERSAL ACCESS TO CYBERSPACE" where it is shown that "*multilingualism in cyberspace is of vital and strategic importance to ensure the right to information and cultural diversity*" and that "*everyone and every nation must have an equal opportunity to benefit from cultural diversity and scientific progress, which must remain, more than ever, a basic human right in the emerging information society*"[1].

[1] http://portal.unesco.org/ci/en/ev.php-RL_ID=4969&URL_DO=DO_TOPIC&URL_SECTION=201.html

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 103–117, 2007.

Yet, mere access to the internet does not open the gates to the "world of knowledge". The e-content is expressed in many languages and this is natural to be so. "Language constitutes the foundation of communication between people and is also part of their cultural heritage. For many people, language carries far-reaching emotive and cultural associations and values embedded in vast literary, historical, philosophical and educational heritage. For this reason the users' language should not constitute an obstacle to accessing the multicultural human heritage available in cyberspace" (ibid.).

Among others, the multilingual language technologies are expected to be most instrumental in lowering as much as possible the language and cultural barriers for harmonious and collaborative development of the information society.

The web language services [3], [13], linguistic grids [9], [14], multilingual collaborative and distributed platforms [12], [27], automatic translation, are seen as key technologies, the most promising in fostering the cross-cultural cooperation. Identifying opinions and emotions expressed in texts, one of the hottest current research area, revealed cross-cultural similarities but also disparities which should be very carefully considered for a smooth intercultural communication [18], [25].

In this paper we will describe another extremely useful technology for multilingual processing, namely the word alignment. Word alignment is not a goal in itself, but an enabling technology which serves all the higher level multilingual applications. After describing the necessary text pre-processing and the alignment procedure, we will exemplify a few cases of text alignment exploitation: lexical semantics knowledge acquisition and validation, cross-lingual studies, and transfer of linguistic (syntactic and semantic) annotations in a multi-cultural cooperation program.

2 Parallel Corpora and Textual Alignment

A bitext is a pair of texts in two languages, so that the texts can be considered reciprocal translations. They are called translation equivalents. By extension, a multi-text is a set of multiple language texts, so that each pair of texts represents a bitext. A large collection of bitexts or multi-texts is called a parallel corpus. Knowing that two or more texts are reciprocal translations is useful, but much more useful is detecting the translation equivalence at finer grained levels.

The automatic identification in a parallel corpus of the segments of texts that represent reciprocal translations is a prerequisite for taking advantage of the implicit linguistic and cultural knowledge embedded into the translations. This problem, known as parallel corpus alignment, can be defined at various levels of text segmentation granularity (paragraph, sentence, phrase, word) with different degrees of difficulty. Two segments of texts from a bitext which represent reciprocal [2]translations make a translation unit. A translation unit may contain, in one or both paired languages, one or more textual units (paragraph, sentence, phrase, word) and one distinguishes between the 1:1 and non-1:1 alignment translation units. While at

---

[2] M is the index of the mth sentence in one part of the bitext which is presumably aligned to the the nth sentence (of index N) in the other part of the bitext.

the paragraph granularity level the non-1:1 alignments are exceptional (most aligners assume the number of paragraphs to be the same in the two sides of a bitext, and thus only 1:1 alignments are considered), at the sentence or phrase level they are quite rare (usually no more than 5-10% of the total number of translation units). At the word level, the non-1:1 alignments are more frequent and their number strongly depends on the language pair and on the type of translation (literal versus free translation). Another source of increased difficulties for fine-grained alignments is that while at the paragraph and (to a large extent) sentence level the ordering of the textual units is preserved in both sides of a bitext (discourse coherence requirement), at the finer grained level this is not true in general (the word or phase ordering being ruled in each language by its syntax).

Depending on the alignment granularity, required accuracy, and the purpose of the alignment, the input textual data might need pre-processing steps in all languages of the parallel corpus (e.g. sentence splitting, tokenization, POS-tagging and lemmatization) or at least in one of the languages of the corpus (e.g. chunking, dependency linking/parsing, and word sense disambiguation).

3 Preprocessing Steps

Text segmentation. The first pre-processing step in most NLP systems deals with text segmentation. In our processing chain this step is achieved by a modified version (much faster) of the multilingual segmenter developed within the MULTEXT project which has tokenization resources for many western European languages, further enhanced in the follow up MULTEXT-EAST project with corresponding resources for Bulgarian, Czech, Estonian, Hungarian, Romanian and Slovene. Our segmenter is able to recognize paragraphs, sentence and clause boundaries, dates, numbers and various fix phrases, and to split clitics or contractions (where the case). We significantly updated the tokenization resources for Romanian and English (the languages we have been most interested in lately).

Sentence alignment. We developed a sentence aligner [6] inspired by Moore's program [21] which removes its 1:1 alignment restriction, the assumption on the monotonic ordering of the sentences in the two languages, as well as the upper limit on the number of sentence-pairs that can be aligned. It has a comparable precision but a better recall than Moore's aligner.

The sentence aligner consists of a hypothesis generator which creates a list of plausible sentence alignments from the parallel corpus and a filter which removes the improbable alignments. The hypothesis generator uses a character based preliminary sentence aligner, similar to Church and Gale's CharAlign [11], which creates a list of possible alignments, represented as pairs of sentence identifiers <N, M>. Each pair of this list is supplemented by pairs obtained by local variations of the sentence indexes <M±k, N±k> with k ranging from 1 to a user-defined upper limit (the default value is 1[3]).

[3] For this case, besides the pair <N, M> the alignment candidates list will also contain the pairs <N-1,M>, <N+1,M>, <N,M-1>, <N,M+1>, <N-1,M+1>, <N+1,M-1>,.

The filter is an SVM binary classifier [8] initially trained on a Gold Standard. The features of the initial SVM model are: the word sentence length, the number of non-word tokens, and the rank correlation for the first 25% of the most frequent words in the two parts of the training bitext. This model is used to preliminary filter alignment hypotheses generated from the parallel corpus. The set of the pairs of sentences that remained after this filtering is used as the input for an expectation maximisation algorithm which builds a word translation equivalence table by a similar approach to the IBM model-1 procedure [4]. The SVM model is rebuilt (again from the Gold Standard) this time including, as an additional feature, the number of word translation equivalents existing in the sentences of a candidate alignment pair. This new model is used by the SVM classifier for the final sentence alignment of the parallel corpus.

POS-tagging. It is generally known that the accuracy of POS-tagging depends on the quality of the language model underlying the morpho-lexical processing, which, on its turn, is highly dependent on the quality and quantity of the training data and on the tagset of the language model. For languages with a productive inflectional morphology the morpho-lexical feature-value combinations may be very numerous, leading to very large tagsets with unavoidable training data sparseness threat. The lack of sufficient training data affects the robustness of the language models which, consequently, will generate an increased number of tagging errors at the run time. To cope with the tagset cardinality problem we developed the tiered-tagging methodology [34] and implemented it using the TnT trigram HMM tagger [2]. The methodology involves the use of a reduced hidden corpus tagset, automatically constructed from the large targeted lexical tagset, and a procedure to map back the reduced tagset into the large one, used in the final annotated text. The two tagsets (the lexical and corpus tagsets) are related by a subsumption relation. When the reduction of the cardinality of the large tagset is information lossless (that is just redundancy elimination) the mapping from the reduced tagset to the large one is deterministic and it is simply ensured by looking up a wordform dictionary. For tagset reduction with information loss, which ensures a much significant reduction of the lexical tagsets, the recovering of the left out morpho-lexical information, although to a large extent deterministic, requires an additional preprocessing to solve some non-deterministic cases. In the previous version of the tiered tagging approach we used several hand-crafted rules (regular expressions defined over the reduced tagset, with a span of ±4 tags around the ambiguously mapped tags).

Recently, we have re-implemented the tiered tagging methodology, by relying on a combination between an HMM tagger, called TTL [15], which produces also the lemmatization, and a maximum-entropy tagger [5].The HMM tagger works with the reduced tagset while the ME-tagger ensures the mapping of the first tagset onto the large one (the lexical tagset) dispensing on the hand-written mapping rules.

Lemmatization is in our case a straightforward process, since the monolingual lexicons, developed according to MULTEXT-EAST morpho-lexical specifications [7], contain for each word, its lemma and the morpho-lexical tag. Knowing the word-form and its associated tag, the lemma extraction is simply a matter of lexicon lookup for those words that are in the lexicon. For the unknown words, which are not tagged as proper names, a set of lemma candidates is generated by a set of suffix-stripping rules induced from the word-form lexicon. A four-gram letter Markov model (trained on lemmas in the word-form dictionary) is used to choose the most likely lemma.

Chunking. By means of a set of language dependent regular expressions defined over the tagsets, our chunker accurately recognizes the (non-recursive) noun phrases, adjectival/adverbial phrases, prepositional phrases and verb complexes (analytical realization of tense, aspect mood and diathesis and phrasal verbs) both for Romanian and English.

Word Alignment. The word alignment [23], [24] of a bitext is an explicit representation of the pairs of words $<w^i_{L1}\ w^j_{L2}>$ occurring in the same translation units that represent mutual translations (called translation equivalence pairs). Either of w^i_{L1} or w^j_{L2} may be NULL (this is the case of *null alignments* where one word in one part of the bitext is not translated in the other part). When w^i_L , w^j_{L2} or both appear in several pairs of the same translation unit they correspond to *multi-word expression alignments*.

The input raw texts, pre-processed as described in the previous section, are fed into the word alignment engine, called COWAL [31], [32] which is a wrapper of two stand-alone aligners (YAWA and MEBA). COWAL merges the alignments produced by each stand-alone aligner and then uses a trained SVM classifier to prune the unlikely alignment links. The classifier is based on the LIBSVM kit [8] used with the default parameters (C-SVC classification and radial basis kernel function). The classifier was trained with positive and negative hand-validated examples of word alignment links.

The usefulness of the aligner combination has been convincingly demonstrated on the occasion of the Shared Task on Word Alignment organized by the ACL2005 Workshop on “Building and Using Parallel Corpora: Data-driven Machine Translation and Beyond” [20]. We participated (on the Romanian-English track) with the two standalone aligners and the combined one [32]. Out of the 37 competing systems, MEBA was rated the 20th and TREQ-AL, (the former version of YAWA), was rated the 21st, but COWAL, their combination, was the winning system.

Table 1. Combined alignment

Aligner	Precision	Recall	F-measure
YAWA	88.80%	74.83%	81.22%
MEBA	92.15%	73.40%	81.71%
COWAL	87.26%	80.94%	83.98%

Meanwhile, both stand-alone aligners have been improved (see Table 1) in various ways and trained on more data, but the combined aligner still performs better than both of them.

COWAL is now embedded into a larger platform, called MTkit that incorporates the tools for bitexts pre-processing, a graphical interface that allows for comparing and editing different alignments, as well as a word sense disambiguation module (described in the next section).

4 Exploiting the Alignments

In the following, there will be shown a few examples of how we used the word alignments. The most obvious application of the word alignment is building translation lexicons [30]. The aligned corpora we worked with were the Ro-En sub-corpus of the „1984" multilingual [7] corpus (about 110,000 tokens per language), a partial translation in Romanian of SemCor2.0 (about 177,000 tokens per language), the journalistic parallel corpus Ro-En) used in the word-alignment competition at ACL2005 [20] (about 1,000,000 words per language) and the four-language (English-Romanian-French-German) sub-corpus of the 21-language parallel corpus Acquis Communautaire[4] (about 8 million tokens per language). In our experiments we were interested only in open-class words, the alignment of which is significantly more accurate than the „all words" alignment. This is to say that most of the alignment errors in the evaluation shown in Table 1 were related to functional words and punctuation. From the parallel En-Ro corpora we extracted a large translation lexicon (about 500,000 entries) used in various applications, some of which will be mentioned in the next sections.

4.1 Aligned Wordnets Validation

Once the translation equivalents identified, it is reasonable to expect that the words of a translation pair <w^i_{L1}, w^j_{L2}> share at least one conceptual meaning stored in an interlingual sense inventory. In the BalkaNet project [35] we used the Princeton WordNet (PWN) [10] as an interlingual index [37]. Based on the interlingually aligned wordnets, obtaining the sense labels for the words in a translation pair is straightforward [16]:

a) one has to identify for w^i_{L1} the synset S^i_{L1} and for w^j_{L2} the synset S^j_{L2} so that S^i_{L1} and S^j_{L2} are projected over the same concept. The index of this common interlingual concept (ILI) is the sense label of the two words w^i_{L1} and w^j_{L2}.
b) if no common interlingual projection will be found for the synsets to which w^i_{L1} and w^j_{L2} belong, the senses of the two words will be given by the indexes of the most similar interlingual concepts corresponding to the synsets of the two words. The semantic-similarity score is computed as SYM (ILI_1, ILI_2) = 1/1+k where k is the number of PWN links from ILI_1 to ILI_2 or from both ILI_1 and ILI_2 to the nearest common ancestor.

In case none of the two cases above holds, then it is very likely that there are some problems which can be categorized as follows:

i) the translation pair is wrong (either because of human translator or because of the word aligner), so it is natural not to find any ILI matching for the two words of the pair;
ii) one or both words do not have implemented the relevant senses;
iii) one of both words are missing from the relevant existing synsets;

[4] http://langtech.jrc.it/JRC-Acquis.html

iv) one or both synsets to which the words of the current translation pair belong are not correctly linked to the relevant ILI;
v) the two words in the current translation pair have different POS. Since all the BalkaNet wordnets were aligned to PWN version 2.0 preserving the POS of the synsets, all the cross-pos translations will fit this case.

For the semantic validation of the wordnets created during the BalkaNet project the cases i) and v) were not relevant. The alignment pairs of the ii), iii) and iv) types extracted from the various bilingual sub-corpora of the "1984" corpus were validated by native speakers, with very good command of English. As a result, many synsets were extended with missing literals, the missing synsets were added, and wrong interlingual projections were corrected. The final report of the BalkaNet[5] project gives a detailed quantitative and qualitative account of the errors and incompletenesses that were detected by this procedure (and corrected by each partner).

Since the BalkaNet project finished, we have consistently extended the Romanian wordnet [36] (currently it contains more than 39,000 synsets, and this number is steadily growing). We repeated the semantic validation procedure several times until we haven't noticed any problematic case in the Ro-En sub-corpus of "1984".

Recently, we have applied a slightly modified variant of this procedure, this time using a partial translation in Romanian of the SemCor2.0. The two sets of documents were sentence aligned (8276 aligned sentence pairs), word aligned and the alignment correctness for the closed-class categories was estimated (based on a random set of links) at 98.5%. With this accuracy, the i) problem was not expected to have a relevant contribution to the number of problems that could (and did) show up.

With the word senses available in the English part, the validation procedure of the Romanian wordnet proceeded the following way:

For each pair of aligned words we checked whether the ILI number of the sense marked on the English word corresponded to the one of any synsets containing the Romanian translation equivalent or to any hypo/hypernyms of it.

Out of the 178499 tokens in the English part[6], only 79595 tokens were sense marked-up (the rest were functional words or punctuation) but 3535 were not translated into Romanian and 3694 translation pairs had different part of speech. For the remaining 72366 POS-preserving translated English content words, the sense correspondence was found for 48392 Romanian tokens (66.87%). We thoroughly analyzed the 23934 cases of the unmatched senses and found two main situations:

a) Romanian translation equivalent was not in the Romanian wordnet:	11930 cases
b) The Romanian synset with the ILI of the English word exists, but does not include the Romanian word (incomplete synset) :	12044 cases

While the a) situation was somehow expected (given that the Romanian wordnet contains much less synsets and literals than PWN, leaving room for long-time development efforts for our team), the second situation was worrying and we concentrated our evaluation on that case. As expected, a large number (more than 4000) of the translation equivalence pairs coming under this rubric were alignment

[5] http://www.ceid.upatras.gr/Balkanet/resources.htm
[6] The Romanian part contains 175603 tokens.

errors but still, numerous incomplete synsets (more than 7000) were detected. The extension of all these synsets with the missing literals is one of our current and future activities.

4.2 WordNet-Based Sense Disambiguation

The WSD task can be stated as being able to associate to an word (*w*), ambiguous in a text or discourse, the sense (s_k) which is distinguishable from other senses ($s_1 \ldots s_{k-1}$ $s_{k+1} \ldots s_n$) and is prescribed for that word by a reference semantic lexicon. The task of word sense disambiguation (WSD) requires one reference sense inventory, in terms of which the senses of the target words will be labeled. A meaningful discussion of the performances of a WSD system cannot dispense of clearly specifying the sense inventory it uses, and the comparison between two WSD systems that use different sense inventories is frequently more confusing than illuminating. Essentially, this is because the differences in the semantic distinctions (sense granularities), as used by different semantic dictionaries (sense repositories), make the difficulty of the WSD task range over a large spectrum. For instance, the discrimination of homographs (more often than not having different parts of speech), is much simpler than the metonymic distinctions.

It is straightforward to turn the previous validation procedure into a WSD engine and we claim [33] that state-of-the-art (and even better) performances can be achieved in sense disambiguation of the words in a parallel text, provided the respective bitext is word aligned and aligned wordnets exists for the considered languages.

As mentioned, we used the PWN version 2.0 (PWN2.0) synset identifiers as the reference sense inventory. Through the 1-1 mapping, existing between the synsets in the Romanian wordnet and PWN2.0, the SUMO/MILO [22] and DOMAINS [19] labels became available in our wordnet (as in all the other wordnets which are aligned to PWN2.0). Since both SUMO/MILO and DOMAINS synset labeling is POS insensitive, their use is extremely helpful in assigning senses to the words in a cross-pos translation equivalence pair. For instance, if we consider the En-Ro translation equivalence pairs <*fire*-verb *concediat*[7]-adj> the POS preserving ILI-based alignment between PWN2.0 and RoWN is not really helpful. However, if one uses instead of synset identifiers the SUMO labels, a match would be found since the senses *fire*:4 for the verb and *concediat:1* for the adjective belong to synsets labeled by the same SUMO concept *TerminatingEmployment.* An extra-bonus is that one can infer that the first sense of the adjective *fired* (the translation of *concediat*) is derived from the fourth sense (and no other one) of the verb *fire*. This is a useful type of information which is not yet encoded in any version of PWN.

4.3 Annotation Transfer as a Cross-Lingual Collaboration Task

Having a parallel corpus aligned at the word and phrase level may be the starting point on significant geographical and cross-lingual distribution of the tasks aimed at creating a multiple layer annotated corpus. One can imagine a cross-cultural initiative, which agreed on some parallel corpora (e.g. AcquisCom) containing the languages of interest, and where each partner is willing to annotate his/her language

[7] Concediat (Ro) = fired (En).

part of the multilingual corpora with the information for which the appropriate tools exists (POS, multi-word expressions, sense labels, parse tree annotations, argument/frame structures, etc). Based on the assumption that correcting annotations is easier and cheaper than creating them from scratch, word alignment technology could be used to transfer information from the tokens in one language to their translation equivalents in the other language. The transfer could be controlled by language specific rules, the writing of which is certainly less demanding than the direct annotation.

An example at hand is transferring the word senses. Hand word sense disambiguation of a large text is an extremely labor intensive work, prone to human errors and extremely expensive.

SemCor2.0 is an English corpus, with (most of) the content words being sense disambiguated and carefully validated. It is not surprising that several research teams (see for instance http://multisemcor.itc.it/index.php) decided to translate as much SemCor documents as possible and then to transfer the senses into the translations. As we have shown in section 4.1, we partially translated the SemCor2.0 and used word-alignments to check-out the completeness of our wordnet. We showed that more than 12,000 Romanian words translating the English sense annotated words were absent from our wordnet. By sense transfer, we can extend the target wordnet with these 12,000 synsets. It is true that these synsets are partial (they are mono-literals) and would certainly require extensions with other synonyms but much effort is already saved. Similar considerations apply for the automatic extension of the 7000 incomplete synsets discovered by the experiment we described in section 4.1.

Another type of annotation transfer, more difficult and less reliable, but extremely useful, refers to the syntactic/semantic relations annotated in the source language. When the same type of syntactic/semantic annotation exists in both languages, the annotation transfer allows for annotation validation in one or both languages of the bitext, or provides evidence for corpus-supported comparative/contrastive studies.
In [1] it is described an experiment aimed at assessing the possibility of statistically inducing a dependency grammar for Romanian by semi-automatic transfer of the dependency relations from a parsed English text. The major assumption to be evaluated was the so-called Direct Correspondence Assumption (DCA): checking whether a dependency relation that holds between two English words remains valid between the Romanian translation equivalents in the aligned bitext. In what follows we present the main lines of the strategy employed in this experiment and some preliminary results.

The aligned English-Romanian bitext used for this study was extracted from the *1984* parallel corpus; the English part of the bitext was parsed with the FDG parser [28] by a partner in Wolverhampton and validated by another partner in Iaşi.

From the entire bitext only 1537 sentence pairs (about 25%) were retained for the proper experiment. We discarded very long sentences, the non-1:1 translation units, and those translation units with fewer aligned words than an empirical threshold or containing slang and non-grammatical language (*proles* language).

The relations between English words with a NULL translation equivalent were not taken into account for the evaluation of the transfer accuracy (in Table 2 below, these are counted in the Lost column), while the rest of relations were mechanically transferred into the Romanian part of the bitext.

Table 2. Percentage of correctly transferred relations

No.	Rel.	RO	Lost	EN	Acc. (%)	No.	Rel.	RO	Lost	EN	Acc (%)
1	qn	10	0	12	83.33	11	cc	94	2	155	61.44
2	neg	10	0	13	76.92	12	pm	44	1	75	59.46
3	oc	3	0	4	75.00	13	obj	79	2	137	58.52
4	dat	3	0	4	75.00	14	mod	114	1	201	57.00
5	cnt	8	0	11	72.73	15	ha	41	0	74	55.41
6	ad	25	0	35	71.43	16	cla	8	0	15	53.33
7	pcomp	218	9	316	71.01	17	tmp	23	0	46	50.00
8	det	126	173	355	69.23	18	man	16	0	32	50.00
9	comp	70	1	112	63.06	19	subj	121	72	319	48.99
10	attr	151	4	245	62.66	20	v-ch	35	48	143	36.84

Two experts independently evaluated the validity of the transferred relations with disagreements negotiated and, agreed one way or another.

There were identified three types of relation transfer: for the first type, the transfer is possible and correct without amendments; the second type refers to correct link transfer but incorrect labeling of the links; it needs mapping rules for switching the names of the correct link dependencies (e.g. the rule responsible for an active voice construction in English, translated by a passive voice construction in Romanian switches the obj and subj labels in the target language sentence); the third category of transfers refers to the "lexicalized" dependencies (relations whose governor (rarely the dependent) is instantiated by a specific word) where the transfer is always wrong (both the dependency link and its name), due to the different behavior of corresponding predicates in the considered languages (e.g. like/plăcea).

Table 2 gives information on the correctness of the unconditional transfer for the relations[8] from the source part of the bitext. Due to the enclitic definite articulation in Romanian, half of the English determiners (the occurrences of the) are not explicitly translated and consequently, half of the det relations are lost. The large number of the subj relations that are lost is due to the pro-drop nature of Romanian. One also may notice that this relation has a low correct transfer figure (48.99%) which is correlated with the low correct transfer figure of the relation obj. The simple mapping rule, mentioned before, for dealing with passive/active voice alternation in the aligned sentences would improve with almost 50% the success rate.

We consider these results extremely encouraging, and one of our future research topics will be the design of a set of transfer rules for correcting the role assignment for the dependency links which were correctly transferred (second type, see above). The "lexicalized" dependencies will be collected as they would be detected and stored (with the correct transfer information) as exceptions from the general transfer procedure.

A similar transfer experiment, but this time involving the valency frames existing in the Czech wordnet of the Balkanet project was carried on with Romanian wordnets as target. We used 601 valency frames, kindly offered by the Czech partner in Balkanet and the Czech-Romanian aligned sub-corpus of the *"1984"* parallel corpus. The manual validation of the automatic transfer of the Czech valency frames from the

[8] The phr relation is not included being specific for English phrasal verbs.

Czech verbs to their Romanian translation equivalents revealed a surprisingly high matching (80%), given the differences between Slavic and Romance languages.

5 Collocations Analysis in a Parallel Corpus

Once a parallel corpus has been word-aligned a very interesting cross-lingual study can be achieved in the area of multilingual terminology, multiword expressions and collocational patterns. Within the context of a trilateral project (University Marc Bloch from Strasbourg, IMS Stuttgart University and RACAI) we experimented on a large four-language parallel corpus (En-Ro-Fr-Ge) extracted from the Acquis Communautaire (AcqCom) multilingual corpus. For the word-alignment we used the English text as a hub language and after generating the En-Ro, En-Fr and En-Ge alignments, by transitivity we computed the alignments Ro-Fr, Ro-Ge and Fr-Ge. These last three alignments were combined, as discussed in Section 3, with the corresponding Ro-Fr, Ro-Ge and Fr-Ge alignments directly generated from the parallel corpus, thus obtaining more accurate alignments.

For the texts in each language in this parallel corpus the collocations were independently extracted (RACAI did it for Romanian and English, IMS for German and University Marc Bloch of Strasbourg for French). Our collocation extraction algorithm is similar to Smadja and McKeown's approach [26]. Based on the word alignment of the different bitexts, one could extract the translations in one language of the collocations detected in the other languages. At the time of the writing of this paper, we performed the partial analysis of the collocations in Romanian and English with respect to their translations in English and Romanian respectively. We selected the best scored 20.000 independently extracted collocations in English ($\mathrm{COLLOC_{EN}}$) and Romanian ($\mathrm{COLLOC_{RO}}$). Then, by translation equivalence relations, found by the word aligner, we identified the translations into Romanian of the English collocations ($\mathrm{TR_{RO}}$-$\mathrm{COLLOC_{EN}}$) and the translations into English of the Romanian collocations ($\mathrm{TR_{EN}}$-$\mathrm{COLLOC_{RO}}$).

Given that the corpus contains specific uses and specialized language most of the collocations represent specific multi-word terms and most of them have word by word translation (Member State = Stat Membru, administrative transparency = transparență administrativă, act of accession = act de aderare, enter into force = intra în vigoare, etc.). The vast majority of these collocations were found in the intersection of the sets:

$$\text{SURE-COLLOC}_X = \text{COLLOC}_X \cap \text{TR}_X\text{-COLLOC}_Y \qquad (1)$$

with X=English &Y=Romanian or X=Romanian &Y English respectively.

However, the most interesting collocations were those not found in the previous intersection sets:

$$\text{INTERESTING-COLLOC}_Z = \text{COLLOC}_Z \setminus \text{SURE-COLLOC}_Z \qquad (2)$$

with Z=English or Romanian

The multiword expressions in the lists $\text{INTERESTING-COLLOC}_Z$ were hand validated for termhood, cleaned-up and classified into three major cases:

a) aligner failure to detect the equivalence, due to preprocessing error and its imperfect RECALL (ex: "in vitro", "in vivo"; in Romanian these words were both wrongly tagged and lemmatized)
b) aligner failure to detect the equivalence due to a free human translation of the original text
c) aligner failure to detect the equivalence due to a non-word-by-word translation of the terms (especially those containing light-verbs).

An example of the case b) is given by the following original English text:

"Whereas under Article 6 of the abovementioned Regulation the time when a transaction is carried out is considered as being the date on which occurs event, as defined by Community rules or, in the absence of and pending adoption of such rules, by the rules of the Member State concerned, in which the amount involved in the transaction becomes due and payable." , which was translated as:
"Întrucât, conform art. 6 din regulamentul menţionat anterior, se consideră ca moment al realizării operaţiei data la care intervine faptul generator de creanţă legată de valoarea aferentă operaţiei respective, aşa cum este el definit de reglementarea comunitară sau, dacă aceasta nu există şi urmează a fi adoptată, de reglementarea statului membru interesat."

In this example the scattered English text: *"(the) event. in which the amount.. becomes due and payable"* coresponds to the Romanian term *"situaţie generatoare de creanţe*" (a literal translation would be "(a) situation generating dues")

The last category is the most interesting as it outlines the multiword expressions which, due to their structural differences, are the hardest to translate by a simple-minded word-by-word approach. They range from legalese jargon (e.g. adversely affect <-> *a aduce atingere*[9]; legal remedy <-> *cale de atac*[10]; to make good the damage <-> *a compensa daunele*[11] etc.) to constructions which are language and culture specific: to shake hands <-> *a da mâna*[12]*;* piece of cake <-> *floare la ureche*[13]; Failing to use the exact wording of such a multiword expression, usually, is the major error source for language comprehension/production by language learners, as well as for other human beings in need to communicate but constrained to use a foreign language.

An interesting preliminary contrastive report on the light-verbs based collocations, with a case study of the verb *a face* (to do/make) for the French-Romanian AcqCom data is presented in [29].

We plan to develop a multilingual collocation dictionary, placing the major emphasis on the "hard" collocations (those existing in at least one inventory INTERESTING-COLLOC$_Z$ with Z one of the project languages) providing structural descriptions, translations in all considered languages, morpholexical restrictions on constituents (such as obligatory definiteness/indefiniteness, singular/plural, obligatory case, etc). We aim at a unified description of the collocational patterns in the four languages (with a perspective to extend our work to all the languages represented in

---

[9] A mot-a-mot translation would be *to bring a touch.*
[10] A mot-a-mot translation would be *way to attack.*
[11] A mot-a-mot translation would be *to compensate the damages.*
[12] A mot-a-mot translation would be *to give the hand.*
[13] A mot-a-mot translation would be *a flower at the ear.*

the Acquis Communautaire corpus) and the development of a comprehensive multilingual dictionary, essential for dealing with the hard topic of collocation translation.

6 Web Services

We implemented a NLP web-services platform which currently ensures the basic preprocessing steps (tokenization-including multiword expression recognition, tiered tagging, lemmatization, language identification, sentence alignment, wordnet browsing) for English and Romanian corpora, as well as a search engine for the English-Romanian parallel corpora.

The services are implemented using standard technology (SOAP/WSDL/UDDI) on a dedicated bi-processor server with a reasonable high-speed internet connection (100Mb/s). The NLP web-services will be continuously extended with new services (word alignment, collocation extraction, translation, QA in open domains, summarization, etc). Although most of the present (and near-future) services are available only for Romanian and English, we plan to add as many new languages as possible. The CLARIN initiative (http://www.clarin.eu/), recently included into the European Roadmap for Research Infrastructures[14], has been adhered by more than 60 institutions from 30 European countries and it is supposed to be the major collaborative work environment that will create, adapt and maintain the language resources and tools we could add to the NLP web services platform.

The current web-services were already used for mass data processing by various remote project partners in the ROTEL project (http://rotel.racai.ro), LT4L project (http://www.lt4el.eu/) and the cross-lingual (Romanian-English) CLEF 2006 QA task (http://www.clef-campaign.org/2006/working_notes/CLEF2006WN-Contents.html).

7 Conclusions

The recent advances in NLP technology, demonstrate the tremendous benefits of collecting and adequately encoding large parallel corpora and multilingual semantic lexicons and ontologies. Building, in a concerted way, this kind of resources, for as many languages and as large as possible, should be a constant objective for an internationally established research infrastructure. Initiatives like Global WordNet (http://www.globalwordnet.org/), furthered by Wordnet Grid proposal [9], Language Grid (http://langrid.nict.go.jp/), the previously mentioned CLARIN initiative and a few others are pioneering this ever increasing need.

The multilingual tools we discussed in this paper have been tested on several languages and showed that when large and good quality language resources are available, rewarding results can be obtained with limited cross-linguistic expertise. Although our experiments considered only Indo-European languages, we are confident that even with more distant pairs of languages, provided the adequate resources are available, the alignment system and the linguistic knowledge induction applications should work reasonably well.

[14] http://cordis.europa.eu/esfri/roadmap.htm

References

1. Barbu-Mititelu, V., Ion, R.: Cross-language Transfer of Syntactic Relations Using Parallel Corpora. In: Proceedings of the Workshop on Cross-Language Knowledge Induction, EUROLAN’ 2005, Cluj-Napoca,Romania, pp. 46–51 (2005)
2. Brants, T.: TnT a statistical part-of-speech tagger. In: Proceedings of the 6th ANLP Conference, Seattle, WA, pp. 224–231 (2000)
3. Bertagna, F., Monachini, M., Soria, C., Calzolari, N., Huang, C.-R., Hsieh, S.-K., Marchetti, A., Tesconi, M.: Fostering Intercultural Collaboration: a Web Service Architecture for Cross-Fertilization of Distributed Wordnets. In [17] (2007)
4. Brown, P.F., Della Pietra, S.A., Della Pietra, V.J., Mercer, R.L.: The mathematics of statistical machine translation: Parameter estimation. Computational Linguistics 19(2), 263–311 (1993)
5. Ceauşu, A.: Maximum Entropy Tiered Tagging. In: Proceedings of the Eleventh ESSLLI Student Session, Malaga, Spain, pp. 173–179 (2006)
6. Ceauşu, A., Ştefănescu, D., Tufiş, D.: Acquis Communautaire sentence alignment using Support Vector Machines. In: Proceedings of the 5th LREC Conference, Genoa, Italy, pp. 2134–2137 (2006)
7. Erjavec, T.: MULTEXT-East Version 3: Multilingual Morphosyntactic Specifications, Lexicons and Corpora. In: Proceedings of the 4th LREC Conference, Lisbon, Portugal, pp. 1535–1538 (2004)
8. Fan, R.-E., Chen, P.-H., Lin, C.-J.: Working set selection using the second order information for training SVM. Technical report, Department of Computer Science, National Taiwan University (2005), www.csie.ntu.edu.tw/ cjlin/papers/quadworkset.pdf
9. Fellbaum, C., Vossen, P.: Connecting the Universal to the Specic: Towards the Global Grid. In [17] (2007)
10. Fellbaum, Ch.: WordNet: An Electronic Lexical Database. MIT Press, Cambridge (1998)
11. Gale, W.A., Church, K.W.: A Program for Aligning Sentences in Bilingual Corpora. In: Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics, Berkeley, California, USA, pp. 177–184 (1991)
12. Hashimoto, C., Bond, F., Flickinger, D.: The Lextype DB: A Web-based Framework for Supporting Collaborative Multilingual Grammar and Treebank Development. In [17] (2007)
13. Hayashi, Y.: Conceptual Framework of an Upper Ontology for Describing Linguistic Services. In [17] (2007)
14. Inaba, R., Murakami, Y., Nadamoto, A., Ishida, T.: Multilingual Communication Support Using the Language Grid. In [17] (2007)
15. Ion, R.: Metode de dezambiguizare semantică automată. Aplicații pentru limbile engleză şi română, PhD Thesis, Romanian Academy, Bucharest, Romania, pp. 145 (2006)
16. Ion, R., Tufiş, D.: Multilingual Word Sense Disambiguation Using Aligned Wordnets. Romanian Journal on Information Science and Technology, Tufiş D. (ed.) Special Issue on BalkaNet, Romanian Academy, 7(2-3), 198-214 (2004)
17. Ishida, T., Fussell, S.R., Vossen, P.T.J.M. (eds.): IWIC 2007. LNCS, vol. 4568. Springer, Heidelberg (2007)
18. Koda, T.: Cross-cultural Study of Avatars’ Facial Expressions and Design Considerations within Asian Countries. In [17] (2007)
19. Magnini, B., Cavaglià, G.: Integrating Subject Field Codes into WordNet. In: Proceedings of LREC2000, Athens, Greece, pp. 1413–1418 (2000)

20. Martin, J., Mihalcea, R., Pedersen, T.: Word Alignment for Languages with Scarce Resources. In: Proceeding of the ACL2005 Workshop on Building and Using Parallel Corpora: Data-driven Machine Translation and Beyond. Association for Computational Linguistics, Ann Arbor, Michigan, pp. 65–74 (2005)
21. Moore, R.C.: Fast and Accurate Sentence Alignment of Bilingual Corpora. In: Richardson, S.D. (ed.) AMTA 2002. LNCS (LNAI), vol. 2499, pp. 135–244. Springer, Heidelberg (2002)
22. Niles, I., Pease, A.: Towards a Standard Upper Ontology. In: Proceedings of the 2nd International Conference on Formal Ontology in Information Systems (FOIS-2001), Ogunquit, Maine, pp. 17–19 (2001)
23. Och, F., Ney, J.: Improved Statistical Alignment Models. In: Proceedings of ACL2000, Hong Kong, China, pp. 440–447 (2000)
24. Och, F.J., Ney, H.: A Systematic Comparison of Various Statistical Alignment Models. Computational Linguistics 29(1), 19–51 (2003)
25. Schneider, P., Mattenklott, A.: Emotion Eliciting Events in the Workplace: An Intercultural Comparison. In [17] (2007)
26. Smadja, F.A., McKeown, K.R.: Automatically extracting and representing collocations for language generation. In: Proceedings of the 28th annual meeting on Association for Computational Linguistics, Pittsburgh, Pennsylvania, pp. 252–259 (1990)
27. Sornlertlamvanich, V., Charoenporn, T., Robkop, K., Isahara, H.: Collaborative Platform for Multilingual Resource Development and Intercultural Communication. In [17] (2007)
28. Tapanainen, P., Järvinen, T.: A dependency parser for English. Technical Report no. TR-1, Department of General Linguistics, University of Helsinki, Finland (1997)
29. Todiraşcu, A.: Towards an automatic extraction of collocations; the case of the verb MAKE/DO (în Romanian). In: Proceedings of the National Workshop on Romanian Language Processing, Iaşi, pp. 95–101 (November 3-4, 2006)
30. Tufiş, D., Barbu, A., Ion, R.: Extracting Multilingual Lexicons from Parallel Corpora. Computers and the Humanities 38(2), 163–189 (2004)
31. Tufiş, D., Ion, R., Ceauşu, A., Ştefănescu, D.: Improved Lexical Alignment by Combining Multiple Reified Alignments. In: Proceedings of the 11th Conference of the European Association for Computational Linguistics (EACL), Trento, Italy, pp. 153–160 (2006)
32. Tufiş, D., Ion, R., Ceauşu, A., Stefănescu, D.: Combined Aligners. In: Proceeding of the ACL2005 Workshop on Building and Using Parallel Corpora: Data-driven Machine Translation and Beyond. Association for Computational Linguistics, Ann Arbor, Michigan, pp. 107–110 (2005)
33. Tufiş, D., Ion, R., Ide, N.: Fine-Grained Word Sense Disambiguation Based on Parallel Corpora, Word Alignment, Word Clustering and Aligned Wordnets. In: Proceedings of the 20th COLING Conference, Geneva, pp. 1312–1318 (2004)
34. Tufiş, D.: Tiered Tagging and Combined Classifiers. In: Matoušek, V., Mautner, P., Ocelíková, J., Sojka, P. (eds.) TSD 1999. LNCS (LNAI), vol. 1692, pp. 28–33. Springer, Heidelberg (1999)
35. Tufiş, D., Cristea, D., Stamou, S.: BalkaNet: Aims, Methods, Results and Perspectives. A General Overview. Romanian Journal on Information Science and Technology, Tufiş, D. (ed.) Special Issue on BalkaNet, Romanian Academy, 7(2-3), 9-34 (2004)
36. Tufiş, D., Barbu-Mititelu, V., Bozianu, L., Mihăilă, C.: Romanian WordNet: New Developments and Applications. In: Proceedings of the 3rd Conference of the Global WordNet Association, Jeju, Republic of Korea, pp. 337–344 (2006)
37. Vossen, P. (ed.): A Multilingual Database with Lexical Semantic Networks. Kluwer Academic Publishers, Dordrecht (1998)

Multilingual Communication Support Using the Language Grid

Rieko Inaba[1], Yohei Murakami[1], Akiyo Nadamoto[1], and Toru Ishida[1,2]

[1] Language Grid Project, National Institute of Information and Communications Technology, 3-5 Hikaridai, Seikacho, Soraku-gun, Kyoto, 619-0289, Japan
[2] Department of Social Informatics, Kyoto University Yoshida-Honmachi, Kyoto, 606-8501, Japan
{rieko.inaba,yohei,nadamoto}@nict.go.jp, ishida@i.kyoto-u.ac.jp

Abstract. Our proposed “Language Grid” infrastructure supports multilingual communication by combining in new way language resources, such as machine translators, morphological analyzers, and dictionaries specific to user communities. We developed the Language Grid as a language infrastructure on the Internet. The Language Grid enables user communities to combine two or more machine translators and their community dictionaries by workflows, and to easily create new multilingual services specific to the communities. Because the quality of language services is not often defined, however, we need to confirm that the created multilingual service is really useful. We need to extend the process of general usability testing to the multilingual environment. For example, cooperation between user communities and language grid providers can significantly improve the accuracy of machine translation: it turns out that machine translations can be useful for interactive communication in the field of inter-cultural collaboration.

Keywords: Multilingual Communication, Collaboration Tools, Language Grid, Usability.

1 Introduction

The language barrier is one of the biggest problems remaining on the Internet. The Internet population has become widely varied, and the percentage of people using the most common language, English, has fallen to about 35% [1]. Many people are actively working on inter-cultural collaboration, and the demand for machine translation on the Internet is growing. However, multilingual communication using machine translation is still quite problematic, particularly in regards to quality and usability. We have been working on both of these problems and have developed a language infrastructure on the Internet called the “Language Grid,” which can be customized for user communities.

Because the de facto language of multilingual activities is English, a burden is placed on non-English-speaking people- they cannot smoothly communicate with each other. Consider the case of a non-profit organization (NPO) encouraging

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 118–132, 2007.

children around the world to communicate through the Internet [2,3]. Not only the children but also the staff of the NPO would have trouble communicating smoothly with each other because of the language barriers. In theory, they could use machine translators; this would enable, for example, a Japanese staff member and an Austrian staff member to communicate using their native languages. Since developing machine translators is costly, it would be difficult to develop one for every possible language pair. A more practical approach would be to cascade two machine translators (Japanese->English and English->German). Incorporating community-oriented dictionaries would increase the quality of the translations, but this requires the participation of the relevant communities.

Fortunately, the Web 2.0 movement [4], which is supported by, among others, Wikipedia, Flicker, and Blog, is well underway. Moreover, the Internet is moving from the producer-consumer paradigm to the open collaboration paradigm. For example, Yakushite Net (http://www.yakushite.net) [5] increases the quality of machine translations by getting users to help develop community-oriented dictionaries. Users in a community can edit their dictionary and keep it up-to-date with the terms specific to their community. As a result, Yakushite Net can produce high-quality translations.

Our Language Grid infrastructure is based on Web 2.0 [6]. It combines multiple language services, including machine translators on the Internet, with community-oriented dictionaries. Our multilingual communication environment provides intercultural collaboration tools for supporting community-oriented language services. These services include language resources, such as dictionaries, thesauri, and corpuses, and language processing functions, such as morphological analysis, translation, and paraphrasing. In this paper, we focus on two research issues.

Customizing multilingual translation using Web-service and workflow technologies
Users can customize multilingual communication through Web-service and workflow technologies; the Language Grid combines multiple language services and community-oriented dictionaries. We will describe our approach to an effective multilingual communication environment based on workflow-based Web-service composition

Multilingual communication tools with usability testing
When users work with information and communication technologies, usability testing is quite important, and much research has already been done on usability testing. Multilingual communication tools, however, require new research on the usability of composite language services including machine translators. To identify the usability issues related to machine translation, we observed various instances of multilingual collaboration, particularly the use of multilingual communication tools.

The remainder of this paper is organized as follows. Section 2 discusses issues related to multilingual communication, Section 3 explains ongoing research and development of the Language Grid, and Section 4 describes the results gleaned from several case studies. We conclude with a short summary in Section 5.

2 Issues on Multilingual Communication

2.1 Quality of Machine Translation

The quality of machine translation is an important factor in ensuring the mutual understanding of people communicating using machine translators. People communicating multilingually can generally catch the basic ideas from the background and context. However, when machine translators are used, this may not be possible because machine translators often output incorrect translations. This is because most machine translators were developed to handle declarative sentences, like those in news articles, which makes it difficult for them to translate dialog sentences.

Nomura et al. [7] tested communication among five languages on the Internet in a study entitled "Intercultural Collaboration Experiments (ICE2002)." The target languages were Chinese, Japanese, Korean, Malay, and English. They developed open software on a multilingual bulletin board system (BBS). A participant input dialog sentences in his/her mother language and received them back translated into the five languages. However, the machine translations were problematic, which impaired communication. They thus implemented a "self-initiated repair" feature. The user first inputs sentences in his/her mother language and then reviews the sentences translated into one of the output languages. The user can then modify, as needed, the original input sentences in their mother language to yield correct sentences in the output language. The accuracy of the machine translation program is thereby improved. In this experiment, the participants used English as the output language for self-initiated repair. The average number of repair iterations was seven or eight, and the accuracy of machine translation was improved by about 60%. They found that self-initiated repair worked well for bilingual users, but was ineffective for non-bilingual users.

We addressed the problem of improving the accuracy of machine translation for non-bilingual users. The approach we took is called "back translation." The input sentences are first translated into the target language (outward) and then automatically translated back into the input language (homeward). The user can then assess the accuracy of the machine translation in his/her mother language. We have developed a multilingual communication tool that supports back translation.

2.2 The Asymmetry and Heterogeneity of Translation

The two key problems with back translation are asymmetry and heterogeneity [8, 9].

First, languages are not completely symmetrical. For example, "banque" in French would normally be translated into "bank" in English, but "bank" might be translated into rive ("river bank") in French. Even if we use the same translation engine for the initial and back translation (Japanese->English and English->Japanese), there may be asymmetry between the original input and the back translation. This means that the more something is translated, the greater the number of translation errors. Therefore, if back translation is used for the self-repair, the cycle of translation->back translation->repair->translation can lead to increasingly incorrect translations. Use of back translation in a multilingual communication tool thus requires consideration of translation asymmetry.

Second, there are differences in the outputs of different machine translators, i.e., there is heterogeneity. This is even more of a problem when the translations are between different pairs of languages. For example, Japanese->English machine translation will generally be more accurate than Japanese->Chinese translation of the same text. While the Japanese, English, and Chinese communicators in this example would be able to communicate, the quality of the translations they receive would be uneven. We thus need to consider how to create a multilingual communication environment in which the participants can understand the "conversation" equally well.

Moreover, to enable the easy use multiple of languages in a "conversation," we need an open-software-based language service that connects multiple language resources on the Internet.

To meet these requirements, we combine community-oriented language services with machine translators to improve translation accuracy. This requires customizing the machine translation engines to enable integration of community-oriented dictionaries. The services should run on community servers because a community can have various types of multilingual collaboration. This is the concept of our Language Grid infrastructure.

2.3 Usability

The opportunities for multilingual collaboration are increasing with the spread of the Internet. However, participants do not necessarily have a common language. Therefore, the demand for multi-language collaboration tools is growing. However, most such tools are designed and evaluated without undergoing intercultural usability testing due to aggressive development schedules, limited budgets, and/or lack of knowledge on how to conduct such research.

Usability testing related to multilingual communication typically involves identifying the target languages, user locations, and types of communication to be evaluated. Most such evaluations have focused on the usability of multilingual operating systems, the addition of multilingual user interfaces, etc. There have been few evaluations of the functions needed for communication involving many languages. We focused our usability evaluation on the tools required to support multilingual communications, aiming to clarify the optimal evaluation technique.

The elements of a multilingual communication tool are a machine translation system that enables mutual understanding and groupware that expands the range of communications. Multilingual communication tasks involve language translation processes.

3 Language Grid Research and Development

3.1 Language Grid

We have proposed the Language Grid. The Language Grid treats existing language services as atomic components and enables users to create new language services by combining the appropriate components. Generally, the word *grid* is defined as "a system coordinating distributed resources with each other; it uses an open standard protocol for creating high quality services". The language grid has two main functions: the *horizontal language grid* and the *vertical language grid*. The *horizontal language grid* combines existing language services using Web service and workflow technologies. The horizontal language grid benefits a wide range of users by providing standard language services for about 10 Asian languages and 20 other languages from around the world. The vertical language grid, on the other hand, layers community language services, such as medical parallel texts and pictogram dictionaries, to support intercultural activities.

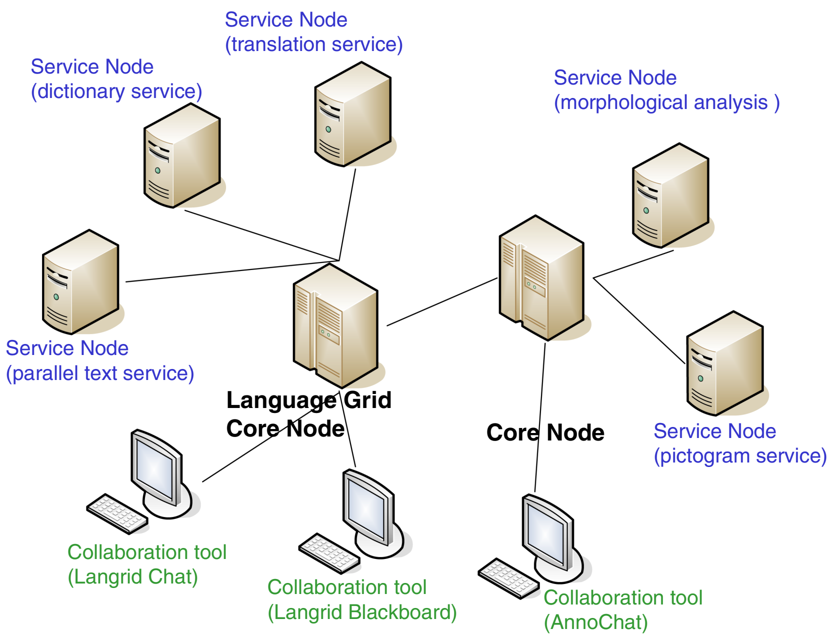


Fig. 1. Concept of the Language Grid

We have two goals for the Language Grid: one is “Everyone can create community oriented services and register them with the Language Grid.”, and the other is “Everyone can use language services on the Language Grid easily”. These enable users to solve the problems given in Section 2 and to create high quality translation services.

When we realize the Language Grid, we have to standardize interfaces of the language services and to increase the accessibility and usability of the language services. In order to promote such standardization, language service providers implement their own language service as a Web service with a standard interface, deploy them on the Language Grid, and register its WSDL description and profile on the Language Grid. Meanwhile, language service users locate the services they need on the Language Grid and then invoke them. At the same time, they can construct a new composite language service by applying the workflow, and then deploy the services. Also, even when a language service is implemented without the standard interface, it can be integrated into the Language Grid by a third party (the wrapper implementer), who deploys a wrapper to adjust its interface to the standard interface. In this case, the service entity of the language service is located outside the Language Grid and is invoked by the wrapper.

3.2 Multilingual Collaboration Tools [6]

The Language Grid project is developing several tools for multilingual collaboration support using the language service registered on the Language Grid. The users input texts in native languages to these tools and the outputs are displayed in other languages as desired. Multilingual translation is realized by workflows coordinating several language services. In multilingual translation, users employ back translation in order to check whether the translation of the input text is correct or not. They also provide a tool to create community oriented dictionaries. Users can add unknown words to the dictionary in order to improve the translation accuracy.

Multilingual chat tool - Langrid-Chat [10]
Langrid Chat is a tool that it uses multilingual translation developed on the Language Grid. When a user inputs a sentence in his/her native language, the back translation results, via the target languages set by the user, are returned in real time. The user can then confirm whether the translation is reasonable.

Multilingual sharing blackboard –Langrid Blackboard
Langrid Blackboard is a multilingual blackboard tool that enables users to share various types of objects, such as labels, arrows, and pictures. It provides two windows to the user: a sharing window and an input window. In the sharing window, the user can use comment labels, image labels, and grouping labels. The user can also create labels to freely input his or her ideas in the input window. Such labels are shared among the participants and are displayed at the same position on each participant's display. This tool enables users to converse in their native languages.

Repair support for multilingual back translation
The chat and blackboard tools provide a multilingual back translation function that can be used to check translation results in multiple languages. They also enable users to fix each translation result individually. For example, the Korean back translation might be good while the English or Chinese one is not. It would be surprising if all the initial back translations from one input sentence were good given the problem of homogeneity described above. As the number of target languages continues to increase, this will become even more of a problem.

3.3 Multilingual Translation Workflow

There are several problems in supporting multilingual communication using a multilingual communication tool based on machine translation: translation path, translation quality, and morphological between machine translation. We avoid these problems by having the users construct and maintain community-oriented translation services. Such services combine multiple machine translators, community-oriented dictionaries, and various kinds of language resources.

An example language service that supports communities is "Yakushite Net" [11], a Web-based collaborative translation environment. It enables people with deep knowledge of a particular domain to collaborate in enhancing specialized dictionaries for online machine translation and thus enable more accurate translations specific to a community. However, current mechanisms require customizing the machine translation engine in order to add dictionary entries created by a community. We need a method that can customize not only the machine translation engine itself but also the workflows used to coordinate the language services, including machine translation. Such a method would enable us to replace one language service with another that has the same interface, thereby improving translation scalability.

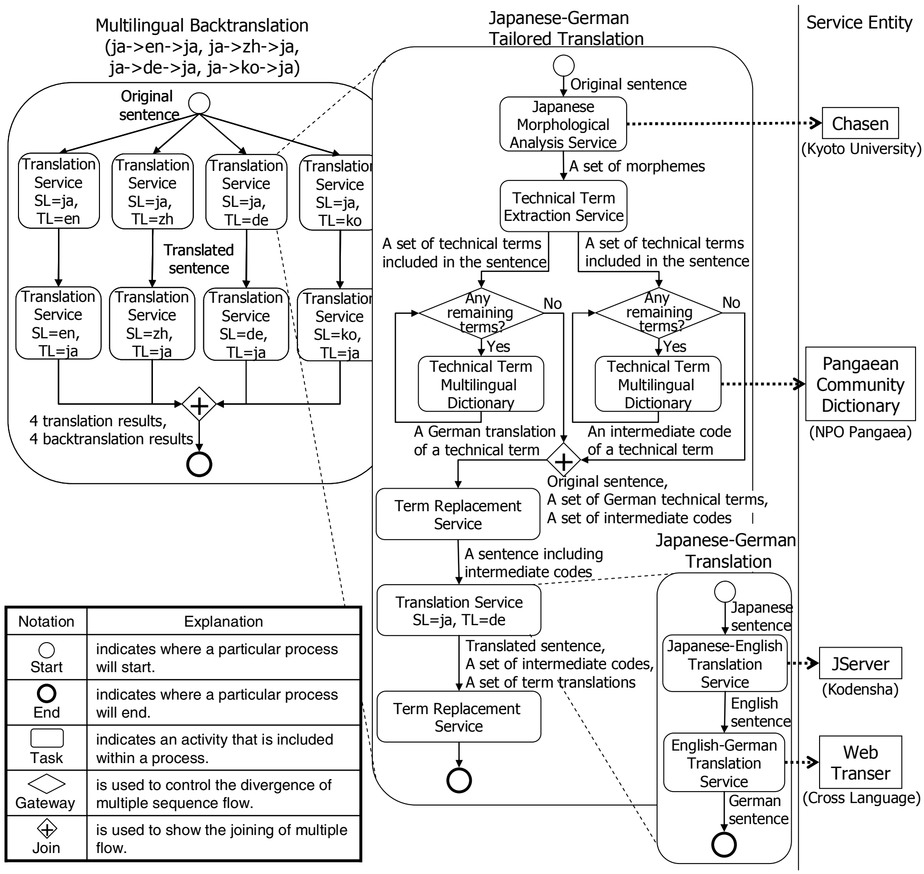


Fig. 2. Workflow for multilingual communication support

An example workflow is illustrated in Figure 2; it is the translation service for Pangaea, an NGO based in Japan. It combines a morphological analysis service, a bilingual dictionary of technical terms, and a machine translation service, which translates Japanese into English, Korean, and/or German and provides the corresponding back translations. Simultaneous multiple back translations (Japanese->English->Japanese, Japanese->Korean->Japanese, and Japanese->German->Japanese) are initial performed. The workflow for each back translation combines the community dictionary and multiple community-oriented machine translations. The composite service divides each input sentence into a set of morphemes using the morphological analysis service. Next, it extracts technical terms from the set of morphemes and creates a set of the technical terms included in the input sentence. Next, it obtains accurate translations of the technical terms from the bilingual dictionary. In the same way, it also obtains the intermediate codes of the technical terms, which are strings that have little effect on machine translation. It replaces the technical terms in the input sentence with the intermediate codes and translates the sentence, including the intermediate codes, using the machine translation service. Finally, it replaces the translated intermediate codes in the translated sentence with the translations of the technical terms taken from the bilingual dictionary. In the current version of this system, if there is no machine translation service for the target language, the system performs multi-hop translation using multiple machine translators.

To deploy a community-oriented multilingual back translation service on the Language Grid, we have to wrap the existing morphological analysis and machine translation services. On the other hand, we can simply implement the technical term multilingual dictionaries on the standard interface before deploying them on Language Grid. In this way, each language service is available on the Language Grid, and they are coordinated by the workflow. In the example shown in Figure 2, the workflow invokes Chasen [12] for Japanese morphological analysis, the Pangaea community dictionary, and JServer and Web Transer for Japanese->English translation and English->Japanese translation, respectively.

3.4 Evaluation of Usability for Multilingual Environment

A system should be constructed with not only the developer but also the users in mind. We built an evaluation process of model usability for multilingual communication tools. The elements of a multilingual communication tool are (1) collaboration tools, and (2) a machine translation system. Therefore, we conducted two usability tests. Our preliminary study focuses on the evaluation of the user interface and functions on collaboration tools. The second study focuses on translation quality using community words, that is, domain specific words. Multilingual communication tasks involve language translation processes. Both general usability [12] and subject usability by means of a language translation process are described here. We describe a usability evaluation process for multilingual communication tools, which are shown in Figure 1. The left side of the figure shows the evaluation process of the general usability. After the development team completed the prototype, we evaluated its usability with two approaches. One is heuristics testing, which is an analytic technique, and the other is usability testing, which involves experiments. The results provide feedback on the development process.

The goal is multilingual communication. Therefore, in addition to an evaluation of the general usability, a multilingual evaluation is also required. The evaluation items specific to multilingual communication are shown on the right side of Figure 3. This process has three issues regarding usability evaluation.

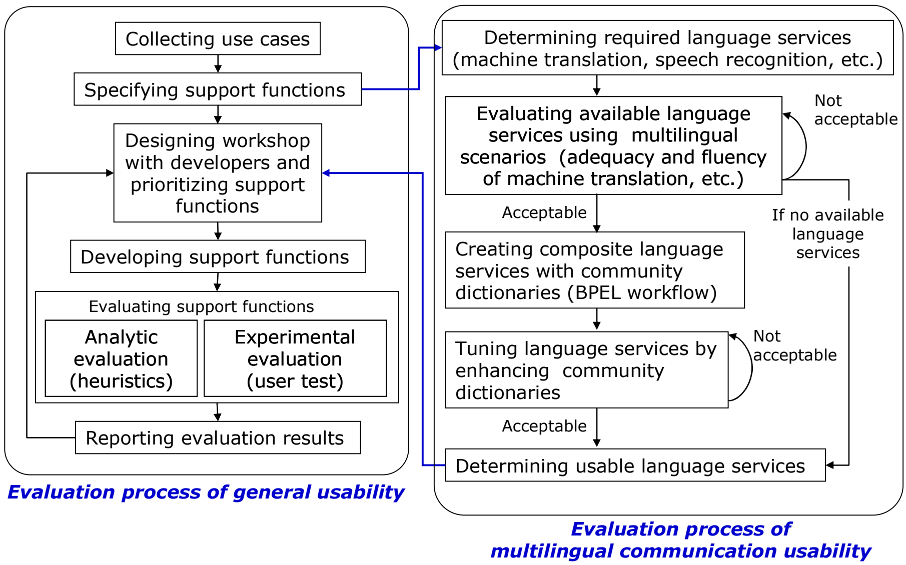


Fig. 3. Usability evaluation process

1) Evaluating available language services using multilingual scenarios

People use the language services on intercultural collaboration. We need multilingual test set to evaluate the accuracy of multilingual translation. Test sets of a parallel corpus based on newspaper articles or similar material are currently available. Moreover, collaboration tools mainly translate in a colloquial style. Furthermore, when we check cooperative machine translation systems with a community dictionary, we need a test set that includes domain words.

A subjective evaluation was conducted using the National Institute of Standards and Technology (NIST) protocol[1]. Judges decide whether each translation of a sentence is adequate. Adequacy refers to the degree to which information present in the original is also communicated in the translation. Thus, for adequacy judgments, the reference translation serves as a proxy for the original source language text. Adequacy is judged on the following five-point scale:

Table 1. A standard of evaluation

	How much of the meaning expressed in the gold-standard translation is also expressed in the target translation?
5	All of it
4	Most of it
3	Much of it
2	Little of it
1	None of it

[1] http://www.ldc.upenn.edu/Projects/TIDES/Translation/TransAssess02.pdf

2) Creating composite language services with community dictionaries
A community dictionary is very important for intercultural collaboration. Users in the community have and use the domain specific words. Therefore, we combine multiple language services and community dictionaries by using workflow technologies shown in the section 3.3.

3) Tuning language services by improving community dictionaries
Language services are being improved because of constant evaluations of their adequacy. A language grid replaces the technical terms an input sentence with the translations of technical terms taken from a bilingual dictionary or a community dictionary. Thus optimal tuning language services are very important to ensure accuracy. An example of these services is shown in Table2. A sentence inputted in Japanese and the results of an initial translation into English and a subsequent translation back into Japanese are shown. The word of "I love map" is domain-specific word in the NPO Pangaea.

Table 2. Example of tuning language services with community dictionary

MT without community dictionary	
Input sentence in Japanese	Please gather children's "I love map".
Translation from Japanese to English	Children's, fond, please collect maps.
Translation from English back to Japanese	Children's, kindly, please collect the map.
MT with community dictionary	
Input sentence in Japanese	Please gather children's "I love map".
Translation from Japanese to English	Please collect children's I love map.
Translation from English back to Japanese	Please collect children's I love map.

We evaluate the tuning language services using 74 domain-specific scenarios. Four people did the evaluations in each language pair. Table3 shows the evaluation results.

Table 3. Evaluation results: Adequacy of tuning language services

1) Japanese- English- Japanese

Evaluator	A	B	C	D
MT without community dictionary	2.83	2.93	3.23	3.03
MT with Community dictionary	4.08	3.51	4.22	3.51

2) Japanese- Korean- Japanese

Evaluator	E	F	G	H
MT without community dictionary	3.79	3.06	3.26	3.46
MT with Community dictionary	4.59	4.18	4.54	4.61

These results show that we should improve language services by using community dictionaries.

4 Experiments

4.1 Multilingual Communication Project

Ways to support the need for multilingual communication among the members of two NPOs, Pangaea [3] and JEARN (Japan Education and Resource Network) [13], were investigated.

Pangaea develops and implements virtual playgrounds where children around the world can meet, communicate, and connect. The staffs in the Pangaea offices in Japan, South Korea, Kenya, and Austria need to communicate using their native languages. To meet this need, we provided Annochat, a chat tool that supports graphic annotation and cooperates with Language Grid. Users can talk by multilingual chatting in real time.

JEARN is the Japanese center of the iEARN (International Education and Resource Network) educational network. It promotes international collaboration projects. For example, at its Natural Disaster Youth Summit, students created disaster safety maps of the area surrounding their schools. These maps will form a "Global Disaster Safety map." The participating countries include Japan, Russia, Taiwan, Armenia, Iran, Trinidad and Tobago, and Senegal. The students are exchanging information and discussing matters related to map creation using a BBS, or "discussion board." However, most participants have poor English skills and thus hesitate to speak up. We proposed enabling them to input to the BBS in their native language using Language Grid.

Pangaea needed a collaboration tool on Language Grid while JEARN wanted to use their original BBS but extended to provide a multilingual input tool. These cases demonstrated that the support tool must be customized to match the user's requirements. They also helped clarify the issues related to workflow and usability.

Management and use of translation tuning knowledge
The community-oriented dictionaries should be incorporated into the machine translation services by using the workflow in order to better support specialized communities and to increase the quality of translation. However, a workflow must be created for each community. Moreover, the structure of the workflow depends on the language resources because the know-how developed in one community cannot necessarily be used in another one. We thus need to be able to manage knowledge about the cooperation workflow.

Intellectual property rights of language resources
Each language service has a different intellectual property policy. For example, a morphological-analysis tool developed by an academic organization is generally provided without charge to NPO users and researchers. In contrast, a machine translation engine developed by a company is generally provided for a fee, even to NPO users. These various types of intellectual property policies must be considered when incorporating language resources into a workflow.

Our investigation also revealed two problems related to usability.

Usability evaluator

To evaluate the possibility of a user's language, we must create test data including community-oriented words and evaluate the accuracy of the translation. To judge whether a language pair, for example, Japanese and Korean, is supported, we need multilingual Japanese and Koreans. The evaluation has two steps: accuracy and specialized community. In the first step, the evaluator does not need to be a community expert who can understand the community words. In the second step, it is necessary. We thus need a community expert for at least part of the evaluation. We would thus like to focus on the differences in the evaluation of the community-oriented words instead of the general words whether the evaluator is a community expert or not. The problem is that dependence on community style also has a setup of the threshold value of the valuation basis of the first step and the second step.

Selection of dictionary

Community words often have two meanings: the community-oriented one and the general one. For example, in the case of Pangaea, “activity” has two possible meanings.

General translation: "activity" -> katsudou (standard Japanese for “activity”)

Community translation: "activity" -> acutyibityi (“activity” rendered into Japanese phonology.)

Since priority is given to the community-oriented dictionary, "activity" is always translated into "acutyibityi." However, there may be cases in which the user actually wants the general meaning. A function is thus needed for determining which is the desired meaning.

4.2 All-for-One Project

At international conferences, the presentations and discussions are usually in English. This is because it is very difficult for people whose native language is not English to think in English. Those who cannot understand the language used in the discussion tend to feel left out since they cannot understand what is being discussed.

The goal of all-for-one collaboration is getting participants who can understand the language to summarize the contents of the discussion for those who cannot understand the language. The participants who cannot understand the language can then ask the input operators of erroneous translation of the summary. Finally, the contents of the discussion are shared among the participants through the Langrid Blackboard collaboration tool. They can share texts on labels and stick the labels on the board as they like. Moreover, they can summarize the discussion contents by classifying the labels by topic. This tool also has a community-oriented dictionary creation function, and words can be added to the dictionary at any time. As a result, collaboration is more efficient.

Figure 4 illustrates the differences between multilingual communication and all-for-one collaboration. The former involves the application of a collaboration tool, and the users are active communicators. In the latter, the users are passive recipients of information. The following issues related to the latter were revealed experimentally.

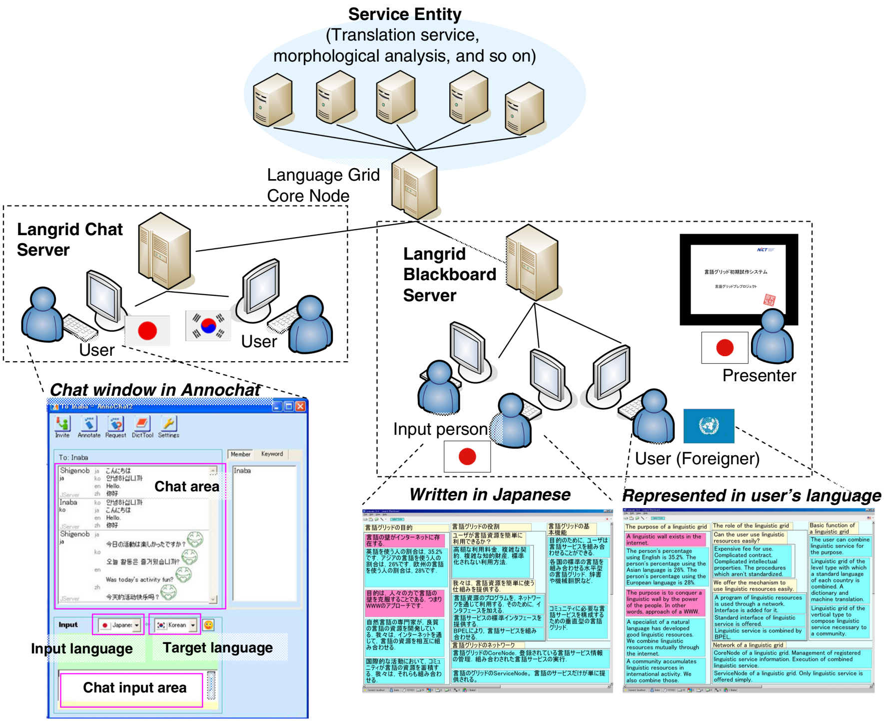


Fig. 4. Multilingual communication project and All-for-One project

Collaboration style

The participants who can understand the language being used are asked to record what they catch in real time. However, without preparation beforehand, they must catch what was said, determine the meaning, and write it on the Blackboard. Tthe accuracy of participants who cannot understand the language used is low. The communication style between the participant and the presenter is thus important. It is also important for the person in charge to get detailed information about what will be said and pass it on to the recorders ahead of time. This means that it is necessary to construct an overall support organization and to implement an effective collaboration style.

Community words construction

The users of the collaboration tool are foreigner audiences and writers. However, the presenter uses community words. Therefore, when the recorder does not have knowledge of the community, the handling of the community-oriented dictionary is problematic.

Translation accuracy vs. speed

In all-for-one collaboration, the recorders attempt to input text at the same speed at which the presenter is talking. Therefore, the translation service used must offer is asked for an early translation speed. However, if we use self-repair to improve translation accuracy, the input speed is reduced. There is thus a trade-off between translation speed and translation accuracy.

Furthermore, 3-hop translation (Japanese->English->French) is needed for Japan->France translation. As one French researcher working in Japan said, "Although there were many mistakes in the grammar of the translated French, it was still possible to catch the meaning from the context of what was being said."

5 Conclusion

To support multilingual collaboration, we need to coordinate machine translators and community dictionaries specific to the collaboration tasks. To improve the accessibility and usability of language services such as machine translations and dictionaries, we applied the Language Grid infrastructure to the following research issues.

Customizing multilingual translation using Web-service and workflow technologies
By creating a workflow to combine language services, such as machine translators, community dictionaries, and morphological analysis, we successfully improved the translation accuracy of texts including specific terms often used in a particular community. Because users are often not allowed to incorporate their community-specific dictionaries into machine translators, we use workflows to realize community-tailored translation. We increase the quality of machine translation output by modifying the translation results through the use of community-oriented dictionaries.

Multilingual communication tools with usability testing
By adding the evaluation of translation accuracy to general usability testing, we successfully identified a translation path that causes translation errors. In the usability testing, we also evaluated the use of community-oriented dictionaries to improve translation results. We found that tuning these dictionaries is the key to community-tailored translation.

We applied our proposed infrastructure to projects involving multilingual collaboration and identified several research issues: how to accumulate know-how for describing workflows, how to satisfy constraints among the intellectual property policies of language services, how to collect subjects for usability testing, and how to detect community-oriented vocabulary during communication.

Acknowledgments. We are grateful for the assistance of Takashi Yoshino, Tomohiro Shigenobu, and other members of the Language Grid Association. Translation services for this research were provided by Kodensha Co., Ltd. and Cross Language Inc.; morphological analysis services were provided by Kyoto University, Kookmin University, and Stuttgart University.

References

1. Andrews, T., Curbera, F., Dolakia, H., Goland, J., Klein, J., Leymann, F., Liu, K., Roller, D., Smith, D., Thatte, S., Trickovic, I., Weeravarana, S.: Business Process Execution Language for Web Services (2003)
2. Takasaki, T.: PictNet: Semantic Infrastructure for Pictogram Communication. In: Proceedings of the Third International WordNet Conference (GWC-06), pp. 279–284 (2006)

3. Mori, Y.: Atoms of Bonding: Communication Components Bridging Children Worldwide. In: Ishida, T., Fussell, S.R., Vossen, P.T.J.M. (eds.) IWIC 2007. LNCS, vol. 4568, Springer, Heidelberg (2007)
4. O'Reilly, T.: What Is Web2.0, http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html
5. Murata, T., Kitamura, M., Fukui, T., Sukehiro, T.: Implementation of collaborative translation environment: Yakushiutenet. In: Proceedings of the Ninth Machine Translation Summit System Presentation, pp. 479–482 (2003)
6. Ishida, T.: Language Grid: An Infrastructure for Intercultural Collaboration. In: Proceedings of the IEEE/IPSJ Symposium on Applications and the Internet (SAINT-06), pp. 96–100, 2006 (keynote address)
7. Nomura, S., Ishida, T., Yasuoka, M., Yamashita, N., Funakosh, K.: Open Source Software Development with Your Mother Language: Intercultural Collaboration Experiment 2002. In: Proceedings of the International Conference on Human-Computer Interaction (HCI2003), pp. 1163–1167 (2003)
8. Yamashita, N., Ishida, T.: Effects of Machine Translation on Collaborative Work. In: Proceedings of the International Conference on Computer Supported Cooperative Work (CSCW-06), pp. 515–523 (2006)
9. Yamashita, N., Ishida, T.: Automatic Prediction of Misconceptions in Multilingual Computer-Mediated Communication. In: Proceedings of the International Conference on Intelligent User Interfaces (IUI-06), pp. 62–69 (2006)
10. Fujii, K., Shigenobu, T., Yoshino, T.: Application and Evaluation of Intercultural Communication Tool AnnoChat using Machine Translation. In: Proceedings of the Information Processing Society of Japan, pp. 67–72, 2005 (in Japanese)
11. Shimohata, S., Kitamura, M., Sukehiro, T., Murata, T.: Collaborative Translation Environment on the Web. In: Proceedings of the MT Ssummit VIII (2001)
12. Asahara, M., Matsumoto, Y.: Extended models and tools for high-performance part-of-speech tagger. In: Proceedings of the18 the Conference on Computational Linguistics, pp. 21–27 (2000)
13. Naya, Y.: How Intercultural Disaster Reduction Education Change Students: A Case Study of an Evening Course Senior High School in Hyogo, Japan. In: Ishida, T., Fussell, S.R., Vossen, P.T.J.M. (eds.) IWIC 2007. LNCS, Springer, Heidelberg (2007)
14. Papineni, K., Roukos, S., Ward, T., Zhu, W.: BLEU: a Method for AutomaticEvaluation of Machine Translation. In: Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL), pp. 311–318 (2002)

Hanzi Grid

Toward a Knowledge Infrastructure for Chinese Character-Based Cultures

Ya-Min Chou[1], Shu-Kai Hsieh[2], and Chu-Ren Huang[3]

[1] Jin-Wen Institute of Technology, Taiwan
milesymchou@yahoo.com.tw
[2] National I-Lan University, Taiwan
shukai@gmail.com
[3] Institute of Linguistics, Academia Sinica, Taiwan
churenhuang@gmail.com

Abstract. The long-term historical development and broad geographical variation of Chinese character (Hanzi/Kanji) has made it a cross-cultural information sharing platform in East Asia. In this paper, we propose a theoretical framework for the knowledge representation of Hanzi in the cross-cultural context. Our proposal is mainly based on two resources: Hantology and Generative Lexicon Theory. Hantology is a comprehensive Chinese character-based knowledge resource created to provide a solid foundation both for philological surveys and language processing tasks, while Generative lexicon theory is extended to catch the abundant knowledge information of Chinese characters within its proposed qualia structure. We believe that the proposed theoretical framework will have great influence on the current research paradigm of Hanzi studies, and help to shape an emergent model of intercultural collaboration.

Keywords: Chinese characters, ontological knowledge resource.

1 Introduction

1.1 Motivation

Chinese character (Hanzi) has existed for over thousands of years. The long-term historical development and broad geographical variation of Hanzi has made it a valuable resource for multi-linguistic and cross-cultural mediation in Asia, especially among the *Sinosphere* (a new-coined term also known as 漢字文化圈 Chinese character cultural sphere), which denotes a grouping of regions and countries where Chinese characters were adopted and integrated to their languages, or historically under Chinese cultural influence ([1]).

However, due to the lack of proper research framework in information science, the integration of heterogeneous knowledge grounded in Hanzi and its variants has been a thorny problem. To achieve that ultimate goal, the first attempt is to provide a linguistic proper and computational interoperable framework, which can facilitate collecting and integrating the usage and knowledge of Chinese characters and their variants, in different spatial and temporal dimensions. In this paper, we propose a Knowledge-driven Hanzi-centered framework for diachronic

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 133–145, 2007.

and cross-cultural knowledge representation. We hope that the construction of this proposed framework will facilitate the intercultural collaborative research on Chinese characters.

1.2 Interfacing Ideographic Script and Conceptual Knowledge

In the history of western linguistics, writing has long been viewed as a surrogate or substitute for speech, the latter being the primary vehicle for human communication. Such "surrogational model" [6] which neglects the systematicity of writing in its own right has also occupied the predominant views in current computational linguistic studies. This paper is set to provide a quite different perspective along with the Eastern philological tradition of the study of scripts, especially the ideographic one i.e., Chinese characters (Hanzi).

For phonological writing systems, a character as a writing unit usually represents a phoneme or a syllable. Since the number of phonemes is finite, only a small set of phonetic symbols is needed to represent sounds of words for such languages. A character in the Chinese writing system, however, is a writing unit that represents a compact package of concept and pronunciation. Through over 3000 years of use, the complete Chinese writing system consists of at least 40,000 characters, over 100,000 variants are counted. In theory, it is still possible to invent and add new characters to the inventory today, although in practice, this is a very rare event. Each Chinese character represents one or more different concepts. Like alphabetic or syllabic characters, the Chinese characters serve as a basis of lexical classification. Unlike phonological writing systems, however, the classification is largely conceptual or semantic. In other words, the linguistic ontology of Chinese characters is explicitly marked with logographic features.

From the view of writing system and cognition, human conceptual information has been regarded as being *wired* in ideographic scripts. We believe that the conceptual knowledge information which has been *grounded* on Chinese characters could be used as a cognitively sound and computationally effective ontological lexical resource in performing some NLP tasks, and it will have contribution to the cross-cultural collaboration in the *Sinosphere* within the context of *Semantic Web* as well.

1.3 Hanzi and Conventionalized Conceptualization

Since Chinese characters act as information bearers for over thousands of years, some researchers proposed that the whole set of Chinese characters can be viewed as an *encyclopedia* in essence. In terms of knowledge representation, we prefer to refer to it as a kind of ontological knowledge. But, can an ontology be psychologically real and be evidenced by shared human experience? This is one of the critical issues that linguistic ontologies, such as WordNet [5], tries to answer. The successful applications of WordNet seem to give a positive reply to this question. However, all the conceptual relations (or lexical semantic relations) of WordNet are annotated by experts, not conventionalized. Hence there is no direct evidence of the psychological reality.

Based on our observation, we find that Chinese writing system can be treated as a linguistic ontology since it represents and classifies lexical units according to semantic classes. Having been used continuously for over 3000 years, it has conventionalized a system of semantic classification. The system is richly structured and robust, and adopted by other languages belonging to different language families. For example, Chinese characters have been incorporated into the writing systems of Japanese (called *Kanji*), Korean (called *Hanja*), and Vietnamese (called *Chunom*), which belong to Japonic, Altaic (debated) and Austro-Asiatic language family, respectively. [1]

2 Hanzi as a Multi-levels Knowledge Resource: Review of Current Works

In general, a Chinese character is an *ideogram* composed of mostly straight lines or "poly-line" strokes. A number of characters contain relatively independent substructures, called components (or glyphs), and some common components (traditionally called radicals) are shared by different characters. Thus, the structure of Chinese characters can be seen to consist of a 3-layer affiliation network: *character, component (glyph)* and *stroke*.

Linguistically, a Hanzi is regarded as an ideographic symbol representing *syllable and meaning* of a "morpheme" in spoken Chinese, or, in the case of polysyllabic word, one syllable of its sound. Namely, character, morpheme and syllable are *co-extensive*.

The fact that characters can be investigated from different angles, resulting in the various approaches to the knowledge mining from them. For example, there are several studies on the creation of Chinese characters database. One important study is Chinese glyph expression database which consists of 59000 glyph structures [11]. The glyphs of Chinese characters are decomposed into 4766 basic components. Each Chinese character can be expressed by the basic components. Chinese glyphs database also contains oracle bone, bronze, greater seal and lesser seal scripts. The largest Chinese characters database is Mojikyo font database which contains more than 110000 characters [10]. Both Chinese glyph expression database and Mojikyo font database contain only glyph knowledge. Yung created an ancient pronunciations database for Chinese characters. Hsieh and Huang [9] proposed an ontological lexical resource based on Chinese characters called HanziNet, in which Chinese characters are located within the context of upper level ontology.

These previous studies unveil many single dimensions of Chinese characters. However, each Chinese character consists of glyphs, scripts, pronunciations, senses, and variants dimensions. To meet the need of computer applications as well as the Chinese philological studies, Chou and Huang [2] propose a language

[1] Besides Japanese, Korean, and Vietnamese, a number of Asian languages have historically been written with Chinese characters, or with characters modified from Han characters. They include: Khitan language, Miao language, Nakhi (Naxi) language (Geba script), Tangut language, Zhuang language and so on.

resource called Hantology (Hanzi Ontology). The construction of Hantology focuses on the comprehensive and robust description of linguistic and conceptual knowledge encoded in Chinese writing system through the decomposition and composition of Chinese characters.

The linguistic knowledge described in Hantology includes various information concerning glyph, script, pronunciation, sense, and variants of Chinese characters. In addition, Hantology has been mapped to SUMO (Suggested Upper Merged Ontology) [12], and it is fully encoded in OWL for shareability and for future Semantic Web applications as well.

3 Radical-Centered General Theoretical Framework

In this and following sections, we will propose a radicals-centered general framework for the knowledge representation of Hanzi.

Generally speaking, each Chinese character is composed of two parts: a radical representing semantic classification, and a phonetic indicating phonological association. This generalization applies to the majority of Chinese characters, though not all. A minority estimated at less than 20% of all Chinese characters show other forms of composition. However, it is still true that these characters contain at least one semantically significant component. A small set of examples based on the radical 馬 (*ma3*, 'horse') are given below to show the range of assigned meanings. In these examples, 馬 is both a character and a radical denoting 'horse':

驩 : a kind of horse
驫 : many horses
騎 : to ride a horse
驍 : a good horse
驚 : to be scared (referring to a horse)

These Chinese characters shown above suggest that radicals are indeed concept-based. However, it also been shown the *conceptual clustering is more complex than a simple taxonomy*. In this paper, we focus mainly on the knowledge structure of radicals of Chinese characters, whose significance in representing semantic taxonomy was first observed by Shen Xu [17]. The following will elaborate on these observations.

3.1 Bootstrapping Conceptual Representation with Chinese Radicals

Any formal account of a conceptual system faces the dilemma of choosing a representational framework. Since a representational framework is itself build upon certain conceptualization, any choice is potentially an *a priori* distortion of the account. A possible solution to this dilemma is a shared upper ontology that is conceptually complete and yet general and robust enough to cover different conceptual systems under consideration. Take the Hantology for example, the Suggested Upper Merged Ontology [12] was adopted in this resource. All concepts expressed in Chinese characters are mapped to SUMO representation in the hope that the mapping can be transformed to a specialized ontology later.

One of the first implications of adopting SUMO representation is the fact that we are now able to formally represent knowledge inference based on the linguistic knowledge provided by radicals. Recall that the radical part of a Chinese character encodes semantic classification. For instance, all characters containing the 魚(‘fish’) radical can be assigned to this SUMO’s concept with the same knowledge. Applying the linking between SUMO and WordNet the following inference is possible whenever an English word is classified as a hyponym of fish in WordNet. However, no labor intensive manually classification is needed for Chinese. The same inference can be achieved automatically with logographic information, by assigning the default inference rule to all characters with the 魚 (‘fish’) radical.

```
(subclass Fish ColdBloodedVertebrate) (disjointDecomposition
ColdBloodedVertebrate Amphibian Fish Reptile) (=>
   (instance ?FISH Fish)
   (exists
      (?WATER)
      (and
         (inhabits ?FISH ?WATER)
         (instance ?WATER Water))))
```

3.2 The Ontology of a Semantic Radical: A Generative Lexicon Approach

One illuminating discovery that we make while trying to map radicals to ontology nodes is that each radical actually represents a cluster of concepts that can be associated to the core meaning by a set of rules. We take the 艸(cao3, ‘grass’) radical for instance. It is generally accepted that 艸 represents the concept ‘plant’.

Of the 444 characters containing the semantic symbol 艸, there is no doubt that they are all related to the concept ‘plant’. But what is interesting is that the conceptual clustering is not simply of taxonomic classification. As seem in figure 1, there are four productive relations described by the radical: being a kind of plant (e.g., 蘭 (‘orchid’)), being a part of a plant (e.g., 葉 (‘leaves’)), being a description of a plant (e.g.,落 (‘fallen’)), and being the usage of a plant (e.g. 藥 (‘medicine’)). The concepts of most radicals that represent concrete objects can be classified into name, part, description and usage. For example, the concepts represented by radical 馬 (‘horse’), 牛 (‘cow’), and 木 (‘wood’) also could be divided into the same four classes.

We observe that this is similar with theory of generative lexicon, where *formal, constitutive, telic* and *agentive* are the four parts of the qualia structure

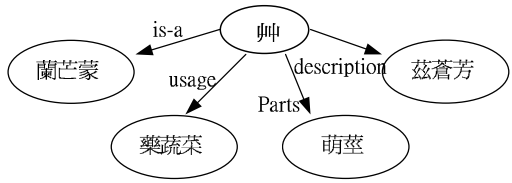


Fig. 1. Conceptual classes represented by semantic radical 艸 (grass), with derived characters 蘭 (orchid), 葉 (leaves), 落 (fallen), 藥 (medicine) and so on

of a word which describe motivated semantic changes and coercions [13]. It is interesting to note that all except the Agentive aspect were attested with the conceptual clustering of Chinese characters derived from the grass radical. Since Pustejovsky's Agentive aspect is strongly associated with artifacts and other human creations, it is not unreasonable that the radicals based on natural objects lack any obvious semantic extension on how it was created. In addition, the descriptive attributes can be subsumed by the formal aspect of the qualia structure.

Indeed, the Agentive aspect is attested by a different radical that is conceptually associated with man-made objects. The radical that we take as example is 金 (jin1,'metal'). Since metals are not useful to human in its natural form, they are shaped by human to become different tools. In the conceptual clusters classified according to the semantic radical 金 , there is a substantial sub-set defined by how a metal object was made, as in Fig. 2.

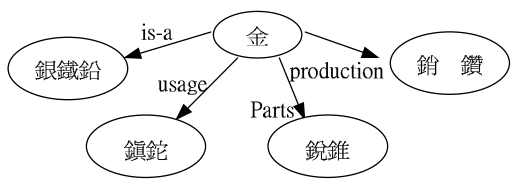


Fig. 2. Conceptual classes represented by semantic radical 金(gold), with derived characters 銀(silver), 鉛(lead), 鑽(diamond), 銳(sharp) and so on

It is also interesting to observe that there is no instantiation of the constitutive aspect for the semantic radical 金. This can be easily explained since metal in its natural form is a mass and does not have any components. Hence we show that the seeming idiosyncrasies in the conceptual clustering under each radical are actually dependent on real world knowledge. Hence we find the conceptual structure of encoded by semantic radicals in the Chinese writing system supports Pustejovesky's theory of Generative Lexicon and qualia structure. These are the same principles used for deriving Chinese characters 3000 years ago suggests that there is cognitive validity.

4 A Sketch on Hanzi-Based Intercultural Collaborative Projects

This section depicts the proposed implementation framework and underlying reasons.

4.1 Hantology as a Prototypical Cross-Cultural Knowledge Platform

In this paper, we propose to elaborate the overall design of Hantology [4] as a prototypical Cross-cultural Knowledge Platform. The methodological and technical considerations are as follows:

First, Hantology is the first comprehensive linguistic ontology of ideographic writing systems. This approach significantly augments knowledge available to the glyph-based Chinese encoding systems. It also allows this systemic knowledge to be applied to facilitate natural language processing. Most important of all, within this framework, the diachronic changes and synchronic variations can be handled by the same mechanism. In other word, cross-lingual variations of Hanzi can be handled with the same mechanism dealing with the variations of Tang usages. (see Fig. 3 and Fig. 4).

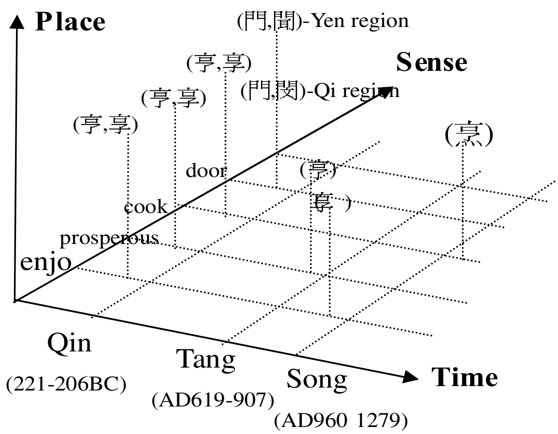


Fig. 3. Character variants: Temporal and locational dependencies

In addition, the glyphs of Chinese characters have undergone historical changes and regional variations, the glyph of each character is different on different period. These relationships are described in Hantology. The descriptions of glyphs include kaishu, lesser-seal, bronze and oraclebone scripts. If two glyphs have evolution relationships, then, hasAncientGlyph and isAncientGlyphOf predicates are used. hasAncientGlyph and isAncientGlyphOf predicates both have inversed and transitive features that are able to infer evolution relationships. The statements of hasAncientGlyph is shown as follows (also see Figure 5):

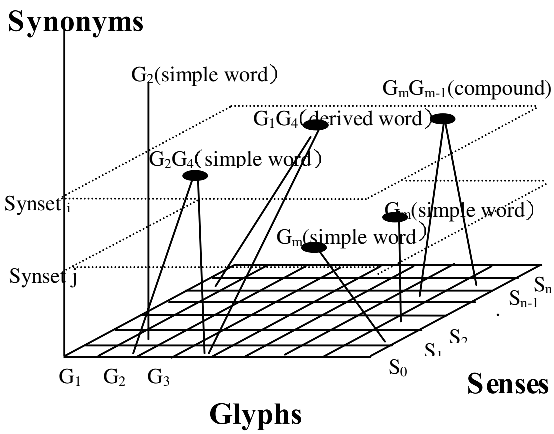


Fig. 4. Words generated from chinese characters

```
if hasAncientGlyph(G$_i$, G$_j$ )
  then  isAncientGlyphOf(G$_j$,G$_i$)

if hasAncientGlyph(G$_i$, G$_j$ ) and hasAncientGlyph(G$_j$,G$_k$)
  then hasAncientGlyph(G$_i$, G$_k$ )
```

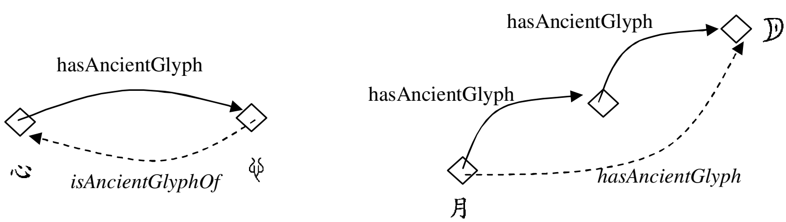


Fig. 5. Ancient glyphs and inferred glyphs

Second, a linguistic context for describing the relation of character variants is proposed in Hantology. Chinese character variants are an important characteristic of Chinese texts. Unfortunately, so far, the relations of variants have not been properly represented. For this, we proposed a linguistic context for describing the relation of variants. Evaluation results show that this linguistic context provide significant improvement over previous counterpart schemes.

Language always changes over time. Any linguistic ontology should not ignore the variation of language. Hantology is the first linguistic ontology describing the variation of languages. The aspects of variation described by Hantology include orthographic form, pronunciation, sense, lexicalization and variants relation. This approach can systematically illustrate the development of Chinese writing system.

Third, Hantology is knowledge-based and equipped with technologies suitable for web services. Fig. 6 and Fig. 7 shows the OWL semantic model of glyph in Hantology and knowledge structure of animal-related radicals interplayed with SUMO ontology.

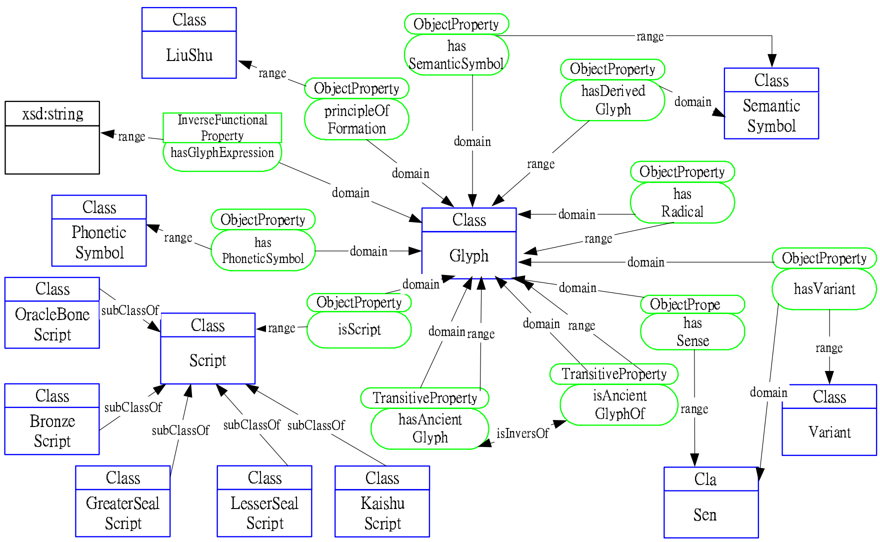


Fig. 6. OWL semantic model of glyph in Hantology

To illustrate the major contents of Hantology, we use the character 臭 ('scent') as an example. Fig. 8 shows the glyphs, pronunciations and variants for 臭. The principle of formation is 會意 ('ideographic compound'), one of the six criteria for character formation 六書 (liu4shu1,'six methods'). Glyph evolution shows the derivational history of the script. 臭 originated as a verb and referred to the act of smelling by nose. There are four variants for the sense of smell. 後作 in the figure means 臭 is replaced by 嗅 to express the sense of smelling later. The first citation appears in period of Tang dynasty(619AD-907AD). For the aim of intercultural colloboration, an example based on Japanese Kanji is shown in Fig. 9.

Lastly, the missing characters and variants retrieval problems are solved. It is an essential requirement to properly represent characters and symbols for any information processing and philological studies as well. However, current Chinese computer systems fail to meet this requirement for decades. Consequently, users

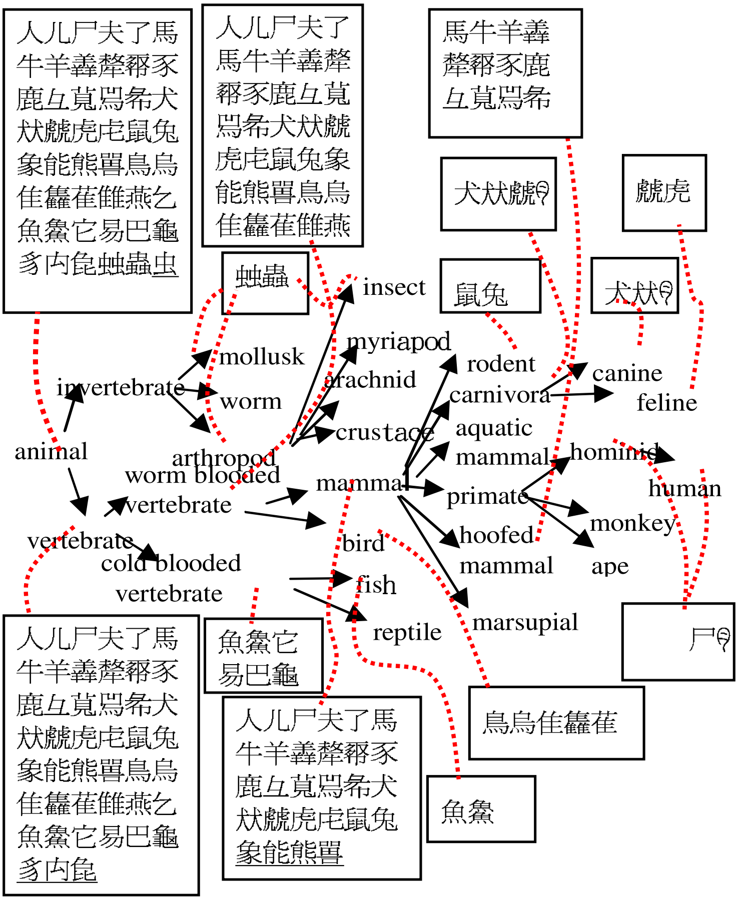


Fig. 7. Knowledge structure of animal-related radicals

always have to face the missing characters and variants retrieval problem. We propose to change the representation of Hanzi to increase the knowledge owned by computers. By integrating missing characters with Hantology, the missing characters and variants problem are solved successfully.

4.2 Meaning-Bearing Radicals as Inter-Lingual-Index

The *Inter-Lingual-Index* (ILI) was proposed by the project of EuroWordNet [15], which is used to connect synsets across all the languages. We propose to adopt *Meaning-bearing Radicals* (MBRs, 意符) to provide an efficient mapping across the CJKV languages on the one side, and other European languages on the other sides.

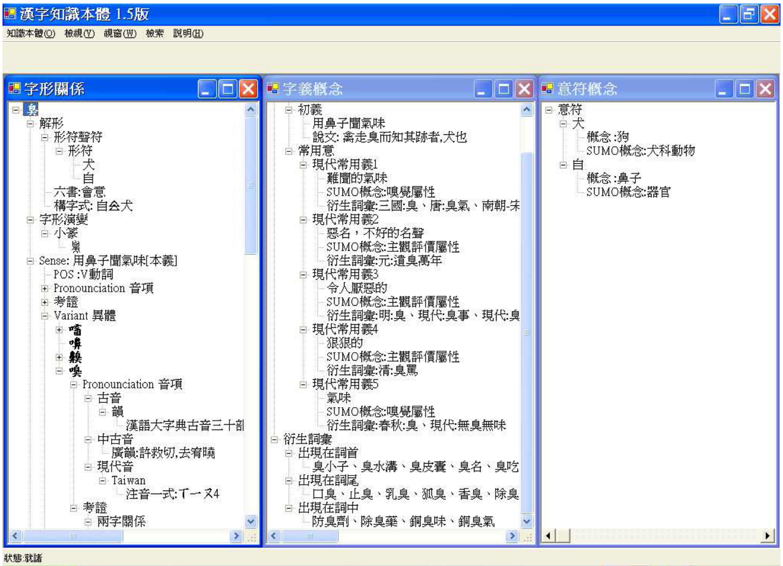


Fig. 8. The glyphs and variants knowledge for 臭

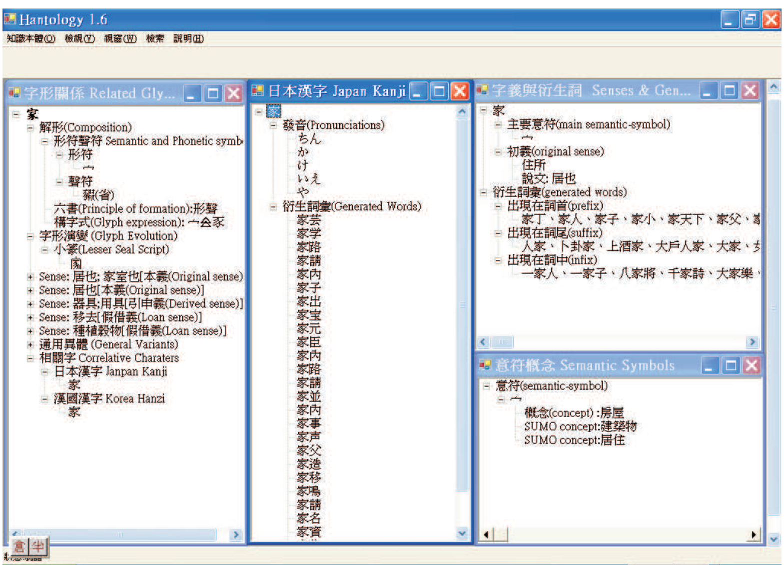


Fig. 9. The kanji example for 家

Formally, given a set of Radicals $\mathcal{R}$, it is also a *power set* of $\mathcal{R}$, written $\mathcal{P}(\mathcal{R})$. Hence we can have $\mathcal{P}(\mathcal{R}) = \{\{c, j\}, \{c, k\}, \{c, v\}, \{c, j, k\}, ...\}$, where c, j, k, v denote Chinese, Japanese, Korean and Vietnamese, respectively. These information can thus facilitate more comparative and contrastive studies of characters variants. [2]

4.3 From Hanzi Ontological Lexical Networks to Hanzi Grid

To achieve the goal of collaborative research, there are many software and hardware implementation architecture available, such as Wiki, Semantic Web and Grid. In order to facilitate the process of automatic learning of qualia structure encoded in Meaning-bearing Radicals, we propose to adopt the *LexFlow* Grid computing environment proposed by Soria et al [14]. These are already in preparation.

5 Conclusion

Chinese characters explicitly encode conventionalized conceptualization. It is well-established practice in computational linguistics to manipulate lexical and inter-lexical level knowledge, such as the very active research based on WordNet. However, the knowledge encoded on Chinese characters is intra-lexical and are embedded in the orthography. In this paper, we focused on how to represent the knowledge structure formed by Chinese characters in the cross-lingual and cross-cultural context. We adopt Hantology as a prototypical formal representation of the linguistic ontology conventionalized with the Chinese writing system. We propose that the radicals, the semantic symbols, do form a robust and well-accepted conceptual system, and can be used as ILI. The historical depth of Hantology will allow us to examine how knowledge systems evolve through time. In addition to the foster the knowledge exchange among the Sinosphere, the richly encoded knowledge at the basic writing level will also support multilingual (CJKV) content processing of texts without higher the syntactic processes of segmentation, chunking, or parsing.

References

1. Bradley, D., LaPolla, R., Michailovsky, B., Thurgood, G. (eds.): Language variation: Papers on variation and change in the Sinosphere and in the Indosphere. Pacific Linguistics, Canberra (2003)
2. Chou, Y.M., Huang, C.R.: Hantology: an ontology based on conventionalized Conceptualization. In: Proceedings of Ontolex Workshop (2005)
3. Chou, Y.M., Huang, C.R.: The construction of the knowledge structure of meaning components of Hanzi. In: The 6th Chinese Lexical Semantics Workshop, Xia-Men, China (2005)

[2] More information such as CJK shared Hanzi and Radicals 中日韓共用漢字表 are available at http://140.111.1.40/fulu/fu5/fu6.htm

4. Chou, Y.M.: Hantology: The knowledge structure of Chinese writing system and its applications. Ph.D Thesis. National Taiwan University (2005)
5. Fellbaum, C.: WordNet: An electronic lexical database. MIT Press, Cambridge (1998)
6. Harris, R.: Rethinking writing. The Athlone Press (2000)
7. Huang, C.R., Chang, R.Y., Li, S.B.: Sinica BOW (Bilingual Ontological Wordnet): Integration of bilingual wordNet and SUMO. In: Proceedings of the 4th LREC, Lisbon, Portugal (2004)
8. Huang, C.R.: Knowledge representation with Hanzi: The relationship among characters, words, and senses (in Chinese). In: International Conference on Chinese Characters and Globalization, Taipei (2005)
9. Hsieh, S.K., Huang, C.R.: When Conset meets Synset: A preliminary survey of an ontological lexical resource based on Chinese Characters. In: poster paper at COLING/ACL 2006, Sydney (2006)
10. Ishikawa, T.: Mojikyo Font Database (1999), `http://www.mojikyo.org/`
11. Juang, D.M., Hsieh, C.C.: The construction and applications of Chinese characters database (in Chinese). In: International Conference on Chinese Characters and Globalization, Taipei, Taiwan (2005)
12. Niles, I., Pease, A.: Toward a standard upper ontology. In: Proceedings of the 2nd International Conference on Formal Ontology in Information Systems, Ogunquit, Maine (2001)
13. Pustejovsky, J.: The generative lexicon. MIT Press, Cambridge (1995)
14. Soria, C., Tesconi, M., Marchetti, A., Bertagna, F., Monachini, M., Huang, C.R., Calzolari, N.: Towards agent-based cross-lingual interoperability of distributed lexical resources. In: COLING/ACL post-conference Workshop Multilingual Languageg Resources and Processing, Sydney (2006)
15. Vossen, P.: EuroWordNet: A multilingual database with lexical semantic networks. Kluwer Academic Publishers, Dordrecht (1998)
16. Wong, S.H., Pala, K.: Chinese characters and top ontology in EuroWordNet. In: Singh, U.N. (ed.) Proceedings of the First Global WordNet Conference 2002, Indian (2002)
17. Xu, S.: (121). Shuo-Wen-Jie-Zi (The Explanation of Words and the Parsing of Characters). Cited edition, ZhongHua (2004)

Fostering Intercultural Collaboration: A Web Service Architecture for Cross-Fertilization of Distributed Wordnets

Francesca Bertagna[1], Monica Monachini[1], Claudia Soria[1], Nicoletta Calzolari[1], Chu-Ren Huang[2], Shu-Kai Hsieh[2], Andrea Marchetti[3], and Maurizio Tesconi[3]

[1] Istituto di Linguistica Computazionale-CNR
via Moruzzi 1, Pisa, Italy
francesca.bertagna@ilc.cnr.it,
monica.monachini@ilc.cnr.it,
claudia.soria@ilc.cnr.it,
nicoletta.calzolari@ilc.cnr.it

[2] Academia Sinica
Nankang, Taipei, Taiwan
churen@gate.sinica.edu.tw,
shukai@gmail.com

[3] Istituto di Informatica e Telematica-CNR
via Moruzzi 1, Pisa, Italy
andrea.marchetti@iit.cnr.it,
maurizio.tesconi@iit.cnr.it

Abstract. Enhancing the development of multilingual lexicons is of foremost importance for intercultural collaboration to take place, as multilingual lexicons are the cornerstone of several multilingual applications. However, the development and maintenance of large-scale, robust multilingual dictionaries is a tantalizing task. In this paper we present a tool, based on a web service architecture, enabling semi-automatic generation of bilingual lexicons through linking of distributed monolingual lexical resources. In addition to lexicon development, the architecture also allows enrichment of monolingual source lexicons through exploitation of the semantic information encoded in corresponding entries. In the paper we describe our case study applied to the Italian and Chinese wordnets, and we illustrate how the architecture can be extended to access distributed multilingual WordNets over the Internet, paving the way to exploitation in a cross-lingual framework of the wealth of information built over the last decade.

Keywords: distributed language resources, interoperable lexical resources, integration of WordNets.

1 Introduction

Enhancing the development of multilingual lexicons is of foremost importance for intercultural collaboration to take place, as multilingual lexicons are the cornerstone of several multilingual applications (such as cross-language QA and IR,

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 146–158, 2007.

machine translation, terminology management, multilingual computing, etc.). Nevertheless, large-scale multilingual lexical resources are not as widely available and are very costly to construct: the work process for manual development of new lexical resources or for tailoring existing ones is too expensive in terms of required effort and time to be practically attractive.

The previous trend in lexical resource was oriented to maximization of effort by building large-scale, general-purpose lexicons. However, these lexical resources are not always satisfactory despite the tremendous amount of work needed to build them and the richness and degree of sophistication of the information contained therein. Often lexical resources are unbalanced with respect of the type of lexical information encoded, focusing on a particular type and not providing enough coverage of other aspects. In some other cases, lexical resources are too much or too little detailed for the specific purposes of an application. On the other hand, the market is increasingly calling for new types of lexical resources: lexicons that can be built rapidly, possibly by combining certain types of information while discarding other, and tailored to specific needs and requirements. Rather than building new lexical resources, the new trend focuses on trying to exploit the richness of existing lexicons.

To meet these needs, lexical resources need to be made available, to be constantly accessed by different types of users, who may want to select different portions of the same resource, or may need to combine information coming from different resources. This scenario no longer leaves space to static, closed, and locally managed repositories of lexical information; instead, it calls for an environment where lexical resources can be shared are reusable, and are openly customizable. At the same time, as the history of the web teaches, it would be a mistake to create a central repository containing all the shared lexical resources because of the difficulties to manage it. Distribution of resources thus becomes a central concept: the solution proposed by the lexical resource community thus consists in moving towards distributed language services, based on open content interoperability standards, and made accessible to users via web-services technologies.

There is another, deeper argument in favor of distributed lexical resources: language resources, lexicons included, are inherently distributed because of the diversity of languages distributed over the world. It is not only natural that language resources to be developed and maintained in their native environment. Since language evolves and changes over time, it is not possible to describe the current state of the language away from where the language is spoken. Lastly, the vast range of diversity of languages also makes it impossible to have one single universal centralized resource, or even a centralized repository of resources. Having lexical resources available as web services would allow to create new resources on the basis of existing ones, to exchange and integrate information across repositories, and to compose new services on demand: an approach towards the development of an infrastructure built on top of the Internet in the form of distributed language services is presented in [2].

This new type of language resources can still be stored locally, but its maintenance and exploitation can be a matter of agents being choreographed to act over them. Admittedly, this is a long-term scenario requiring the contribution of many different actors and initiatives (among which we only mention standardization, distribution and international cooperation). The first prerequisite for this scenario to take place is to ensure true interoperability among lexical resources, a goal that is long being addressed to by the standardization community and that is now mature. Although the paradigm of distributed and interoperable lexical resources has largely been discussed and invoked, very little has been made in comparison for the development of new methods and techniques for its practical realization. Some initial steps are made to design frameworks enabling inter-lexica access, search, integration and operability. An example is the Lexus tool [3], based on the Lexical Markup Framework [4], that goes in the direction of managing the exchange of data among large-scale lexical resources. A similar tool, but more tailored to the collaborative creation of lexicons for endangered language, is SHAWEL [5]. However, the general impression is that little has been made towards the development of new methods and techniques for attaining a concrete interoperability among lexical resources. In this paper we present a tool, based on a web service architecture, fostering the integration and interoperability of computational lexicons, focusing on the particular case of mutual linking and cross-lingual enrichment of distributed monolingual lexical resources. As a case-study, we have chosen to work with two lexicons belonging to the WordNet family, the ItalWordNet [6] and Sinica BOW [7]. The development of this application is intended as a case-study and a test-bed for trying out needs and requirements posed by the challenge of semi-automatic integration and enrichment of practical, large-scale multilingual lexicons for use in computer applications. The paper is organized as follows: Sect. 2 describes the general architectural design of our project; Sect. 3 describes the tool taking care of cross-lingual integration of lexical resources, while a case-study involving an Italian and Chinese lexicons is presented in Sect. 4. Section 5 briefly explains how this tool can be integrated in a more general framework for the semi-automatic management of lexical resources.

2 An Architecture for Integrating Lexical Resources

Designing a general architecture able to turn into reality the vision of shared and distributed lexical repositories is a very challenging task. We designed a distributed architecture to enable a rapid prototyping of cooperative applications for integrating lexical resources. This architecture is articulated in three layers:

- The lower layer consists of a sort of *meta–wordnet*, i.e. a grid of local wordnets realized as a virtual repository of XML databases residing at different locations and accessible through web services. Basic software services are also necessary, such as an UDDI server for the registration of the local wordnets and web services dedicated to the coherent management of the different versions of WordNet the databases refer to.

– The middle layer hosts diverse applications that exploit the wordnets grid. The so-called MultiWordNet Service (MWS, Section 3) was built as a proof of concept of the possibility to mutually enrich wordnets in a distributed environment; other, more advanced NLP applications (in particular multilingual) can be developed by exploiting the availability of the WordNet grid.
– A higher layer, called "cooperative layer" or *LeXFlow* is intended as an overall environment where all the modules realized in the lower layers are integrated in a comprehensive workflow of human and software agents.

Figure 1 illustrates the general architecture. In this paper we concentrate on the description of the middle layer (see Sect. 3). A more detailed description of the cooperative layer can be found in [8] and [9].

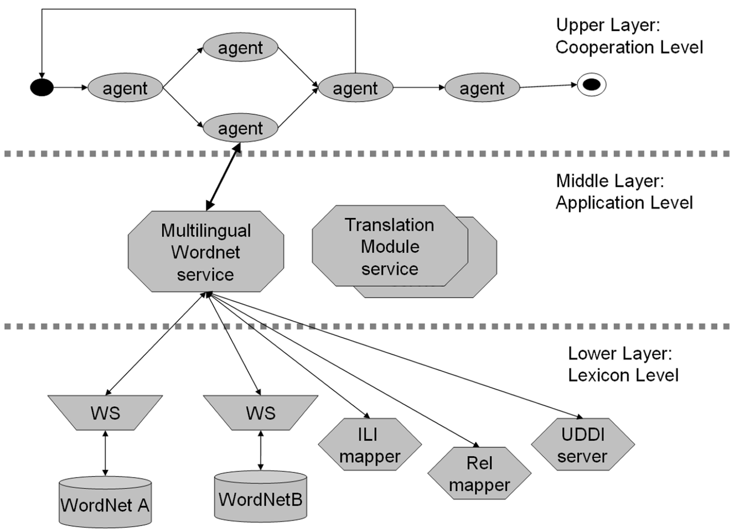


Fig. 1. A three–layered architecture for integrating lexical resources

In Sect. 5 we illustrate how the general LeXFlow environment could accommodate the tool described as a module of a general architecture geared towards lexicon management.

3 Multilingual WordNet Service

In this section we present a tool that addresses the issue of lexicon augmentation or enrichment focusing on mutual enrichment of two wordnets. This module, named "Multilingual WordNet Service" is responsible for the *automatic cross-lingual fertilization of lexicons* having a WordNet-like structure. Put it very simply, the idea behind this module is that a monolingual wordnet can be enriched by accessing the semantic information encoded in corresponding entries of other monolingual wordnets.

Since each entry in the monolingual lexicons is linked to the Interlingual Index (ILI, cf. Sect. 3.1), a synset of a WN(A) is indirectly linked to another synset in another WN(B). On the basis of this correspondence, a synset(A) can be enriched by importing the relations that the corresponding synset(B) holds with other synsets(B), and vice-versa. Moreover, the enrichment of WN(A) will not only import the relations found in WN(B), but it will also propose target synsets in the language(A) on the basis of those found in language(B). The various WN lexicons reside over distributed servers and can be queried through web service interfaces. The overall architecture for multilingual wordnet service is depicted in Fig. 2.

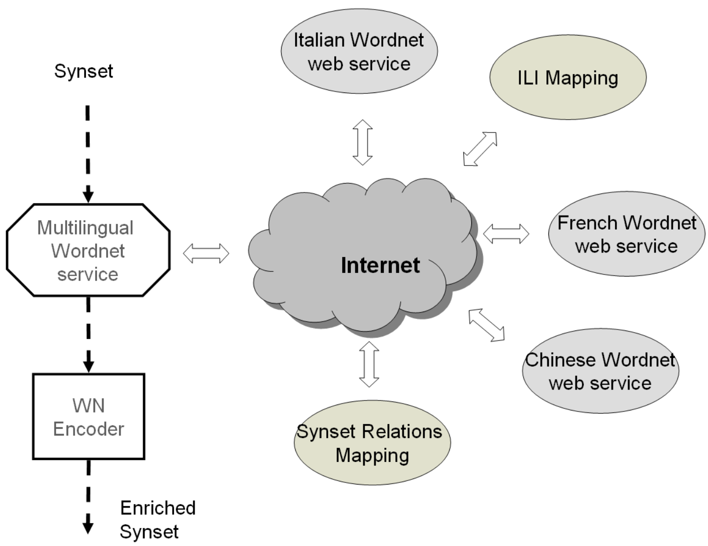


Fig. 2. Multilingual wordnet service architecture

Put in the framework of the general LeXFlow architecture, the Multilingual wordnet Service can be seen as an additional external software agent that can be added to the augmentation workflow or included in other types of lexical flows. For instance, it can be used not only to enrich a monolingual lexicon but to bootstrap a bilingual lexicon.

3.1 Linking Lexicons Through the ILI

The entire mechanism of the Multilingual WN Service is based on the exploitation of Interlingual Index [11], an unstructured version of WordNet used in EuroWordNet [10] to link wordnets of different languages; each synset in the language-specific wordnet is linked to at least one record of the ILI by means of a set of equivalence relations (among which the most important is the EQ_SYNONYM, that expresses a total, perfect equivalence between two synsets). Figure 6 describes the schema of a WN lexical entry. Under the root "synset" we find both internal relations ("synset relations") and ILI Relations, which link to ILI synsets. Figure 3 shows the role played by the ILI as set of pivot nodes allowing the linkage between concepts belonging to different wordnets.

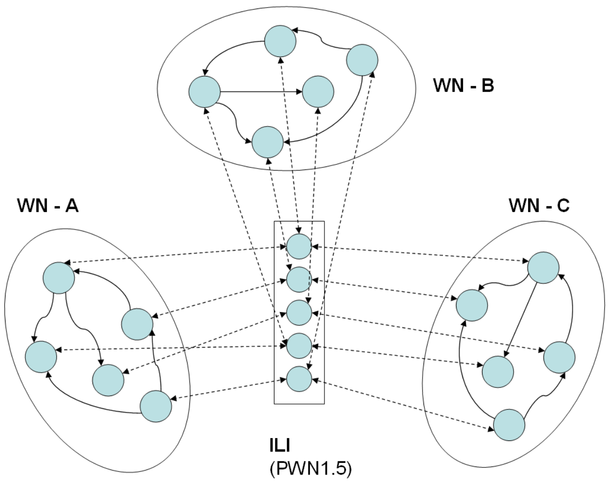


Fig. 3. Interlingual linking of language-specific synsets

In the Multilingual WN Service, only equivalence relations belonging to the types EQ_SYNONYM and EQ_NEAR_SYNONYM have been taken into account, being them the ones used to represent a translation of concepts and also because they are the most exploited (for example, in IWN, they cover about the 60% of the encoded equivalence relations). The EQ_SYNONYM relation is used to realize the one-to-one mapping between the language-specific synset and the ILI, while multiple EQ_NEAR_SYNONYM relations (because of their nature) might be encoded to link a single language-specific synset to more than one ILI record. In Fig. 4 we represented the possible relevant combinations of equivalence relations that can realize the mapping between synsets belonging to two languages. In all the four cases, a synset "a" is linked via the ILI record to a synset "b" but a specific procedure has been foreseen in order to calculate different "plausibility scores" to each situation. The procedure relies on different rates assigned to the two equivalence relations (rate "1" to EQ_NEAR_SYNONYM relation and rate "0" to the EQ_SYNONYM). In this way we can distinguish the four cases by assigning respectively a weight of "0", "1", "1" and "2". The ILI is a quite powerful yet simple method to link concepts across the many lexicons belonging to the WordNet-family. Unfortunately, no version of the ILI can be considered a standard and often the various lexicons exploit different version of WordNet as ILI. This is a problem that is handled at web-service level, by incorporating the conversion tables provided by [11]. In this way, the use of different versions of WN does not have to be taken into consideration by the user who ac-cesses the system but it is something that is resolved by the system itself. This is why the version of the ILI is a parameter of the query to web service (see Sect. below).

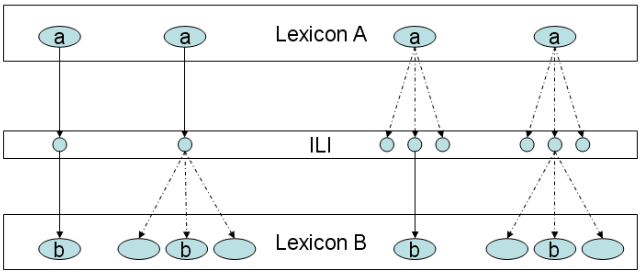


Fig. 4. Possible combinations of relations between two lexicons A and B and the ILI

3.2 Description of the Procedure

On the basis of ILI linking, a synset can be enriched by importing the relations contained in the corresponding synsets belonging to another wordnet. In the procedure adopted, the enrichment is performed on a synset-by-synset basis. In other words, a certain synset is selected from a wordnet resource, say WN(A). The cross-lingual module identifies the corresponding ILI synset, on the basis of the information encoded in the synset. It then sends a query to the WN(B) web service providing the ID of ILI synset together with the ILI version of the starting WN. The WN(B) web service returns the synset(s) corresponding to the WN(A) synset, together with reliability scores. If WN(B) is based on a different ILI version, it can carry out the mapping between ILI versions (for instance by querying the ILI mapping web service). The cross-lingual module then analyzes the synset relations encoded in the WN(B) synset and for each of them creates a new synset relation for the WN(A) synset. If the queried wordnets do not use the same set of synset relations, the module must take care of the mapping between different relation sets. In our case-study no mapping was needed, since the two sets were completely equivalent. Each new relation is obtained by substituting the target WN(B) synset with the corresponding synset WN(A), which again is found by querying back the WN(A) web service (all these steps through the ILI). The procedure is formally defined by the following formula:

$$\text{Let} \quad a_j \in A \tag{1}$$

$$\text{Let} \quad Ba_j = \{b_i \mid b_i \in B \quad \text{and} \quad (b_i mathrmILI a_j)\} \tag{2}$$

$$\forall b_i \in Ba_j \tag{3}$$

$$\text{Let} \quad R_i = \{b_i r_k b_p \mid b_i, b_p \in B \quad \text{and} \quad (r_k \in R_A \cap R_B)\} \tag{4}$$

$$\forall b_i r_k b_p \in R_i \tag{5}$$

$$\text{Let} \quad Ab_p = \{a_i \mid a_i \in A \quad (a_i \text{ILI} b_p)\} \tag{6}$$

$$\forall a_t \in Ab_p \tag{7}$$

$$a_j r_k a_t \quad \text{is} \quad \text{a} \quad \text{candidate} \quad \text{relation} \tag{8}$$

where A, B are the lexicons; a_j, b_i the synsets; $a_j r_p a_i$ is the synset relation r_p between a_j and a_i; b_i ILI a_j means that b_i is connected via ILI to a_j; R_A and R_B is the relation space of lexicons A and B, and $R_A \in R_B$ is the common relation space of A and B.

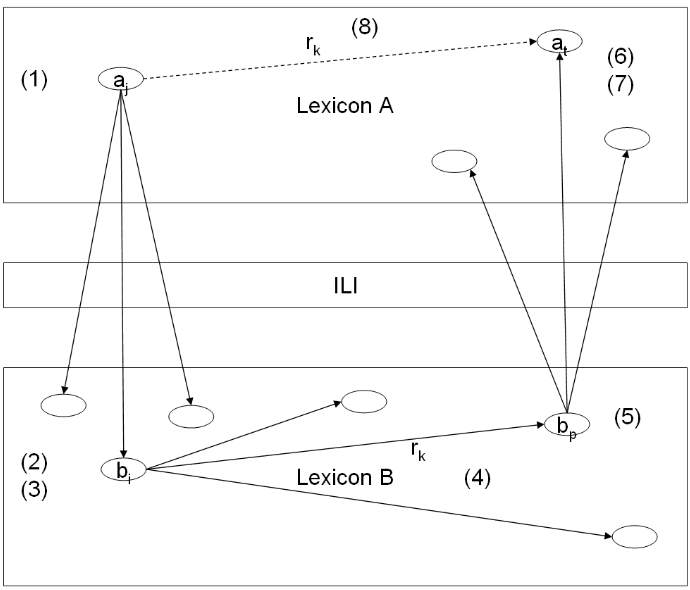


Fig. 5. Finding new relations

Every local wordnet has to provide a web service API with the following methods:

1. GetWeightedSynsetsByIli(ILIid, ILIversion)
2. GetSynsetById(sysnsetID)
3. GetSynsetsByLemma(lemma)

The returned synsets of each method must be formatted in XML following the schema depicted in Fig. 6. The scores returned by the method "GetWeightedSynsetsByIli" are used by our module to calculate the reliability rating for each new proposed relation.

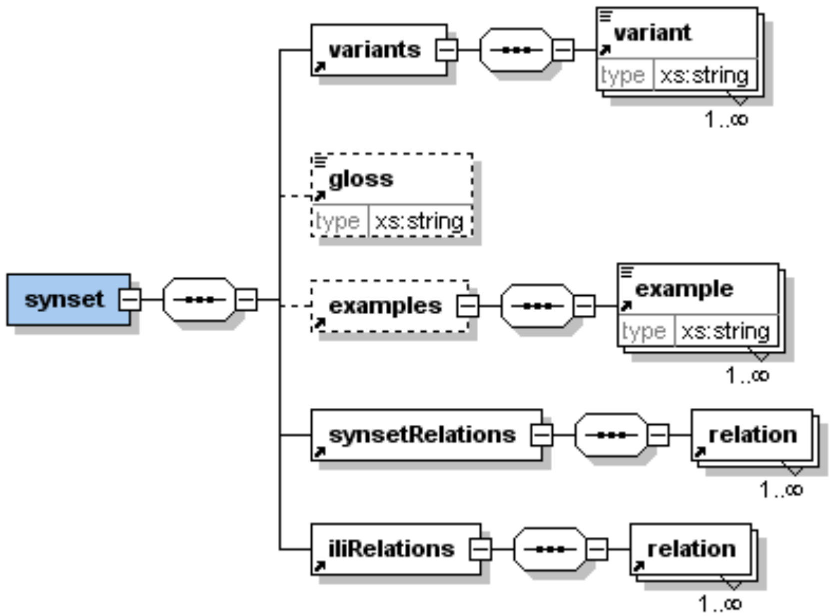


Fig. 6. Schema of wordnet synsets returned by WN web services

4 A Case Study: Cross-Fertilization Between Italian and Chinese Wordnets

We explore this idea with a case-study involving the ItalWordNet [6] and the Academia Sinica Bilingual Ontological Wordnet (Sinica BOW, [7]). The BOW integrates three resources: WordNet, English-Chinese Translation Equivalents Database (ECTED), and SUMO (Suggested Upper Merged Ontology). With the integration of these three key resources, Sinica BOW functions both as an English-Chinese bi-lingual wordnet and a bilingual lexical access to SUMO. Sinica Bow currently has two bilingual versions, corresponding to WordNet 1.6. and 1.7. Based on these bootstrapped versions, a Chinese Wordnet (CWN, [12]) is under construction with handcrafted senses and lexical semantic relations. For the current experiment, we have used the version linking to WordNet 1.6. ItalWordNet was realized as an extension of the Italian component of EuroWordNet. It comprises a general component consisting of about 50,000 synsets and terminological wordnets linked to the generic wordnet by means of a specific set of relations. Each synset of ItalWordNet is linked to the Interlingual-Index (ILI). The two lexicons refer to different versions of the ILI (1.5 for IWN and 1.6 for BOW), thus making it necessary to provide a mapping between the two versions. On the other hand, no mapping is necessary for the set of synset relations used, since both of them adopt the same set. For the purposes of evaluating the cross-lingual module, we have developed two web-services for managing a subset of the two resources. Figure 7 shows a very simple example where our procedure discovers and proposes a new meronymy relation for the Italian synset {*passaggio,strada,via*}. This synset is equivalent to the ILI {*road, route*} that is ILI-connected with BOW synset {*dao_lu, dao, lu*} (7, A) . The Chinese synset has a meronymy relation with the synset (wan) (B). This last synset is equivalent to the ILI {*bend, crook, turn*} that is ILI-connected with Italian WordNet synset {*curvatura, svolta, curva*} (C). Therefore the procedure will propose a new candidate meronymy relation between the two Italian WordNet synsets (D).

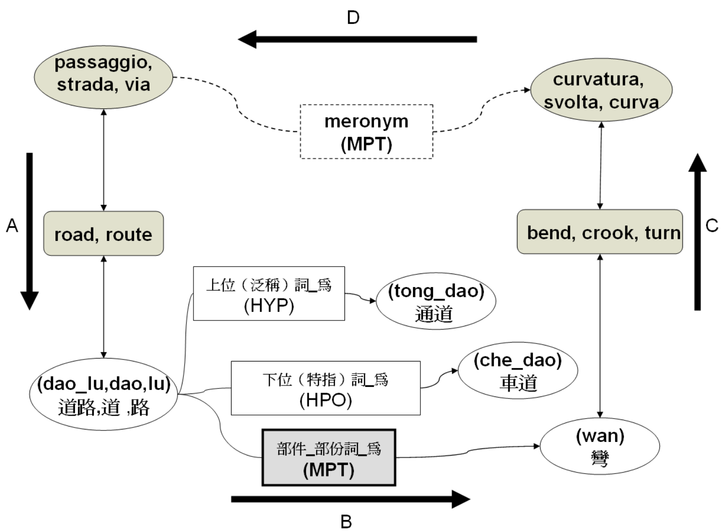


Fig. 7. Example of a new proposed meronymy relation for Italian

4.1 Considerations and Lessons Learned

Given the diversity of the languages for which wordnets exist, we note that it is difficult to implement an operational standard across all typologically different languages. Work on enriching and merging multilingual resources presupposes that the resources involved are all encoded with the same standard. However, even with the best efforts of the NLP community, there are only a small number of language resources encoded in any given standard. In the current work, we presuppose a de-facto standard, i.e. a shared and conventionalized architecture, the WordNet one. Since the WordNet framework is both conventionalized and widely followed, our system is able to rely on it without resorting to a more substantial and comprehensive standard. In the case, for instance, of integration of lexicons with different underlying linguistic models, the availability of the MILE [13] was an essential prerequisite of our work. Nevertheless, even from the perspective of the same model, a certain degree of standardization is required, at least at the format level. From a more general point of view, and even from the perspective of a limited experiment such as the one described in this paper, we must note that the realization of the new vision of distributed and interoperable language resources is strictly intertwined with at least two prerequisites. On the one side, the language resources need to be available over the web; on the other, the language resource community will have to reconsider current distribution policies, and to investigate the possibility of developing an "Open Source" concept for LRs.

5 LeXFlow

This MWNS can run as an individual system, but it has to be seen more as a software module to be integrated into the general LeXFlow architecture [8], developed with the aim to make the vision of an infrastructure for access and sharing of linguistic resources more tangible. LeXFlow was born as an adaptation to computational lexicons of XFlow, a cooperative web application for the management of document workflows (DW, [14]) and can be considered as both an architecture for proving new cooperation methods among lexicon experts and a general, versatile framework enabling automatic lexical resource integration. The novelty of LeXFlow is that it enables the cooperation of agents, either human or software agents and allows different agents to interact, even residing over distributed places. Since it allows the independent and coordinated sharing of actions over portions of lexicons, LeXFlow naturally lends itself as a tool for the management of distributed lexical resources. The other software modules available, at present, in the flow (defined Lexical Workflow Type, LWT), are a platform for interoperability and integration of monolingual semantic lexicons with differently conceived architectures and diverging formats (such as two Italian lexicons from the SIMPLE [15] and WordNet [6] families) and an automatic acquirer of lexical information from corpora or from the web. This LWT, called "lexicon augmentation", explicitly addresses dynamic augmentation of semantic lexicons, thus allowing the two lexicons to interact by reciprocally enriching themselves and, moreover, to integrate information coming from corpora.

6 Conclusions

Our proposal to make distributed wordnets interoperable has the following applications in processing of lexical resources:

- Enriching existing resources: information is often not complete in any given wordnet: by making two wordnets interoperable, we can bootstrap semantic relations and other information from other wordnets.
- Creation of new resources: multilingual lexicons can be bootstrapped by linking different language wordnets through ILI.
- Validation of existing resources: semantic relation information and other synset assignments can be validated when it is reinforced by data from a different wordnet.

In particular, our work can be proposed as a prototype of a web application that would support the Global WordNet Grid initiative (www.globalwordnet.org/gwa/gwa_grid.htm). Any multilingual process, such as cross-lingual information retrieval, must involve both resources and tools in a specific language and language pairs. For instance, a multilingual query given in Italian but intended for querying English, Chinese, French, German, and Russian texts, can be sent to five different nodes on the Grid for query expansion, as well as performing the query itself. In this way, language–specific query techniques can be applied in parallel to achieve best results that can be integrated in the future. As multilingualism clearly becomes one of the major challenges of the future of web-based knowledge engineering, WordNet emerges as one leading candidate for a shared platform for representing a lexical knowledge model for different languages of the world. This is true even if it has to be recognized that the wordnet model is lacking in some important semantic information (like, for instance, a way to represent the semantic predicate). However, such knowledge and resources are distributed. In order to create a shared multilingual knowledge base for crosslingual processing based on these distributed resources, an initiative to create a grid–like structure has been recently proposed and promoted by the Global WordNet Association, but until now has remained a wishful thinking. The success of this initiative will depend on whether there will be tools to access and manipulate the rich internal semantic structure of distributed multilingual WordNets. We believe that our work on LeXFlow offers such a tool to provide interoperable web-services to access distributed multilingual WordNets on the grid. This allows us to exploit in a cross-lingual framework the wealth of monolingual lexical information built in the last decade.

References

1. Calzolari, N., Soria, C.: A New Paradigm for an Open Distributed Language Resource Infrastructure: the Case of Computational Lexicons. In: Proceedings of the AAAI Spring Symposium on Knowledge Collection from Volunteer Contributors (KCVC05), vol. 78, pp. 110–114 (2005)
2. Ishida, T.: Language Grid: An Infrastructure for Intercultural Collaboration. In: IEEE/IPSJ Symposium on Applications and the Internet (SAINT-06), pp. 96–100 (2006)
3. Kemps–Snijders, M., Nederhof, M., Wittenburg, P.: LEXUS, a web-based tool for manipulating lexical resources. In: Proceedings of LREC2006, Genoa, Italy, pp. 1862–1865 (2006)
4. Francopoulo, G., George, M., Calzolari, N., Monachini, M., Bel, N., Pet, M., Soria, C.: Lexical Markup Framework (LMF). In: Proceedings of LREC2006, Genoa, Italy, pp. 233–236 (2006)
5. Gulrajani, G., Harrison, D.: SHAWEL: Sharable and Interactive Web-Lexicons. In: Proceedings of the LREC2002 Workshop on Tools and Resources in Field Linguistics, Las Palmas, Canary Islands, Spain, pp. 1–4 (2002)
6. Roventini, A., Alonge, A., Bertagna, F., Calzolari, N., Girardi, C., Magnini, B., Marinelli, R., Zampolli, A.: ItalWordNet: Building a Large Semantic Database for the Automatic Treatment of Italian. In: Computational Linguistics in Pisa, IEPI, Pisa-Roma, pp. 745–791 (2003)

7. Huang, C., Chang, R., Lee, S.: Sinica BOW (Bilingual Ontological Wordnet): Integration of Bilingual WordNet and SUMO. In: Proceedings of the LREC2004, Lisbon, Portugal, pp. 1553–1556 (2004)
8. Soria, C., Tesconi, M., Bertagna, F., Calzolari, N., Marchetti, A., Monachini, M.: Moving to Dynamic Computational Lexicons with LeXFlow. In: Proceedings of LREC2006, Genoa, Italy, pp. 7–12 (2004)
9. Tesconi, M., Marchetti, A., Bertagna, F., Monachini, M., Soria, C., Calzolari, N.: LeXFlow: a System for Cross-fertilization of Computational Lexicons. In: Proceedings of the COLINGACL 2006 Interactive Presentation Sessions, Sydney, Australia, pp. 9–12 (2006)
10. Vossen, P.: Introduction to EuroWordNet. Computers and the Humanities 32(2-3), 73–89 (1998)
11. Daudé, J., Padró, L., Rigau, G.: A Complete WN1.5 to WN1.6 Mapping. In: Proceedings of NAACL Workshop on WordNet and Other Lexical Resources: Applications, Extensions and Customizations. Association for Computational Linguistics, Pittsburg, PA, USA, pp. 83–88 (2001)
12. Huang, C., Chen, C., Weng, C., Lee, H., Chen, Y., Chen, K.: The Sinica Sense Management System: Design and Implementation. Computational Linguistics and Chinese Language Processing 10(4), 417–430
13. Calzolari, N., Bertagna, F., Lenci, A., Monachini, M. (ed.): Standards and Best Practice for Multilingual Computational Lexicons. MILE (the Multilingual ISLE Lexical Entry). ISLE CLWG Deliverable D2.2 and 3.2. Pisa (2003)
14. Marchetti, A., Tesconi, M., Minutoli, S.: XFlow: An XML-Based Document-Centric Workflow. In: Ngu, A.H.H., Kitsuregawa, M., Neuhold, E.J., Chung, J.-Y., Sheng, Q.Z. (eds.) WISE 2005. LNCS, vol. 3806, pp. 290–303. Springer, Heidelberg (2005)
15. Ruimy, N., Monachini, M., Gola, E., Calzolari, N., Del Fiorentino, C., Ulivieri, M., Rossi, S.: A Computational Semantic Lexicon of Italian: SIMPLE. In: Computational Linguistics in Pisa, IEPI, Pisa-Roma, pp. 821–864 (2003)

MCD: A Joint Semantic Project on East Asian Languages

Hong Zhu and Yang Liu

Institute of Computational Linguistics, Peking University
Beijing 100871, China
{zh,liuyang}@pku.edu.cn

Abstract. The Multilingual Concept Dictionary is a NSFC/KOSEF joint project for natural language processing, now in its second year.[1] The project's key feature is to merge and align various ontologies and lexical resources of East Asian languages into a compatible one, especially for Chinese, Korean and Japanese. This paper as a progress report first introduces our new knowledge representation method for ontology construction, which is called "Upper Ontology Tree" plus "Lexical Semantics N-tuples", and then analyzes the design of our dictionary framework especially the higher-level part, in which Korean and Japanese ontologies are mapped to the shared semantic hierarchy by semi-automatic methods.

Keywords: Lexical resource, MCD, Multilingual, Ontology.

1 Introduction

Nowadays, the understanding and processing of content information in different languages begins to draw people's attention. The study on semantic analysis, such as Word Sense Disambiguation (WSD) and semantic induction/ reasoning, has gradually been put under the spotlight in natural language processing. Since computer itself has no intelligence, a semantic language resource, especially a multilingual language resource, is needed as a medium to help it to fulfill the task of "understanding" human languages and as a material prerequisite to make it "more and more bright".

Multilingual Semantic language resources are useful for many NLP applications, among which the famous one is EuroWordNet [1], which links together languages such as Dutch, Italian, Spanish, German, French, Czech, and Estonian with the American English WordNet [2]. The Suggested Upper Merged Ontology (SUMO) [3] is another one. It is developed by the IEEE Standard Upper Ontology Working Group as a standard upper ontology. It is intentionally limited to concepts that are thought as meta, generic, abstract or philosophical. We may say that it is language-independent. We also find some other influential ontologies in English, such as WordNet, FrameNet [4] and

[1] MCD project is a NSFC-KOSEF Scientific Cooperation Program named as "Construction of the Multilingual Concept Dictionary with Its Application Research". It is also partially supported by the National Hi-tech R&D Program of China, the FANEDD Project of Chinese Ministry of Education, under the grant No. 2004CB318102 and 200514 respectively. Thanks also go to PKU-Fujitsu Young Scholar Foundation.

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 159–169, 2007.

MindNet [5]. Besides, multilingual resources in East Asian languages have become more and more important and also achieved great development, especially for Chinese, Japanese and Korean. EDR [6] and Goi-Taikei [7] in Japanese, CoreNet [8] in Korean, HowNet [9] in Chinese are the main East Asian language resources now available. Chinese Concept Dictionary (CCD) [10] is a WordNet-like Chinese-English bilingual lexical resource constructed by the Institute of Computational Linguistics of Peking University (ICL). Sinca BOW [11], which is developed by Academia Sinica, is another famous bilingual resource that integrates WordNet, English-Chinese Translation Equivalents Database (ECTED), and SUMO.

As a joint semantic project on East Asian languages, The Multilingual Concept Dictionary (MCD) aims at merging and aligning various ontologies and lexical resources of East Asian languages into a compatible one.

In this paper, we discuss the state-of-the-art of the lexical resource and the methodology used in current work. We first introduce in section 2 our ontology construction methodology. Then in next section we introduce our new ontology knowledge representation method called "Upper Ontology Tree" plus "Lexical Semantic N-tuples". In section 4, the design of MCD higher-level part is presented. Finally, in section 5 we put forth the draft design of MCD lower-level part for the next step work. Finally, MCD database's specification is given briefly.

2 Methodology of MCD Construction

Ontologies and lexical resources vary in methodology of construction, knowledge representation and other aspects. Table 1 and Table 2 show a glance at some famous language resources for reference.

Table 1. Some famous knowledge bases (a)

Dictionary	Time	Developer	Language
WordNet	1985-	American Princeton University	English
EDR	1986-1994	Japan Electronic Dictionary Research Institute	Japanese/English
Goi-Taikei	1994-	Nippon Telegraph and Telephone Corporation	Japanese/English
HowNet	1988-	Zhendong. Dong etc.	Chinese/English
CoreNet	1994-	Korea Advanced Institute of Science and Technology	Korean Chinese Japanese
SUMO	2001-	IEEE Standard Upper Ontology Working Group	English Chinese Czech German Hindi Italian
CCD	2000-	Institute of Computational Linguistics	Chinese/English
Sinca BOW	-2004	Academia Sinica	Chinese/English

Table 2. Some famous knowledge bases (b)

Dictionary	Size	Semantic theory basis
WordNet	117,597 synsets	synsets; semantic relations
EDR	410,000 concepts	semantic categories; semantic relation;
Goi-Taikei	400,000 Japanese words	semantic relation; semantic categories; valence theory
HowNet	147,979 records	sememe; semantic role; semantic relation
CoreNet	51,641 senses	Goi-Taikei like;
SUMO	1,000 terms and 4,000 statements	semantic categories; semantic relation
CCD	About 100,000 synsets	WordNet-like
Sinica BOW	About 100,000 concepts	bilingual ontological Wordnet

The information listed here does not reflect the newest since some of language resources are still in progress. You could access the related webs to get the latest message.

The multilingual nature of MCD raises methodological issues for its design. How to choose a proper construction method is a question of significances. As we know, ontology construction is always a labor-intensive and time-consuming procedure. Researchers of MindNet have developed method for automatically construction. However, the language resources introduced above are all manually or semi-automatically created. This is common since ontology building lacks of commonly established methodology. Based on where data resources come from, we classify the language resources construction methodologies into three categories as follows:

(1) Construction by merging publicly available ontological content into a single, comprehensive, and cohesive structure.
(2) Construction by extracting words from corpus or linguistic dictionaries and building concept taxonomy on their own philosophy.
(3) Construction by extracting concept description and concept taxonomy from given corpus

Method (1) is characterized by its concentration on the technologies of aligning and merging different ontologies varied from languages and taxonomies. SUMO and CoreNet are cases in point.

The main procedure to create SUMO is combining various ontologies (e.g. ITBM-CNR, John Sowa's upper-level ontology, Russell and Norvig's upper-level ontology, etc) into a single one. The procedure constitutes "syntactic merge" and "semantic merge". Similar to SUMO, by assigning Goi-Taikei's semantic category to each sense of the nouns in MRD, CoreNet's developers combine these two resources and produce an expanded lexicalized noun thesaurus, which become the basis structure for CoreNet.

WordNet and HowNet are constructed by using method (2), each with its own underlying philosophy. WordNet is widely used and famous for its psycholinguistic theories including separability hypothesis, patterning hypothesis, comprehensiveness hypothesis. Unlike WordNet, HowNet holds that all physical and non-physical matters undergo a constant process of motion and change in a specific space. This change is usually reflected by a change in state, which in turn is manifested by a change in the value of certain attributes. Under this hypothesis, the most pivotal issue is to define a close set of attributes, known as sememes.

EDR's Concept Dictionary is constructed by using method (3). Its concept dictionaries consist of concept descriptions and concept taxonomy. The initial 1,000,000 practical sentences were syntactically and semantically analyzed through manual work. The concept-concept, concept-category, or category-category relations, and terms used to represent concepts are all derived from the corpora.

For pragmatic reasons we have chosen method (1) for MCD construction. The concept framework of this language resource is based on the SUMO and WordNet. WordNet is now mapped to SUMO [12]. Since it adopts the same framework as WordNet, CCD can be mapped onto SUMO using the same pointers. This SUMO-WordNet-CCD framework establishes our MCD base part (see Fig.1.). Goi-Taikei and CoreNet are mapped to the framework to provide the Korean and Japanese lemmas for MCD.

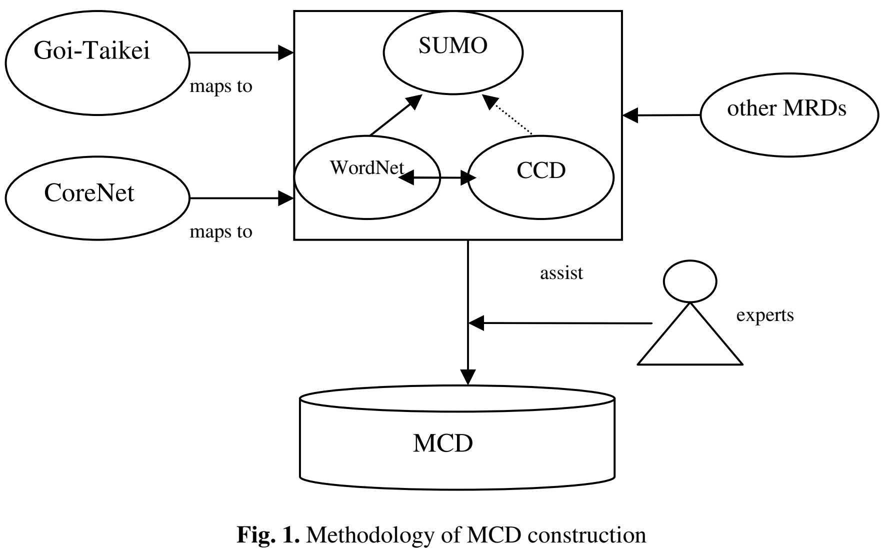


Fig. 1. Methodology of MCD construction

3 A New Knowledge Representation Method

Ontology construction highly relies on knowledge representation methodology. To represent both paradigmatic relation and syntagmatic relation within a dictionary will facilitate most applications. For this reason, to achieve the real natural language understanding, we employ both kinds of relations in MCD.

"Upper Ontology Tree" plus "Lexical Semantics N-tuples" is our new knowledge representation methodology concerning ontology construction which both applies paradigmatic relation and syntagmatic relation on different knowledge functional levels [13]. This methodology divides ontology into two parts, i.e. the higher-level and the lower-level. Two parts have their own orientations in natural language processing. The hyper-level is organized by the hyponymy and the concept taxonomy is just the "Upper Ontology Tree". The lower-level, unlike the higher-level, is designed to represent the syntagmatic information of the concepts which exist in all levels. The syntagmatic relation required here is mainly collocation and co-occurrence, which are extracted from large-scale texts from different areas. This knowledge will be constructed into the "Lexical Semantic N-tuples".

In this new methodology of knowledge representation, the higher-level is beneficial in representing the easy understanding and remembering inheritance hierarchy which is in coarse granularity, while the lower-level can be easily expanded and also provide the possibility of computation by using complicate syntagmatic tuples.

4 Higher-Level Design of MCD

4.1 Coarse Concept Set Prepared for Higher-Level Data

Based on the above exploration of language resource construction, MCD database is designed into higher-level and lower-level parts.

The higher-level part contains general-purpose terms which act as foundation for the lower-level concepts. It not only uses formal language but also is mostly language-independent. The number of concepts is specified between 2,000 and 3,000. The prominent relation is the paradigmatic one, such as "subclass" relation in SUMO and "Hypernym" relation in WordNet.

Unlike the higher-level, the lower-level base has a wide coverage in concepts and reflects the practical use in large-scale running texts. In addition to the paradigmatic relation, syntagmatic relation such as collocation is also included in MCD as important knowledge. In SUMO-WordNet-CCD framework, SUMO provides 1,093 top-level terms in SUO-KIF formal language, which also has OWL formal language for other applications. WordNet has been mapped onto the SUMO so that SUMO terms can be expended to meet the applications in natural language processing such as Information Retrieval.

Currently, we have extracted all concepts associated with hyponymy in SUMO-WordNet-CCD framework, which we called candidate concepts. Among these concepts, a filter algorithm is introduced to generate coarse concept set, which is prepared for our higher-level data. The filter algorithm will be described hereinafter with an example. The current size of the coarse concept set shows in Fig. 2.

The uppermost principle we observe in defining the coarse concept set is that the candidate concepts should be distinguishable for Chinese lemmas, which are presented in CCD as multiple occurrences.

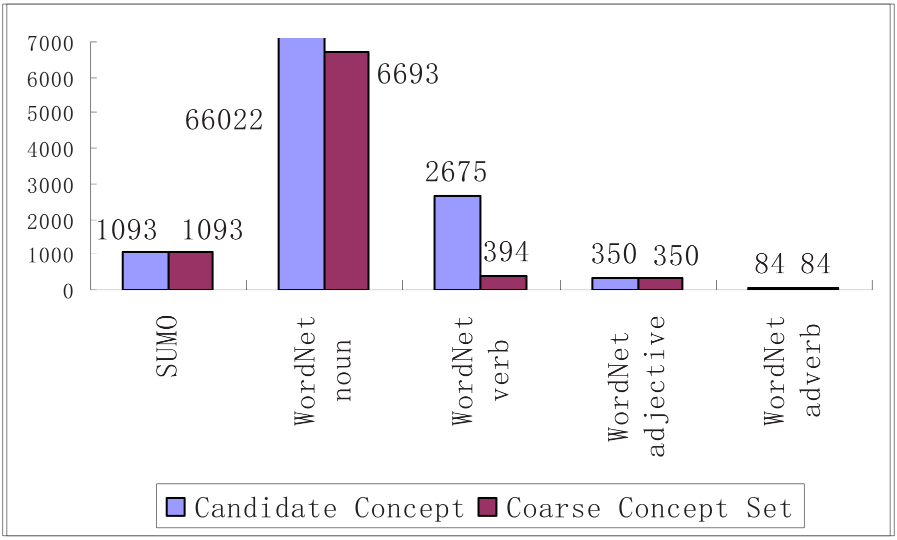


Fig. 2. Current size of candidate concept and coarse concept set

For example, the Chinese word "仪表" (see Fig. 3.) has three occurrences in CCD whose English synsets are (a) {meter}, (b) {appearance} and (c) {appearance visual_aspect}. They all inherit from SUMO concept Entity, but disjoin from SUMO concept Physical and Abstract. Then (b) and (c) are separated by the SUMO concept Attribute and Proposition. In Fig. 3, a SUMO concept is represented in the round angle rectangle, and the CCD concept in the rectangle. The labels, which are placed near the arrowhead line, represent the relation between two concepts they associate. For sense (a), any concept except Entity is distinguishable from (b) and (c), so higher concepts are collected as members of our coarse concept set. With the same principle, sense (b) and (c) can be differentiated by Attribute and Proposition. As a result, the upper concepts comprising Attribute and Proposition are included as members of our coarse set.

4.2 Mapping CoreNet with MCD Frame Base

CoreNet consists of 2,954 hierarchical semantic categories. In addition, more than 50,000 entries are located at the nodes of top-level hierarchy. Since MCD is a collaboration project that ICL keeps with the KAIST, CoreNet is the primary referenced resource for MCD Korean lemmas construction.

A problem arises as to which method should be used to add CoreNet lemmas to MCD, ontologies mapping or merging. Ontologies merging is preferred if these ontologies are mostly consistent. However, this situation is not common since each language resource choose different semantic aspects for categorization, which is hard to translate in other ontologies. After comparing two hierarchies, we propose to design a mapping method between CoreNet and MCD framework.

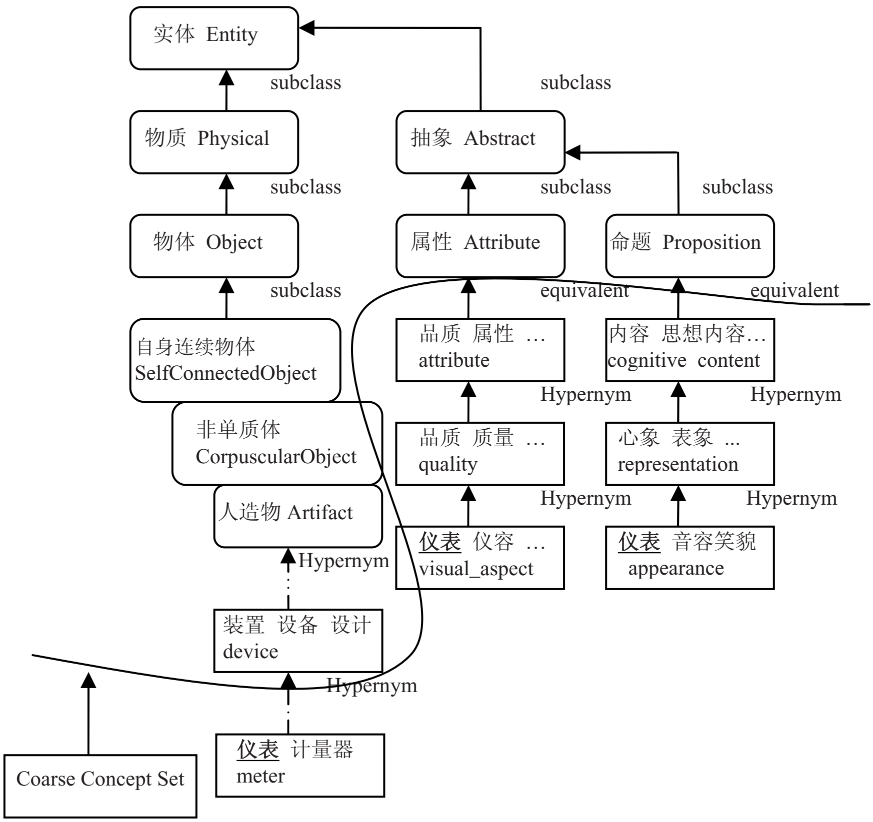


Fig. 3. Generation of coarse concept set

Another problem is concerned with which mapping method is to be applied. Ontology aligning can be done through three kinds of ontology mapping techniques, viz. term-similarity, topology-similarity and instance-similarity. Term-similarity uses similarities of names of terms in ontologies. Bilingual MRD and thesaurus could be used to compute the similarities among terms. Topology-similarity concerns the similarity of hierarchical structures in ontologies. The instance-similarity method compares the overlapping of terms belonging to different categories in ontologies. CoreNet is a Korean-based multilingual language resource, which has Chinese and Japanese lemmas. It contains the same 3,500 terms among coarse concept set's 7,000 terms. So our ontology mapping is carried out in three steps. In the first step, we use CoreNet's Korean-Chinese database to detect the translated terms automatically. We develop editing tools for experts and list all the candidate translation terms in a listbox with their hierarchical information showed alongside. Then Korean expert make the decision that whether the translation terms are appropriate. This semi-automatic method is based on term-similarity method. As a result, in recently, about 4,000 terms are obtained with their Korean equivalents. In the second step, these terms are used as data seeds by applying instance-similarity strategy to return more mapped categories. It is a recursive procedure. We propose to employ a statistical method such as Kappa statistic [14] to achieve our goal. In the last step, the results will be proofread by experts and data evaluation will be carried out.

The manual processing is an important part in the first step of MCD higher-level construction. We have developed an editing tool for experts to operate on. The interface shows in Fig. 4.

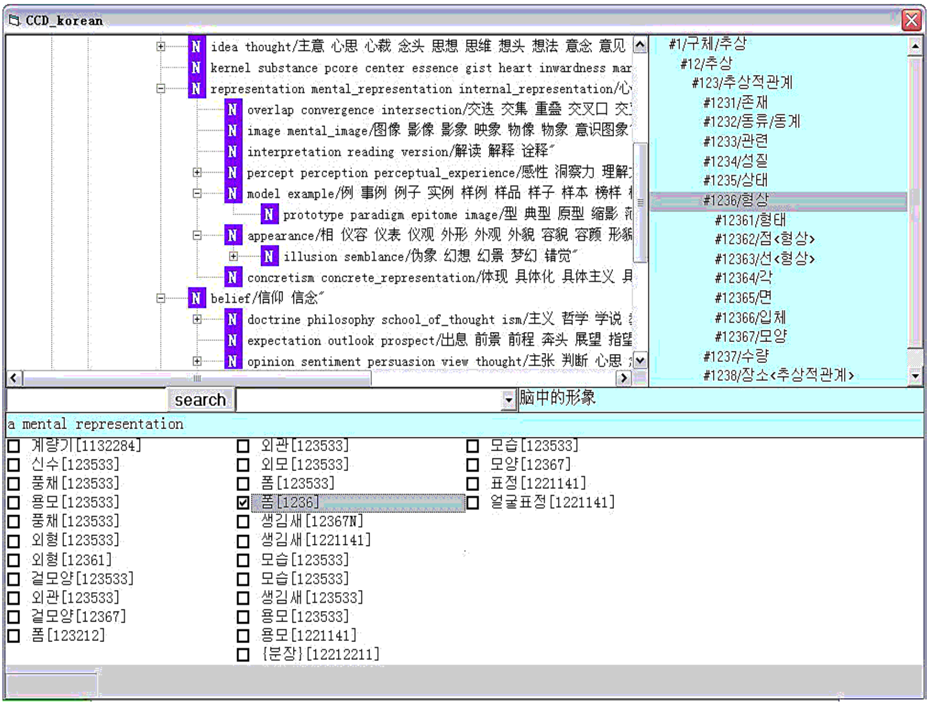


Fig. 4. Editing tool for Korean expert

4.3 Mapping Goi-Taikei with MCD Frame Base

Goi-Taikei consists of 2,710 hierarchical semantic categories. The fact that CoreNet semantic hierarchy originated from Goi-Taikei helps us a lot for Japanese lemmas construction. Since CoreNet semantic categories have pointed onto Goi-Taikei already, MCD concepts, which have Korean equivalents with CoreNet, can be mapped onto Goi-Taikei semantic categories easily. However, we could not totally trust these mapping results. So mapping Goi-Taikei with MCD is carried out in two steps. In the first step, Japanese expert confirm whether the MCD concept can be mapped on the Goi-Taikei semantic category given by Korean expert. If expert find one MCD concept cannot be mapped on the Goi-Taikei's category, he should choose another proper one, and this operation result will feed back to the Korean expert. Then Korean expert will check and modifies his result if it is needed. After the first step, almost 4,000 MCD terms have mapping onto the Goi-Taikei semantic categories. Then in the second step, Japanese experts choose the proper Japanese equivalents for MCD concepts. These procedures can be seen in Fig. 5.

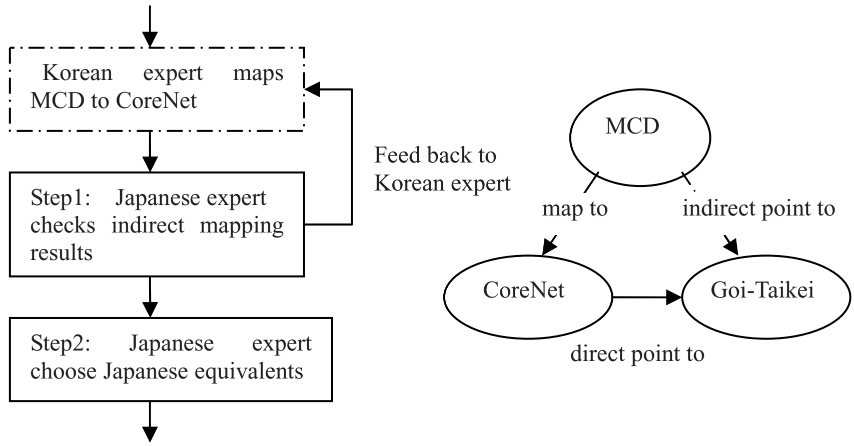


Fig. 5. Procedures of mapping Goi-Taikei with MCD

5 Lower-Level Design of MCD

MCD higher-level base, as we mentioned above, is mainly organized by paradigmatic relation. We call this structure as “Upper Ontology Tree”. With the MCD project ongoing, the research of lower-level representation will be explored simultaneously. The lower-level is designed to represent the syntagmatic information about the concepts which exist at all levels. In addition to the collocation and co-occurrence knowledge extracted from large-scale corpora from different areas, we will also supply other knowledge, such as example sentences from real texts, pointer to the hyper-level node, valence information and so on. These contents are intergraded into the “Lexical Semantic N-tuples”. “Upper Ontology Tree” plus “Lexical Semantic N-tuples” is our new knowledge representation of ontology.

In this new methodology of knowledge representation, the higher-level is beneficial in representing the easily understanding and remembering inheritance hierarchy which is in coarse granularity, while the lower-level is easy to expand and it also provide the possibility of computation among complicate syntagmatic tuples.

6 MCD Database Specification

MCD database now stores in Access format. The database includes 15 fields.

For every record, we mark them with a unique OFFSET and a TYPE, which indicates where it comes from, such as *S* representing that it comes from SUMO. Since MCD is a multilingual dictionary, it contains Chinese, Korean, Japanese lemmas and their descriptive contents. These informations record in CZH, CKR, and CJP fields etc.

The hyponymy can be obtained by the MCDID field. For example, *M001001*'s parent is *M001*. Other relations, such as INSTANCE-OF, PART-OF, ENTAILMENT and CAUSE, are all recorded in MCD database. Besides, we have both pointers with CoreNet and Goi-Taikei. At present, MCD describes 8,614 concepts. Among these, about 5,000 concepts have their Korean equivalents in CoreNet, and around 1,000 concepts have mapped onto Goi-Taikei semantic categories. Part of MCD database can be seen in table 3.

Table 3. Part of MCD database

offset	MCDid	type	czh	cen	cjp	ckr
0001	M001	S	实体	Entity	実体	실체
0002	M001001	S	抽象	Abstract	抽象	추상
0003	M001001001	S	数量	Quantity	数量	수량
0004	M001001001001	S	数	Number	数	수
0005	M001001001001001	S	实数	RealNumber	実数	실수
0006	M001001001001001001	S	有理数	RationalNumber	有理数	유리수
0007	M001001001001001001001	S	整数	Integer	整数	정수
0008	M001001001001001001001001	S	偶整数	EvenInteger	偶数	우정수
0009	M001001001001001001001001001	A	偶数 双数	even	X	짝수
0010	M001001001001001001001002	S	奇整数	OddInteger	奇数	홀수정수
0011	M001001001001001001001002001	A	单数 奇数	odd	X	홀수
0012	M001001001001001001001003	S	质数	PrimeNumber	素数	소수
0013	M001001001001001001001003001	A	质数	prime	X	소수
0014	M001001001001001001001004	S	非负整数	NonnegativeInteger	非負整数	비음정수
0015	M001001001001001001001004001	S	正整数	PositiveInteger	自然数	양의 정수

7 Conclusions

MCD is a Chinese-Korean-Japanese multilingual concept dictionary. It can be used to estimate the semantic similarity between words for the WSD. In computer aided translation, the role of translation memory (TM) and terminology management is very important. MCD could also serve as a complementary method for both TM and terminology management. It can help translators to select target words more accurately at phrase or word level.

This paper is a progress report of MCD. MCD has a lot of challenges that require more consideration in future work. We also plan to perfect the coarse-grained set of the MCD by adding the domain-specific glossary and the collocations, which help text translation in a restricted domain (for instance, translating the Chinese monograph on "computational linguistics" into Korean). With the new knowledge representation, viz. "Upper Ontology Tree" plus "Lexical Semantics N-tuples", the application version of the MCD, built according to the frame of motivated semantic description and the specific domain knowledge will be available for practical use.

References

1. Vossen, P.: Introduction to EuroWordNet. In: Nancy Ide, N., Greenstein, D., Vossen, P. (eds.) Special Issue on EuroWordNet. Computers and the Humanities, 32(2-3), 73-89 (1998)
2. Miller, G.A., Beckwith, R., Fellbaum, C., Gross, D., Miller, K.: WordNet: An On-Line lexical Database. International Journal of Lexicography 3(4), 235–244 (1990)
3. Niles, I., Pease, A.: Towards a Standard Upper Ontology. In: Proceedings of The Second International Conference on Formal Ontology in Information Systems (FOIS-2001), Ogunquit, Maine (2001)
4. Baker, C.F., Fillmore, C.J., Lowe, J.B.: The Berkeley FrameNet project. In: Proceedings of the COLING-ACL, Montreal, Canada (1998)
5. Richardson, S.D., Dolan, W.B., Vandervende, L.: MindNet: acquiring and structuring semantic information from text[A]. In: Proceedings of the COLING- ACL, pp. 1098–1102 (1998)
6. EDR.: EDR Electronic Dictionary Version 1.5 Technical Guide. EDR TR2-007 (1996)
7. Ikehara, S., Miyazaki, M., et al.: Goi-Taikei - A Japanese Lexicon. Iwanami Shoten 1997 (in Japanese)
8. Choi, K.S., Bae, H.S.: Procedures and Problems in Korean-Chinese-Japanese Wordnet with Shared Semantic Hierarchy. In: WordNet Conference, Brno, Czech, pp. 320–325 (2004)
9. Dong, Z.: Knowledge Description: What, How and Who? In: Proceedings of International Symposium on Electronic Dictionary, Tokyo, Japan (1998)
10. Yu, J.S., Yu, S.W., et al.: Introduction to Chinese Concept Dictionary. In: International Conference on Chinese Computing (ICCC2001), pp. 361–367 (2001)
11. Huang, C.R., Chang, R.E, Lee, S.B.: Sinica BOW (Bilingual Ontological Wordnet): Integration of Bilingual WordNet and SUMO. In: Proceedings of the 4th InternationalConference on Language Resources and Evaluation (LREC2004), Lisbon, Portugal (May 26-28, 2004)
12. Niles, I., Pease, A.: Linking lexicons and ontologies: Mapping wordnet to the suggested upper merged ontology. In: Proceedings of the 2003 International Conference on Information and Knowledge Engineering (IKE 03), Las Vegas, Nevada (2003)
13. Liu, Y., Zhu, H.: A Study on NLP-Oriented Lexical Knowledge Representation. In: CLSW’06, Hsinchu (in Chinese)
14. Fleiss, J.L.: Statistical Methods for Rates and Proportions. John Wiley & Sons, West Sussex, England (1973)

Designing an Aural Comprehension Aid for Interlingual Communication

Hideyuki Nakanishi

Department of Adaptive Machine Systems, Osaka University
2-1 Yamadaoka, Suita, Osaka 565-0871, Japan
nakanishi@ams.eng.osaka-u.ac.jp
http://www.media.ams.eng.osaka-u.ac.jp/

Abstract. This study presents an aural comprehension aid to help Japanese travelers hear a counter clerk's questions at fast food restaurants in the US. The prototype of the aid employed a speech recognition method in which a user assists the speech recognizer of the mobile device. The user presses the device's button as promptly as possible when missed words were spoken so that the recognizer perceives the moment, which is utilized for improving recognition accuracy. More than a hundred dialogs between a Japanese traveler and fast-food clerks were recorded and used to evaluate the prototype. The evaluation showed that the proposed method could improve recognition accuracy, though the improvement was not sufficient for practical use.

Keywords: Speech recognition, aural comprehension, mobile device, interlingual communication.

1 Introduction

This study is about a speech recognition technique for mobile devices that help interlingual communication. Since speech-to-speech translation is extremely difficult [9], many researchers have tried to devise simplified ways instead of using a straightforward combination of speech recognition and machine translation technologies. The most typical approach is sentence-based translation based on keyword spotting [6]. In this approach, a speech recognizer retrieves some keywords from the utterance of the person the user is talking with, and then a search engine uses those keywords to search for any of a prepared set of sentences in order to detect the sentence that would have actually been spoken. Several extensions to this method have been proposed. For example, there is a study that used a common-sense database for improving the process of finding the sentence that matches the retrieved keywords [1]. The mobile device developed in another study displays the selected sentence and also a menu of replies to it, which makes conversational exchanges smoother [4]. A commercial product employed an embodied conversational agent that plays a role of in-between and eases the awkwardness caused by recognition and translation errors [7]. Despite these efforts, sentence-based speech translation does not work well, because even

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 170–181, 2007.

keyword spotting is not easy in our noisy living environments [2]. A more robust speech recognition method is necessary to develop mobile speech recognition devices for supporting interlingual communication.

This paper proposes a speech recognition method named "user-assisted speech recognition," which is functionally very limited but much more robust than conventional methods. In this method a user presses the device's button as promptly as possible when words are spoken. Then, the speech recognizer of the device searches the last part of continuously recorded audio for any of a prepared set of phrases (or short sentences). Finally, the recognizer obtains an interval between the moment when the button is pressed and the ending time of each spotted phrase. The possibility that a phrase has actually been spoken is high when its interval is short. In this way, users can assist the speech recognition process, since humans can recognize when an utterance ends, even if they cannot understand the actual utterance because of a language barrier.

The method was implemented and tested as an aural comprehension aid to help Japanese people understand idiomatic phrases spoken in English. The aid is a weaker but more feasible solution than speech-to-speech or sentence-based translation. Speech-to-speech translation basically does not restrict input and output sentences. Sentence-based translation restricts output sentences but can accept a wide variety of input sentences. The aural comprehension aid can accept only a prepared set of input sentences. Even such a limited function is useful for Japanese people who have difficulty in hearing English and who are also not familiar with English idioms. To learn the unfamiliar pronunciation of unfamiliar words in a foreign language, humans have to make the effort to train themselves for it. Speech recognition software that has already been trained for the language can help them to hear until they master the language.

2 Application

The supposed users of the aural comprehension aid are Japanese travelers who try to hear a fast-food clerk's questions in the US, since that is a typical situation where the aid is vital as exemplified below.

Recently, a large number of people write Web-based diaries, some of which present their experiences of traveling abroad, since that is an interesting topic. Since especially funny problems during their travels become attractive content, I was able to find about forty diaries that describe awkward experiences that happened while they were ordering fast food in the US. According to those diaries, fast-food restaurants appear to be a prominent place where Japanese travelers face difficulties in communication. One of the reasons for this is that fast-food clerks are not used to talking with nonnative speakers, whereas hotel clerks are. Another reason is that people have to order orally and quickly in fast-food restaurants, while Japanese travelers can order by reading a menu book and pointing at a menu item with plenty of time in normal restaurants.

Ten of the forty diaries report that the writers of those diaries could not understand the question, “For here or to go?” Three of them answered “Yes, please,” to the question and confused the counter clerk. Since eight of the ten writers did not know this idiom, this problem seemed to be caused by a lack of knowledge rather than a lack of listening ability. Electronic dictionaries can compensate for such a lack of knowledge, and an aural comprehension aid is not necessary to solve this problem. The other two writers, however, could not catch the question, even though they knew the idiom. Thus, a lack of listening ability really counts against the understanding of the question.

Three of the forty diaries report that the writers of those diaries could not hear the question, “Ketchup?” These reports clearly indicate that listening ability is essential to successful understanding, since almost all Japanese people know the word “ketchup.” Two of the three diaries transcribed how the writers heard the word. One writer heard it as “Catch a?” The other writer thought that he was being asked, “Catch up?” Why did they hear the simple word “ketchup” incorrectly? A plausible reason is that fast-food clerks speak too quickly. To understand such quick speech a slow-playback function is helpful and more reliable than an aural comprehension aid. But, the above two transcriptions imply that the difference between Japanese and English pronunciation should be the cause. The two vowels of “ketchup” are different from those of “catch up,” but it is not easy for Japanese people to notice this difference, since vowels in the Japanese language consist of only: “a,” “i,” “u,” “e,” and “o.”

Many other mistakes that can be considered to originate in the difference between Japanese and English pronunciation have been reported. For example, a writer was very shocked when he realized that the question that he heard as “South?” was actually “Sauce?” This mistake happened, probably because of the “th” sound, that is peculiar to English. Another writer mistook “Chips or drink?” for “Chicken?” Since no word in Japanese ends with a final consonant but all words end with a vowel, the writer might not be able to catch the “ps” and “k” sounds. Moreover, there is no concatenated consonants like “dr” in Japanese. Two funny examples transcribed in the diaries are given below.

Example 1. The word “bread” was heard as “beverage.”

Writer:	I’ll have a chicken sandwich.
Clerk:	What kind of bread [→ beverage] do you want?
Writer:	Nothing, thank you.
Clerk:	... What kind of bread? (in a loud voice)

It is not easy for Japanese people to distinguish the “b” sound from the “v” sound. And Japanese people are usually insensitive to the accents of words. These shortages of listening ability might cause the problem in the above example.

Example 2. "Is that all?" was heard as "Zero?"

Writer:	This one.
Clerk:	OK. Is that all? [→ Zero?]
Writer:	One.
Clerk:	One more? Two?
Writer:	No. Only one.
Clerk:	Is that all? [→ Zero?]
Writer:	... One. (in a tearful voice)

It is conceivable that the writer of this example could not catch the "th" sound just after the "z" sound of "Is", and the final "t" sound of "that". Consequently, the writer might hear "Is that" as the first syllable of "zero", since there is no "ae" vowel in Japanese. Additionally, the writer seemed to believe that "all" was the second syllable of "zero", since the Japanese people make no distinction between "l" and "r". Probably, these transformations gave birth to the mistake.

The survey described above indicated that the pronunciation difference was one of the major impediments to understanding what a fast-food clerk said. An aural comprehension aid is useful for coping with pronunciation differences, while the other two impediments, which are a lack of idiomatic knowledge and quick speech, can be eliminated by other functions, e.g., electronic dictionaries and a slow-playback function.

3 Aural Comprehension Aid

3.1 User-Assisted Speech Recognition

In contrast to conventional speech recognition methods in which the speech recognizer tries to detect speech without the user's intervention, in the user-assisted speech recognition method the speech recognizer cooperates with users as follows: The recognizer embedded in a mobile device is used in a situation in which both the recognizer and its user simultaneously hear the speech of a speaker. In a noisy place it is difficult for the recognizer to know when the speech ends, whereas the user often knows it precisely, even though he or she cannot understand the speech. When the user cannot catch someone's words, he or she presses the button of the device as promptly as possible in order to inform the recognizer of the exact time when the missed words were spoken. When the button is pressed, the recognizer searches the last part of continuously recorded audio for any of a set of idiomatic phrases that were prepared beforehand. Each found phrase is accompanied with its confidence score, which represents the probability that the phrase was actually spoken, and the phrase's ending time, which represents when a speaker is supposed to finish speaking the phrase. Then, the recognizer calculates an interval between the moment when the button was pressed and the ending time of each phrase. The larger the interval is, the lower the score is. An example of the formula to reduce scores is:

$$Score = ConfidenceScore - Coefficient * Interval \qquad (1)$$

Finally, the device suggests several phrases in the order of their scores, and then the user guesses which one was actually spoken based on the context of the conversation and the place where the conversation occurred.

When the coefficient is higher, the interval affects the score more seriously. Namely, the accuracy of scoring becomes more improved if a user presses the button with precise timing, but becomes worse if a user's press lags behind the actual moment. Furthermore, in that case, the speech recognizer may spot extra phrases that incorrectly match noises that filled in the last space of the device's audio buffer. Since it is almost impossible for a user to eliminate such a time lag perfectly, a fixed period of time that offsets the time lag should be introduced as follows:

$$Score = ConfidenceScore - Coefficient * (Interval - Offset) \qquad (2)$$

The offset should be adjusted for each user, since the length of time it takes to respond to a missed word varies from person to person. If the offset is too short, the problem described above occurs. If the offset is too long, a speech recognizer may lose a correct phrase that could be spotted if the offset were shorter than the actual time lag.

What occurs if conventional speech recognition methods, which rely only on confidence scores and do not cooperate with users, are used for an aural comprehension aid? The device may continue to suggest incorrect phrases one after another, since the last part of the recorded audio does not necessarily correspond with the unheard phrase and is usually noise. To mitigate the influence of noise, the speech recognizer should suggest only phrases for which the confidence scores are above a certain threshold. In this case a highly scored phrase can continue being suggested until a higher phrase appears. So, a maximum duration in which a phrase can continue to be suggested is necessary to delete a phrase that is highly scored but obsolete. In the manner described above, conventional speech recognition methods may be useful for an aural comprehension aid, but the appropriate values of the threshold and the maximum duration are likely to vary according to the situation, e.g., the place and the conversation. User-assisted speech recognition seems to be a better way to deal with noise, since the coefficient is relatively independent from the situation, and adjusting one value is a simpler task than adjusting two values.

The aural comprehension aid device has to record audio continuously and keep it in the audio buffer, since it is impossible for the device to predict when a user will press its button. Continuous audio recording in everyday life may invade privacy [3]. The device, however, requires only several seconds of audio, which is sufficient to cover a single phrase that has just been spoken. Since recorded audio can be discarded within a very short time, an aural comprehension aid (and user-assisted speech recognition) is fairly free from privacy infringement problems.

1. What kind of bread?	$50 - 25 * 0.4 = 40$
2. Chips or drink?	$35 - 25 * 0 = 35$
3. What kind of drink?	$60 - 25 * 1.2 = 30$
4. What size?	$45 - 25 * 0.8 = 25$

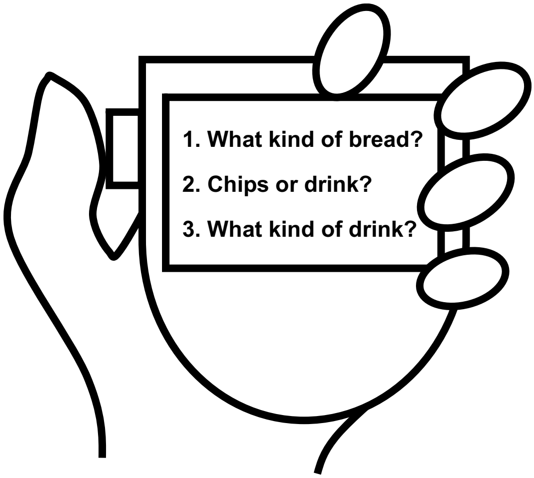


Fig. 1. A result from the proposed method

3.2 Usage

Suppose that a user has just ordered a sandwich at a submarine sandwich restaurant. The user is asked, "What kind of bread?" but cannot hear it, so he or she presses the button of the device immediately. Then, the device's LCD screen displays the result of recognition as Figure 1, in which the coefficient is twenty five and the offset is zero. Since the maximum confidence score is logically one hundred, the score of any spotted phrase becomes zero if its interval is longer than four seconds. In the figure, the four phrases matched the last part of the audio data stored in the device's audio buffer when the user pressed the button. According to the confidence scores the phrase, "What kind of drink?" is the best match, but it becomes ranked as the third after its confidence score, which is sixty, is changed to thirty, because the interval is 1.2 seconds and thus thirty is subtracted from its confidence score. The phrase, "Chips or drink?" is ranked as second in spite of the worst confidence score, which is thirty five, since its interval is zero. The phrase, "What size?" is not presented, because the device depicted in the figure displays only three phrases. Eventually, "What kind of bread?" is suggested as the most probable phrase and the user answers "Wheat bread, please."

In the above example the user picks up the first phrase, "What kind of bread?", since the situation is that the user has just ordered a sandwich. If the situation were such that the user had ordered all of the ingredients of the sandwich and was about to receive it, the user would guess that the second

1. What kind of drink?	60 (– 12.5 * 1.2 = 45)
2. What kind of bread?	50 (– 12.5 * 0.4 = 45)
3. What size?	45 (– 12.5 * 0.8 = 35)
4. Chips or drink?	35 (– 12.5 * 0 = 35)

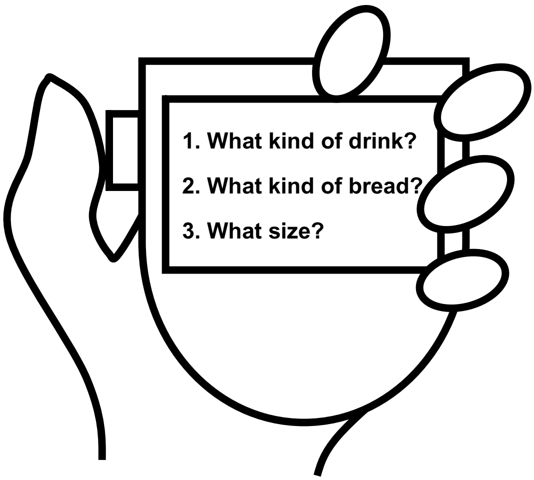


Fig. 2. A result from the conventional method

phrase, "Chips or drink?" had actually been spoken. If the user had ordered a drink, the third phrase, "What kind of drink?" would obviously have been the most probable.

What happens if the device uses only confidence scores? As Figure 2 shows, "What kind of drink?" becomes the first, but the user knows that it is incorrect, because the user did not order a drink. The user, however, is confused by the third phrase, "What size?", since a counter clerk may ask that question as well as the second phrase just after the user says the kind of sandwich. That question would be out of the ranking when the intervals were taken into account. The user is unable to decide between responding, "Wheat bread, please," and responding, "Six-inch, please." Even if the scores reflect the intervals, the same confusion occurs when the coefficient is lower than 12.5, as shown in the figure.

In these two examples it is assumed that the device displays only English text of the phrases. This function, which merely compensates for a lack of listening ability, can easily be extended to compensating for a lack of knowledge. The device should display the Japanese translation of phrases if a user does not have enough knowledge of idioms to understand them. And the device becomes more helpful if it displays presumable answers to such a typical question as "Chips or drink?" Furthermore, the device can be equipped or connected with a database that contains the menus of fast-food chains to display answers to such a question as "What kind of drink?"

It is also assumed that the results of recognition are presented via the device's screen. The results, however, can be presented through other modality such as pictorial icons and text-to-speech. Pictorial icons seem easier for a user to perceive than small text displayed on the device's screen, but the user is likely to misinterpret the displayed icon if many kinds of icons are used, since it is difficult to memorize all of them. Text-to-speech releases a user from looking at the device's screen, but it takes a long time for the user to listen to all three phrases suggested by the device.

3.3 Implementation

In general, speech recognition engines convert speech into the most probable text. This kind of a speech-to-text engine cannot be used to implement user-assisted speech recognition, since it does not provide multiple possible texts that match the same audio segment. User-assisted speech recognition requires a speech-to-phrases engine that searches the same audio segment for multiple phrases. The second problem of general speech recognition engines is that they usually use a language model and corpus to estimate the confidence score of each word. This kind of model-based engine makes the results of recognition vulnerable to informal speech and the noise of living environments. Model-free engines seem more suitable for an aural comprehension aid. The third problem is that some of general engines are speaker-dependent and needs user enrollment. An aural comprehension aid obviously needs a speaker-independent engine.

A speech-to-phrases, model-free speaker-independent engine (*Nexidia* [5]) was incorporated into the prototype of the aural comprehension aid system. Nexidia is a phonetic search engine that uses only phonetic features of a word or phrase to find it in an audio file. The prototype was not implemented on a mobile device but simulated by GUI-based software running on a desktop or notebook computer, since that was sufficient for the evaluation described later. In the test, the user clicked a button on a screen instead of pressing a physical button, and the prototype searched through the audio files stored in the same computer instead of actual surrounding audio. However, the results of the test were not artificial, since the audio files were actual recorded dialogs between a Japanese visitor and a fast-food clerk. The procedure of the test was such that the audio file of a dialog was played, the user clicked a graphical button, and then the prototype searched the file. This simulated environment could produce credible recognition results.

The mobile speech translation devices developed in the two past studies [1,4] presented several recognition results and asked the user to choose the most relevant one, since it was supposed that speech-to-text did not work perfectly. In these studies, however, it was assumed that keyword spotting worked well. The aural comprehension aid system that was developed without such an assumption relies on keyphrase matching, which is more robust than keyword spotting.

The two past studies of mobile speech recognition systems used conventional speech recognition methods and a speech-to-text model-based engine (*IBM ViaVoice*) [1,8]. As a result, in the experiments, the systems developed in those studies dealt with only formal speech uttered in quiet environments. This study used the user-assisted speech recognition method and a speech-to-phrases model-free engine (Nexidia) to develop a prototype of the aural comprehension aid system. The results of the evaluation described later show that the prototype was better than conventional systems for dealing with informal speech in noisy environments.

4 Preliminary Evaluation

A native Japanese person who was staying in Atlanta, Georgia, USA, temporarily from April 2005 to March 2006 recorded more than a hundred dialogs between himself and counter clerks at fast-food restaurants of more than sixty different chains. The kinds of food served at the restaurants that he visited were sandwiches, hamburgers, fried chicken, pizza, burritos, and so on. To record the dialogs, he walked into the restaurants with a digital audio recorder in his breast pocket. He deleted each audio file except the part that contains the dialog shortly after recording it in order to avoid unnecessarily invading anyone's privacy.

Table 1 is a list of one hundred and one phrases that appeared in the collected dialogs. These phrases were used as keyphrases embedded in the prototype of the aural comprehension aid system. The keyphrases include greetings as well as questions, since a user may be unable to understand greetings and may misinterpret a greeting as a question. The keyphrases cover various kinds of fast-food restaurants. If the keyphrases dynamically change in accordance with the kind of a restaurant where the user is ordering, the accuracy of recognition can be improved. For example, it is possible to exclude keyphrases that can be spoken only at sandwich, fried chicken, pizza, or burrito restaurants if a user informs the aural comprehension aid device that the user is now ordering at a hamburger restaurant. If the user specifies the kind of restaurant chain, the accuracy may be further improved. If the device is location-aware, these adaptations can be automated.

4.1 Limitations of the Proposed Method

It was difficult for the prototype to recognize phrases that have only one or two syllables, since a phonetic search engine uses only the phonetic features of a phrase to locate it in an audio file. The longer a phrase is, the more easily a phonetic search engine can find it. Long phrases are generally, however, multifarious and exceptional, thus it is difficult to include all of their variations as keyphrases. For example, it is hard to include phrases in which the last few words vary. All of their variations must be included because it becomes impossible to calculate an interval if a partial phrase not having the last few words is used to find the phrase. The most varying phrases are those that end with a number (e.g. "That'll be seven sixty-eight" and "Your number is sixteen"). A phrase that begins with a number (e.g. "Seventy-five-cent charge to use your card.") is not problematic, since the non-numerical part that follows the number can be

Table 1. Keyphrases

Anything else?	Just one sandwich?	Tomatoes?
Anything else for you?	Just sandwich?	Two side orders?
Anything to drink?	Ketchup?	Two sides?
Banana peppers?	Lettuce?	Vinegar?
Bread bowl?	May I help you?	What kind of mustard?
Cheese on it?	Mayonnaise?	What can I get for you?
Chicken or Steak?	Medium or large?	What else?
Chips or drink?	Medium size?	What kind of bread?
Debit or credit?	More cocktail?	What kind of bread do you want?
Do you need plates or forks?	More ketchup?	What kind of bread would you like?
Do you wanna combo?	Mushrooms?	What kind of cheese?
Drink?	Mustard?	What kind of dressing would you like?
Enjoy your meal.	Onions?	What kind of drink?
Everything?	Oregano?	What kind of fries?
Everything on it?	Original or crispy?	What kind of salad?
For here?	Pickle?	What kind of sandwich?
For here or to go?	Picture ID?	What kind of sauce?
Good bye.	Ready?	What kind of sauce do you wanna?
Good night.	Salt pepper?	What kind of soup do you want?
Gravy?	Sauce?	What sides?
Have a good day.	Sign it.	What size?
Have a good night.	Sign it for me.	What size do you wanna?
Have a nice day.	Sourdough bread?	What size is your drink?
Hello.	Spicy or mild?	What size of fries?
Here we go.	Take your receipt.	What would you like to drink?
Here you are.	Thank you.	What's your name?
Hot sauce?	That'll be all?	Which size do you wanna?
How are you?	That's all?	White or wheat?
How are you doing?	That's it?	White or wheat bread?
How can I help you?	That's it for you?	Whole or half?
How's it going?	This is your receipt.	With cheese?
Is that all?	To go?	You need ketchup?
Is that for here or to go?	To go or for here?	You want something to drink or dessert?
Jalapeno?	Tomato sauce?	

used to find such a phrase. Interrogative sentences usually have some variations. For example, "What kind of bread?" can be followed by "do you want?", "do you like?", or "would you like?" This kind of variation is not too numerous to be included as keyphrases.

4.2 Comparison with the Conventional Method

Since it was found that even the proposed method could not suggest a correct phrase if the recorded audio was very noisy, relatively less noisy audio data were chosen to compare the proposed and the conventional methods. Table 2 shows how the two methods suggested the eight phrases. In this table the phrase shown in boldface is the correct phrase. The left half of the table presents the results obtained by the proposed method. The coefficient was twenty-five and the offset was zero. The right half of the table shows the results produced by a conventional

Table 2. Comparing the methods

Proposed method		Conventional method	
1. **Anything else?**	46	1. **Anything else?**	51
2. What size do you wanna?	44	2. That's it?	49
3. Onions?	40	3. How's it going?	49
1. **What size?**	61	1. **What size?**	66
2. Sign it.	61	2. Sign it.	61
3. What sides?	53	3. What sides?	58
1. **Chips or drink?**	48	1. Is that all?	53
2. What size is your drink?	45	2. To go?	50
3. That's it?	44	3. Have a good night.	51
1. **What kind of bread?**	46	1. Hello.	49
2. What kind of drink?	45	2. What else?	48
3. How are you doing?	45	3. **What kind of bread?**	46
1. What kind of bread do you want?	49	1. That's it?	53
2. **For here or to go?**	49	2. Sign it.	53
3. Sign it.	48	3. What size do you wanna?	52
1. **Is that all?**	50	1. **Is that all?**	60
2. That's it?	45	2. Here you are?	50
3. That's all?	45	3. Hello.	50
1. **White or wheat?**	51	1. **White or wheat?**	56
2. Onions?	48	2. Onions?	53
3. That's it?	44	3. What kind of soup do you want?	52
1. Is that all?	55	1. Is that all?	55
2. That's it?	48	2. That's it?	48
3. **That'll be all?**	46	3. Good night.	47

method which sorted the phrases that appeared in the recorded audio by just their confidence scores. The differences between the left and the right parts of the table indicate how the results were improved by the proposed method. You can see that the results of "Chips or drink?", "For here or to go?", and "That'll be all?" obtained by the conventional method were completely incorrect.

5 Conclusion

This paper showed that simple user intervention could enable mobile devices to recognize short speech in noisy living environments. The user-assisted recognition method proposed in this paper was implemented and evaluated as an aural comprehension aid to help Japanese travelers hear a counter clerk's question at fast food restaurants in the US. The proposed method was able to improve recognition accuracy, though the improvement was not sufficient for practical use.

Acknowledgements. I would like to express my sincere gratitude to Gregory Abowd for providing me an opportunity to stay at Georgia Tech and to conduct this work. I also express my thanks to Julie Kientz for helping me to use the phonetic search engine *Nexidia*.

References

1. Chung, J., Kern, R., Lieberman, H.: Topic Spotting Common Sense Translation Assistant. CHI Extended Abstracts 2005, 1280–1283 (2005)
2. Dobler, S.: Speech Recognition Technology for Mobile Phones. Ericsson Review 3, 148–155 (2000)
3. Iachello, G., Abowd, G.D.: Privacy and Proportionality: Adapting Legal Evaluation Techniques to Inform Design in Ubiquitous Computing. In: Proc. CHI 2005, pp. 91–100 (2005)
4. Imoto, K., Sasajima, M., Shimomori, T., Yamanaka, N., Yajima, M., Masai, Y.: A Multi Modal Supporting Tool for Multi Lingual Communication by Inducing Partner's Reply. In: Proc. IUI 2006, pp. 330–332 (2006)
5. Nexidia: http://www.nexidia.com/
6. Phraselator: http://www.voxtec.com/
7. TalkMan: http://www.jp.playstation.com/scej/title/talkman/
8. Vemuri, S., Schmandt, C., Bender, W., Tellex, S., Lassey, B.: An Audio-Based Personal Memory Aid. In: Davies, N., Mynatt, E.D., Siio, I. (eds.) UbiComp 2004. LNCS, vol. 3205, pp. 400–417. Springer, Heidelberg (2004)
9. Wahlster, W.: Verbmobil: Foundations of Speech-to-Speech Translation. Springer, Heidelberg (2000)

Parallel-Text Based Support System for Intercultural Communication at Medical Receptions

Mai Miyabe[1], Kunikazu Fujii[1],
Tomohiro Shigenobu[2], and Takashi Yoshino[2,3]

[1] Graduate School of Systems Engineering, Wakayama University,
930 Sakaedani, Wakayama, Japan
{s085051,s075065}@sys.wakayama-u.ac.jp
[2] Language Grid Project,
National Institute of Information and Communications Technology,
3-5 Hikaridai, Seika-cho, Soraku-gun, Kyoto, Japan
shigenobu@nict.go.jp
[3] Faculty of Systems Engineering, Wakayama University,
930 Sakaedani, Wakayama, Japan
yoshino@sys.wakayama-u.ac.jp
http://www.wakayama-u.ac.jp/~yoshino/

Abstract. Differences in languages have to be bridged in order for intercultural communication to take place. Although the medical field requires highly accurate translations for promoting intercultural communication, the present level of support is insufficient. In this paper, we propose a system that uses parallel texts to support intercultural communication at hospital reception desks. Using parallel texts that have been translated accurately, the proposed system can provide this level of intercultural medical communication. We conducted an experiment to test the effectiveness of the system, which confirmed that it is possible to provide accurate intercultural communication using parallel texts. The experiment compared our system with the conventional system, which uses a printed list of parallel texts, and showed that the retrieval time of the proposed system is about the same as that of the conventional system.

Keywords: parallel text, intercultural communication, medical reception.

1 Introduction

At the end of 2005, foreigners in Japan accounted for 1.57% of the total population. This percentage has being increasing every year, so opportunities for intercultural communication will increase. Because people speak different languages, we need special support to facilitate intercultural communication, particularly machine translation.

Despite recent advances in machine translation technology, it is still very difficult to obtain highly accurate translations. Inaccurate translation adversely affects communication, and an incorrect machine translation can cause serious problems.

In the medical field, in particular, accurate translations are very important. Medical care directly impacts both life and health. When a foreigner who cannot speak Japanese

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 182–192, 2007.

receives medical care, poor communication between doctor and patient may lead to errors in treatment. The medical field requires accurate communication between people who speak different languages. Therefore, medical translators often need to accompany foreigners seeking medical care, but there are few medical translators. Because of this shortage of medical translators, parallel texts are often used to aid in communication.

Parallel texts are lines of text in one language paired with translations of the text in other languages. In other words, parallel texts are accurate translations prepared in advance that are meant to improve the efficiency and accuracy of medical treatment[1]. Face-to-face communication systems using parallel texts are now in use. One of them is a support system using speech-to-speech translation for foreign travelers[2]. Another topic of research is a tool that supports communication between speakers of different languages that uses parallel texts for speech recognition[3]. In these systems, the user inputs speech, and the system outputs a translated sentence, but the system cannot output sentences that have not been previously registered. Speech translation systems using phrase translation for communication in the medical field have also been proposed[4,5]. Even these systems, however, provide insufficient support for the medical field. We need a common system that can register, share, and use parallel texts among several hospitals.

In this paper, we propose a parallel-text based support system for intercultural communication at hospital reception desks. This system provides reliable communication using correct translations. We describe the proposed system, our experiment with intercultural conversation at a hospital reception desk, and an experiment comparing our system with the conventional method.

2 Support System for Intercultural Communication at a Hospital Reception Desk

2.1 System Configuration

(1) Retrieval and registration of parallel text using a web service
Groups and communities working together can easily create a huge body of parallel texts. Conversation support systems use web services to reserve and share parallel texts by retrieving and registering them. This web service is a software system designed to support interoperable machine-to-machine interaction over a network. It is defined by the W3C. Web services can cooperate with one another and are scalable. They can also use parallel texts easily. Figure 1 shows the configuration of the proposed system. We use two types of web service: an example service and a parallel text service. The example service provides Japanese questions that are used by hospital staff. The parallel text service provides parallel texts of Japanese questions and candidate responses to the questions. These services are divided because they are easy to combine with a parallel text collecting system. Using these services, we can obtain and share the necessary parallel texts. If the necessary parallel text is not found, hospital staff can ask translators to register a new parallel text. The structure of this system enables it to store huge amounts of parallel texts and share and update data among hospitals.

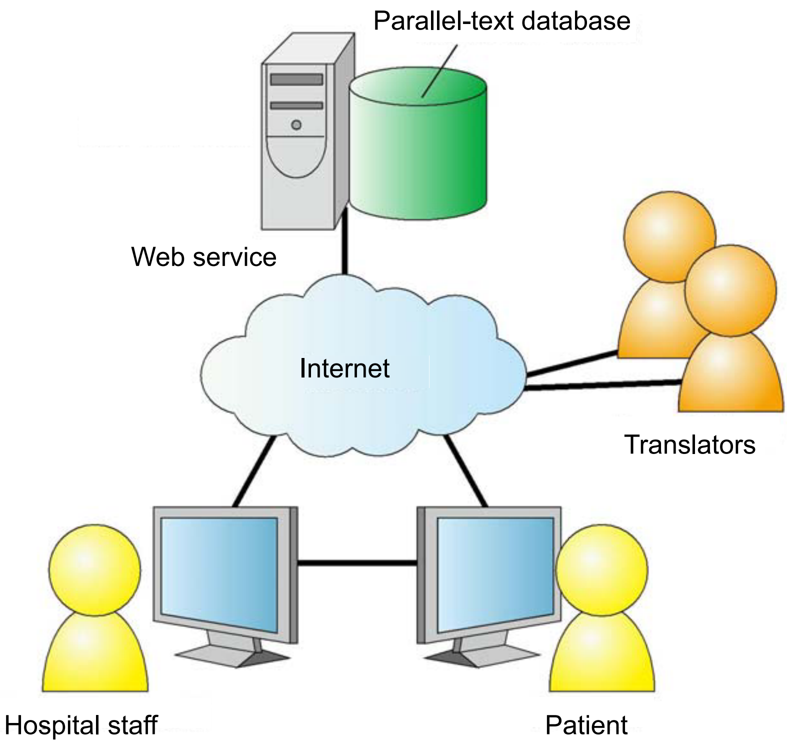


Fig. 1. Configuration of proposed system

(2) User-friendly interface
Hospital reception staff and patients use these systems differently. Hospital staff use the system regularly. Patients, on the other hand, use the system less frequently, meaning they are less expert. Therefore, the system provides two types of system interface: a simple one for patients and a multifunctional one for hospital staff. Use of a touch screen instead of a keyboard makes the system easier to use.

2.2 Staff Side of System

Figure 2 is a screenshot of the staff side of the system. The users of the system are Japanese. Hospital staff communicate with patients using this system. A hospital staff questions a patient and he or she responds. When the hospital staff selects a "scene selection button" on the screen, the system retrieves parallel texts for a given hospital reception scenario from the web service. The parallel texts obtained are shown as a list of questions on the screen. When a user selects and sends a parallel text from the list of questions, the selected question and the candidate responses that are translated into the patient's language are displayed on the patient side of the system.

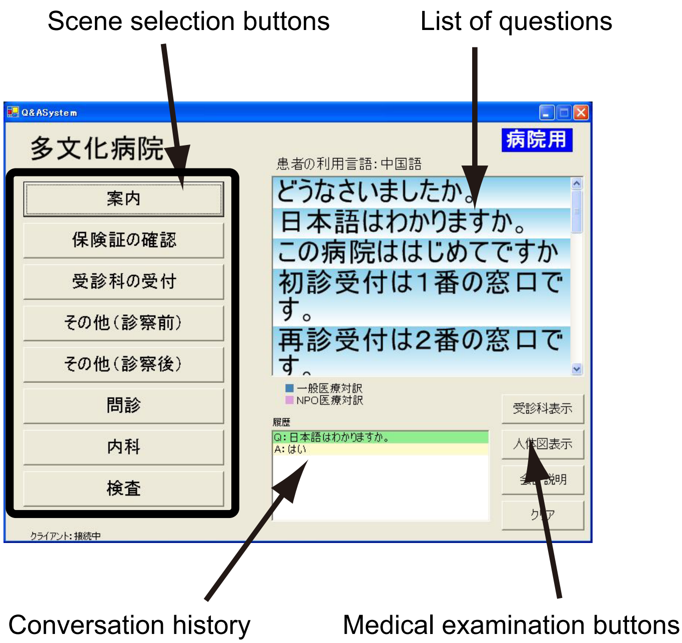


Fig. 2. Screenshot of staff side of system

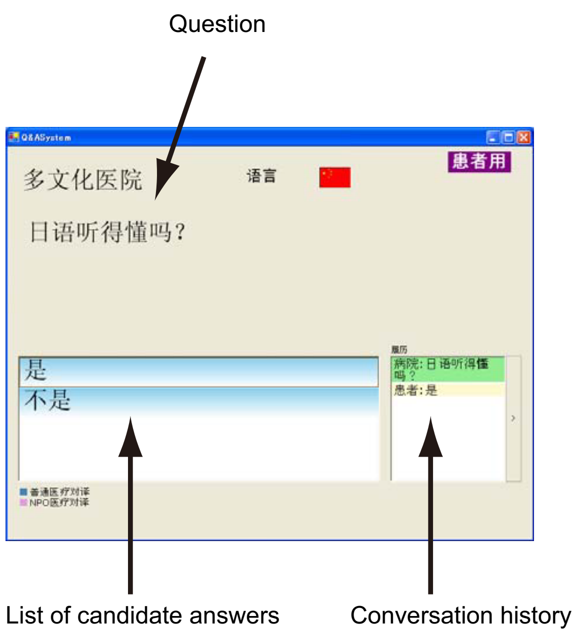


Fig. 3. Screenshot of a patient side of system for Chinese patients

2.3 Patient Side of System

Figure 3 shows a screenshot of the patient side of the system. The users of this version of the system are Chinese. The parallel text that is sent by the staff side of the system is labeled as the question and the candidate responses are as shown in the list of candidate answers. When a user selects and sends a response, the selected response is translated into Japanese and is displayed on the staff side of the system. When a hospital staff clicks the medical examination button on the staff side of the system, the medical examination screen is displayed on the patient side of the system. In the screen, a patient can answer questions about his or her symptoms by using a chart of the human body. Figure 4 shows a screenshot of the medical examination. In this screen, the candidate symptoms are shown in the list of symptoms when the patient clicks the region exhibiting a symptom in the chart of human body. When a patient selects a symptom, the selected region and the symptom that is translated into Japanese is displayed on the staff side of the system.

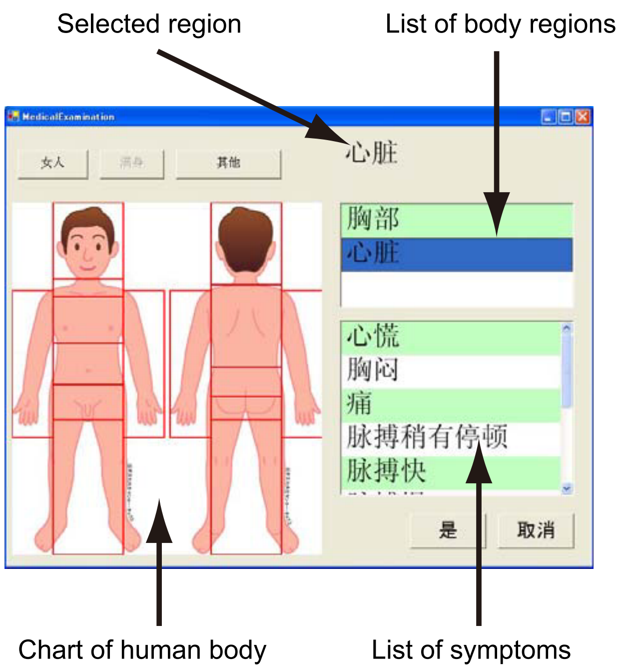


Fig. 4. Screenshot for Chinese patients

3 Experiment on Intercultural Conversation at a Hospital Reception Desk

3.1 Experimental Outline

We experimented with intercultural conversation between Japanese and Chinese individuals at a hospital reception desk to evaluate how well our system supported intercultural communication. There were four subjects: medical translators and a nurse. In this

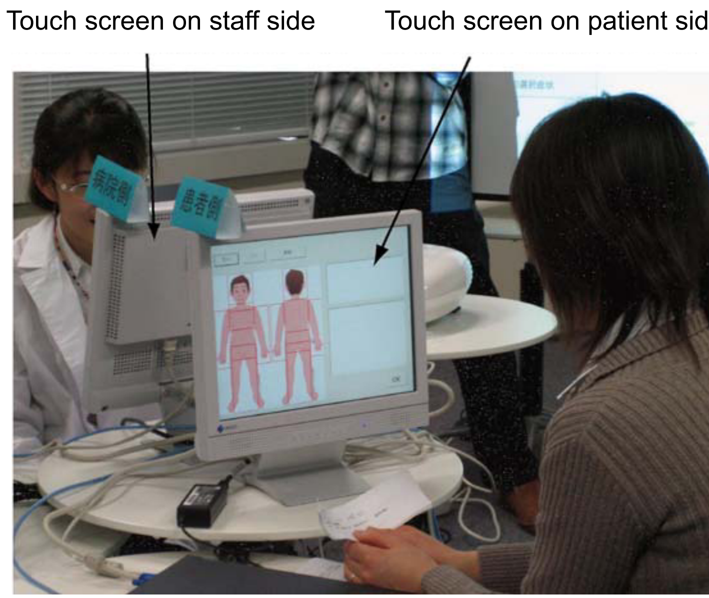


Fig. 5. Users in experiment

experiment, we placed two touch panels so the subjects were face-to-face, and the subjects communicated face-to-face using the conversation support system. We assumed that this communication is similar to that between actual hospital staff and patients and assigned the roles of hospital staff and patient to the subjects. The hospital staff began the usual medical reception and the patient responded based on supposed medical conditions. Figure 5 is a photograph of subjects carrying out the experiment.

3.2 Results

After the experiment, the subjects made the following comments.

- I was able to ask for basic information using this system.
- I feel that the system would be useful at hospitals that are frequented by many foreigners.
- Because medical translators often cannot provide assistance, this system will be helpful.
- Communication between hospital staff and patients is possible via this system.
- Registration of parallel texts is an excellent feature of this system.
- I hope that the system will be able to add and delete parallel texts at each hospital.
- When I was using the system I could not see my interlocutor.
- If we register a required sentence, it will be useful if it is translated so that a new parallel text will be available in the system. Moreover, the new parallel text should be accurate.
- The system is inconvenient because it has no technique by which the accuracy of the translation of a new parallel text can be checked.

The results of the experiment showed that Japanese hospital staff and Chinese patients can communicate accurately via this system. The medical field requires very accurate translations. Based on these comments, we realized that a method of checking the accuracy of the registered parallel texts must be developed.

4 Experiment Comparing the Proposed and Conventional Systems

4.1 Experimental Method

We carried out a conversation experiment to examine the time required for retrieval and the effect of using different interfaces on communication. The subjects were ten students from Wakayama University. They were trained in use of the system before doing the experiment.

In the experiment, we assigned the subjects the roles of hospital staff and patients. We carried out experiments using the conventional and proposed systems. In the conventional system, a hospital staff member and a patient use a printed list of parallel texts for conversation. The subject searches for the specified parallel text from the printed list, and then points out the parallel text to his interlocutor. Figure 6 is a photograph of

Table 1. Procedure at hospital reception

(1)	Are you a new patient?
(2)	Has it been ( ) month/months or more since your last consultation?
(3)	Do you have a health insurance certificate?
(4)	Please show the health insurance certificate.
(5)	Today, we have a translator. Do you wish to apply for translation?
(6)	Have you decided the department to be consulted?
(7)	Which department do you want to consult?
(8)	May I select the ( ) department?
(9)	Did you visit this hospital all by yourself today?
(10)	Do you have any allergies?
(11)	Do you have a fever?

Table 2. Patient answers in experiment

Language: You do not speak Japanese.
New patient: You are not a new patient.
Health insurance certificate: You have it.
Consultation: You consult.
Department of consultation: You decided.
Request for a translator: You want to request a translator.
Other situation: Visit alone, live on your own, have parents, drink alcohol, have a fever

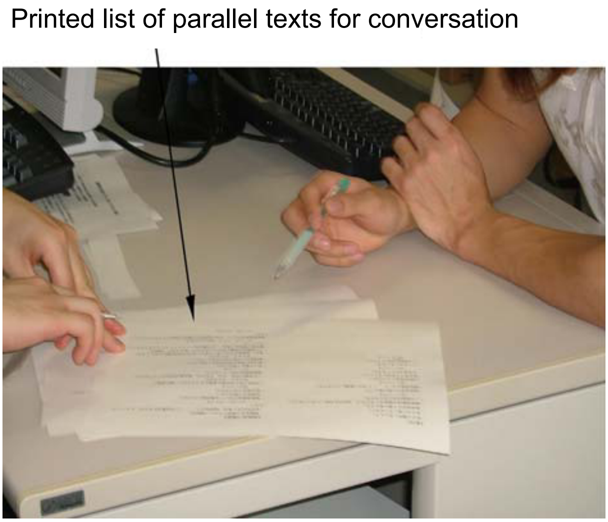


Fig. 6. Conversation using conventional system

a conversation using the conventional method. On the other hand, when our system was used, we conducted two different trials: one that displayed the conversational history and one that hid it.

Table 1 is an example of a list of questions asked of patients by hospital staff at a hospital reception desk. We designed the procedure of hospital reception using parallel text registration. The subjects playing the hospital staff role follow the hospital reception procedure. We instructed them not to look at the next question until the patient responded.

Table 2 lists various patient answers to the questions in the experiment. The patient's answers did not reflect the real medical condition of the subject. When the hospital staff pointed to the question, the patient responded as a patient might.

We did not allow the subjects to converse in Japanese. We allowed only gestures or nods based on the assumption that neither the hospital staff nor the patient could converse with each other in the same language. The experimental procedure was as follows:

(1) Pair the subjects and assign roles (hospital staff or patient).
(2) Hold a dialogue using the conventional method.
(3) Hold a dialogue using the proposed system with the conversational history displayed.
(4) Hold a dialogue via the proposed system with no display of conversational history.
(5) Switch roles and repeat steps (2) through (4).

We used a different hospital reception procedure and different patient answers during each of the six experimental trials. After the experiments, we asked the subjects to fill out a questionnaire.

4.2 Results

Table 3 shows average retrieval time and average response time. The former is the time that it takes a hospital staff member to search for a specified parallel text with each method, and the latter is the time it takes a patient to respond to the question. Tables 4, 5, and 6 are the results of the questionnaire in which a five-point Likert scale was used for the evaluation: 1: Strongly disagree, 2: Disagree, 3: Neutral, 4: Agree, and 5: Strongly agree.

In Table 3, there were few differences between the average retrieval and response times of the two systems. In Tables 4 and 5, there were few differences between the systems.

However, when our system was used, there were some problems in retrieving parallel texts.

First, when users were searching for parallel text, it was difficult for them to understand how parallel texts were grouped. This is because the lists had many parallel texts; thus, it seems that finding a parallel text is difficult and often takes a long time. In this

Table 3. Average retrieval time for questions, and average response time for questions from staff

	Average retrieval time (hospital)			Average response time (patient)		
Subject	Conventional method	History hidden	History displayed	Conventional method	History hidden	History displayed
1	12	10	10	7	5	5
2	7	9	15	5	5	5
3	23	9	8	5	6	5
4	8	11	8	9	8	6
5	18	23	10	6	8	7
6	16	12	12	7	6	5
7	17	14	10	8	7	7
8	9	18	7	6	5	8
9	13	12	10	6	6	4
10	11	20	10	3	7	5
Average	13	14	10	6	6	6

(Seconds)

Table 4. Results of the questionnaire on staff side of system

Questions	Average
(1) I could find the specified parallel text more easily than I could using the conventional system.	4.0
(2) Using this system, for a given patient response time, I did not experience more stress than with the conventional system.	3.6
(3) I could converse more smoothly by using the proposed system than by using the conventional system.	4.3
(4) I referred to the conversation history during the conversation.	3.4
(5) I think the conversation history helped me communicate.	4.2

Table 5. Results of questionnaire on patient side of system

	Questions	Average
(1)	I could respond to questions more easily by using the proposed system than by using the conventional system.	4.2
(2)	Using this system, for a given staff response time, I did not experience more stress than with the conventional system.	3.3
(3)	I could converse more smoothly by using the proposed system than by using the conventional system.	4.0
(4)	I referred to the conversation history during the conversation.	3.9
(5)	I think the conversation history helped me communicate.	4.0

Table 6. Results of the questionnaire on both systems

	Questions	Average
(1)	I felt that the displays were inconsistent.	2.1
(2)	I think that the difference in displays influenced the communications.	1.8
(3)	I used a touch screen.	3.9
(4)	I could operate the touch panel easily.	4.0

experiment, there were 197 registered parallel texts. The maximum number of parallel texts classified per scenario at the reception desk was 65.

Second, it was difficult to operate the scroll bar using the touch panel. On the patient side of the system, the user did not have to operate the scroll bar because most responses were either “Yes” or “No”. The staff side of the system had many parallel texts in the lists; thus, hospital staff needed to be able to operate the scroll bar.

These problems influenced the time it took to retrieve question sentences. To achieve smooth communication, we need to consider how to display the parallel text list more efficiently.

We found that different interfaces have little effect on communication because of the values of questions (1) and (2) in Table 6. These are with regard to the following: “Interfaces do not have to be the same because perspectives are different” and “Different interfaces are intuitive.” Therefore, we determined that interfaces based on standpoint are useful when the users’ standpoints are different.

Using the conventional method, the subjects often gestured and nodded during the conversation. With our system they focused on the display and seldom saw their interlocutor’s face. This is because the user had to focus intently on the display because he or she did not otherwise know when a question or a response was sent. We think that communication lacks emotive content if it is carried out using a non-native language and the faces of the interlocutors are hidden behind a screen. We found that face-to-face communication is also important. We need to consider and overcome this problem in the future.

5 Conclusion

We proposed a parallel-text based support system for intercultural communication at hospital reception desks. We found that:

- The proposed system effectively uses parallel texts to support intercultural communication. The results of the experiment on intercultural conversation show that correct intercultural communication is possible using parallel texts.
- Our system can retrieve a question in about the same time as a conventional system using a printed list.
- An interface based on a user's perspective can improve communication. The comparison of our system and the conventional system showed that the interface was easy to use.

Acknowledgements. The authors express their sincere thanks to Ms. Aguri Shigeno and Prof. Toru Ishida for their comments and support of the prototype system.

References

1. Wang, B., Cheng, X., Bai, S.: Example-Based Phrase Translation in Chinese-English CLIR. In: Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval, pp. 435–436 (August 2002)
2. Ikeda, T., Ando, S., Satoh, K., Okumura, A., Watanabe, T.: Automatic Interpretation System Integrating Free-style Sentence Translation and Parallel Text Based Translation. In: Proceedings of the Workshop on Speech-to-Speech Translation: Algorithms and Systems, Philadelphia, pp. 85–92 (July 2002)
3. Imoto, K., Sasajima, M., Shimomori, T., Yamanaka, N.: A Multi Modal Supporting Tool for Multi Lingual Communication by Inducing Partner's Reply. In: Proceedings of the 11th international conference on Intelligent user interfaces IUI '06, pp. 330–332 (January 2006)
4. Rayner, M., Bouillon, P., Van Dalsem, V., Isahara, H., Kanzaki, K., Hockey, B.A.: A Limited-Domain English to Japanese Medical Speech Translator Built Using REGULUS 2. In: Dignum, F.P.M. (ed.) ACL 2003. LNCS (LNAI), vol. 2922, pp. 137–140. Springer, Heidelberg (2004)
5. Chung, J.-w., Kern, R., Lieberman, H.: Topic spotting common sense translation assistant. In: CHI '05 extended abstracts on Human factors in computing systems, pp. 1280–1283 (April 2005)

Design and Development of a Pictogram Communication System for Children Around the World

Toshiyuki Takasaki and Yumiko Mori

NPO Pangaea, R&D center,
Cocon Karasuma Building 4F,
Shijokarasuma-Sagaru, Shimogyo-ku, Kyoto-city, Kyoto, 600-8411, Japan
{toshi,yumi}@pangaean.org

Abstract. Pangaea develops an intercultural collaboration environment using ICT (Information and Communication Technology) called the Universal Playground where children around the world can foster personal bonds regardless of their location, language, and cultural background, through a playful activity called "Pangaea Activity." Pangaea is a unique organization in that it has ongoing global fields for local children and has developed its own ICT system. This case paper reports how the Communicator, the pictogram communication software, was designed and developed. Development of the software and ICT system comes together through the Pangaea Activity menu, facilitation know-how, and field operation flow in order to bring the best performance toward its mission. As human-resources, funding, and time are limited, internal qualitative evaluations were conducted actively and quantitative evaluations were done in cooperation with external research groups.

Keywords: pictogram, intercultural communication, graphical user interface, usability, development process.

1 Introduction

1.1 Overview

We humans experienced the tragedy of *9/11* at the start of the twenty-first century. Although we should have learned, from our long history of war, to avoid war and instead take action for peace, there is still much strife and conflict in the world today. One of the fundamental reasons for this ongoing strife includes stereotypical threats resulting from the interactions of different religions, races, and nations. Meanwhile, science and technology has advanced rapidly and we now have a borderless cyber world, the Internet. The Internet enables people to obtain information instantly and to communicate with people worldwide through email and by web cameras. The cost of intercultural collaboration by ICT, Information and Communication Technology, is much cheaper than intercultural exchange which requires physical travel. Hence, the opportunity to communicate across cultures is increasing incessantly. In order to reduce the threats mentioned above, we have been making an effort to create the Universal Playground, where children from all around the world can create personal

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 193–206, 2007.

bondings through ICT regardless of their language, physical separation, cultural background and financial state. This paper is a case study in which we demonstrate the design and implementation of a pictogram communication media for use with the Universal Playground. Specifically, we will discuss our choice of pictograms over other media, our choice not to restrict the placement of pictograms in any specific layout, and finally, we will discuss the classification scheme employed by the pictogram dictionary to store the pictograms.

2 Background

2.1 Pangaea

Pangaea is a non-profit research and development project with the goal of creating a 'Universal Playground' for children. It creates an ICT environment where children, aged between nine and sixteen, can develop a personal and emotional 'bond' with each other around the world [15]. Geographical and language barriers as well as differences in social background are major factors that limit the opportunities for children to experience these 'bond.' Pangaea develops an online environment and tools through which children can spontaneously enjoy getting to know each other, share their experiences, and collaborate despite being physically separated. Using the Internet as a catalyst and connector, Pangaea provides a range of opportunities for children to 'bond' through Peace Engineering.

It also creates and implements playful and collaborative activity menus called Pangaea Activities at Pangaea's various local offices. By the end of 2006, Pangaea Activities take place at five locations in Japan; Nairobi, Kenya; Vienna, Austria and Korea. Both simultaneous and non-simultaneous activities are available in these locations. "Meeting", "Communicating", and "Connecting", Pangaea aims to develop content and tools to make these tasks enjoyable to the young participants. As a non-simultaneous activity platform, we have developed PangaeaNet, which is an online Universal Playground where children can communicate with each other and can share their work.

2.2 Pictogram Communication

The Communicator is a communication system made for the purpose of exchanging pictograms in the context of the various Pangaea Activities. Communicator enables children from different countries to communicate with each other in their own mother tongues, develop bonds, and to express their cultural and individual uniqueness freely. Communicator uses expressive and succinct pictograms. We named these pictograms in “pictons” (Fig. 1). More than five hundred pictons have been designed by the Picton Design Team, which consists of thirty volunteers mainly from Tama Art University in Tokyo. Presently, four hundred and fifty pictons are registered in the Picton Dictionary.

The Communicator also facilitates cultural diversity; when a child receives a message containing pictons, the *Picton Translation Function* [18] is invoked, by which the message is displayed in both the recipient's and sender’s picton sets. By looking at two picton messages which have the same meaning, yet possibly a different representation, it is possible for children to learn the cultural and personal background of their partners.

Fig. 1. A random sampling of pictograms designed by Pangaea

3 Case1: Communication by Pictograms

This chapter shows how we decided on pictogram as a shared communication method in Pangaea.

3.1 Motivation

In order to realize communication among Pangaea participants, a common language was needed. In international exchange and intercultural collaboration as well as in international commerce and politics, English is commonly used as a universal language. However, according to the statistics of Global Reach 2005, 64.8% of the users of the Internet have mother tongues other than English, which means that there are probably just as many children that do not understand English. In addition, the dependence on the English language suggests that there exist superior and inferior languages. In the past, we could do no better than to use English as a global language. However, we believed that we could change this with the help of the current ICT. In order to let children foster create bonds amongst themselves, the ideal language to use would be accessible to people around the world, would include rich emotional expressions, and would allow for the inclusion of the cultural identity of each group of participants.

3.2 Preliminary Survey

Before deciding on a specific communication method for the communicator, we set the following standards which the method must fulfill:

(1) It realizes heart-to-heart communication.
(2) It respects cultural diversity and self-identity.
(3) It is realized without learning new languages.
(4) It is made possible by computers and can be transmitted over the Internet.

We then proceeded to evaluate possible global communication methods on these specific points. By reading articles and interviewing researchers, we concluded that there are three candidate methods: using an artificial text-based language, using machine translation, and using a visual language.

First, one well known example of artificial text-based languages is Esperanto. Though Esperanto is useful in that a fluent speaker can convey ideas in detail, it requires learning the language from scratch. As it is time consuming to learn another language like Esperanto, and also given that they do not reflect any culture in particular, we decided not to use artificial languages to facilitate communication.

Second, while machine translation does enable us to exchange messages without learning a new language, there are many unsupported languages such as Khmer, the language of Cambodia. In addition, when we tested Japanese to English machine translation services on vulgar texts such as transcripts from chat sessions and email, we found the quality of translation not to be adequate. Thus, we concluded that while it would be possible to use machine translation for isolated portions of the Pangaea systems (perhaps the UI), we came to the conclusion that we risked losing too much meaningful content if we used machine translation as an exclusive communication tool.

We also evaluated using visual language [12]. We found that the main communication systems which use pictograms can be placed into three categories: universal signs, signs for the handicapped, and smilies. Good examples of universal signs are road signs, direction boards at airports, and the symbols of each sport played in the Olympic games. The second category of pictographic languages includes systems for Augmentative Alternative Communication (AAC) such as Blissymbolics [1, 16], PIC [13], and Elephant memory[8]. AAC assists people with severe communication disabilities to be more socially active in interpersonal interaction, learning, education, community activities, employment, volunteering, and care management. The third category includes pictograms that decorate text messages [17] such as emoticons and pictograms in cell phone. Some teenagers in Tokyo are reported to send email messages only using pictograms with rich emotional expressions. In comparison, Blissymbolics are relatively abstract, and they have a grammar which is more defined than emoticons.

After reviewing the aforementioned communication methods, we determined that a pictographic system would be the most practical and realistic system for Pangaea participants. Although a pictogram system is not as good at conveying precise information, it can be useful to facilitate heart-to-heart communication.

3.3 Pictogram Creation Activity

We conducted pictogram activities to investigate the feasibility of pictogram communication in the context of Pangaea Activities. The subjects were approximately fifty Japanese students nine to fifteen years of age. They drew pictograms on paper about prescribed themes, while the drawing process was overseen by Pangaea staff.

Their drawings expressed concrete images such as pictures and included more information by having more colors and lines than existing pictograms and symbols, in which redundancy is eliminated through simple design. While drawing, subjects started to show curiosity in the pictograms of the people around them as they began to ask each other, “What do you think this pictogram means?” Amongst themselves, they answered correctly more than 60% of the time, implying that more than 60% of the content of the pictograms was readily conveyable to other people. Furthermore, if they were allowed to guess thrice, the percentage of correct answers rose to 85%. When the children combined more than two pictograms to compose a message, we observed that the correct comprehension of the individual pictograms increased.

Also, we observed that some subjects did not like drawing and that they were not interested in drawing pictograms. However, most of them were interested in seeing each other’s pictograms. For example, an eleven year old girl commented that she enjoyed pictographic communication much more than text communication. And a thirteen year old boy commented that he would like to communicate with pictograms as he cannot master all of the languages in the world.

3.4 Lessons Learned

As a result of field testing the use of pictograms to convey meaning among children, we decided to go forward with our plan to use pictograms as the main vehicle of communication in subsequent Pangaea Activities. Though the correct answer rate of pictogram interpretation was high, when we asked a professional book editor to join the activity as a facilitator, the pictograms that he drew were incomprehensible to the children, although the pictograms were carefully and precisely drawn. The opposite also held: some facilitators and staff had trouble understanding the meaning of the children’s pictograms, while the children amongst themselves had much less difficulty understanding what they drew. It is plausible that this is due to the fact that the children are of a younger generation which differs in its ideology and depth of expression from that of the facilitators.

Some people argue that we should have conducted in-depth quantitative testing concerning the feasibility of pictogram communication. Pangaea has an on-going implementation fields almost every week so that we should provide a practical quick and reasonable solution. Based on qualitative observation and comments from children in Japan and some in Austria and Kenya, we had to take into account that there are children who are not good at drawing pictures and that visual representation of pictogram designs depends on cultural or social background [10, 11]. To deal with children who were not good at drawing, we decided to form a pictogram design team to make a standard pictogram set. To deal with social and cultural differences which caused unintended interpretation of some pictograms, we used an online pictogram survey to analyze the role that cultural differences plays in the interpretation of pictograms, as well as to determine which kinds of pictograms were more susceptible to misinterpretation than others.

4 Case2: Pictogram Message Rule

In this chapter, we discuss our decision to give more freedom of expression to children by allowing them to form pictographic text laid out both horizontally and vertically.

4.1 Motivation

After we decided to use pictograms as the communication method in Pangaea Activities, we started to develop a pictogram communication platform on a computer. We designed the overall system architecture, the back-end systems including the pictogram dictionary database and user management. In designing the user interface, one of the most important decisions made was the layout of the pictogram message pane. We had to decide how pictograms were to be laid out on the pictogram message pane. To that end, we had to learn how children compose pictogram messages. In addition, we had to consider if we needed to establish rules such as a grammar for pictogram communication.

4.2 Preliminary Survey

We conducted preliminary survey on the layouts of existing pictogram systems. We surveyed the pictogram systems in cell phones, emoticons, PIC, Blissymbolics, and ASCII Art. We found that cell phones, emoticons, and PIC employ text-editor style layouts. Blissymbolics defines a specific grammar which includes overlapping individual symbols to express new meanings. And a reconstruction idea of Blissymbolics was proposed which provides communication software with GUI [4]. ASCII Art is drawn on a multi-line text field; while some people include it in their email signatures, others use them on newsgroups and bulletin boards.

4.3 Pictogram Message Activity

We found that most editing layouts employ a text-editor style. This style is easier and simpler to develop and implement because the system doesn't need to retain the location data (coordinates) of each pictogram. However, some input methods displayed a canvas-style or extended text-editor style such as ASCII Art. The canvas style gives users the freedom of message composition compared to the text-editor style. And this freedom might induce creative and fun element of message composition and message interpretation for children. Therefore, we decided to introduce the pictogram message activity as a Pangaea Activity in order to observe how children compose pictogram messages, which would serve as a basis for deciding the layout of the interface of the pictogram messaging system. The test subjects were Japanese participants between nine and sixteen years of age. They composed pictogram messages on paper according to prescribed topics. The topics were varied and included "a memory from summer vacation," "a favorite story", "commenting and inquiring about overseas participants pictured in a Pangaea video letter" and also "making comments to the authors of animations created by overseas teenagers." Subjects composed pictogram messages by mounting Post-its (sticky) of their original pictogram designs and/or printed pictogram-cards designed by Pangaea on A3 or A4 paper. Both horizontal layouts and two-dimensional layouts were observed. In the horizontal layout case, pictograms were laid out horizontally in sequence and a new line was started when they have more than two messages, which was a form of the text-editor style.

As for the two-dimensional layout case, some subjects laid out pictograms both horizontally and vertically to indicate spatial relationships between the various

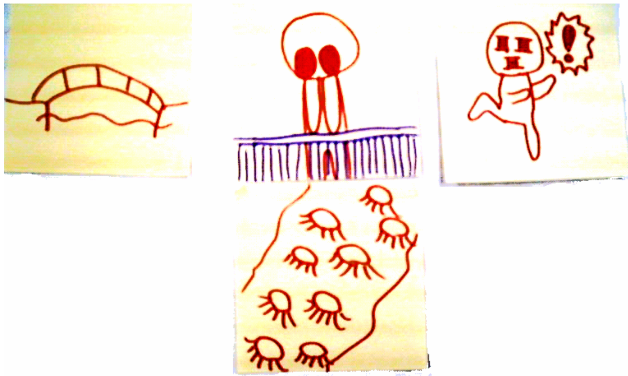


Fig. 2. This is an example of a pictogram message of two-dimensional layout. It means "I was surprised to see smack of jellyfish under a bridge."

concrete and abstract concepts as shown in Fig. 2. We found that pictographic messages with a two-dimensional layout took more time to compose than simply horizontal messages, though the time required also depended on the number of pictograms being used. However, it seemed that the process of thinking through the layout of pictograms in two-dimensions was enjoyable to the subjects and the ones who composed two-dimensional layouts felt like showing the message to other participants willingly. As a result, we decided to implement a Canvas style editor as we believed that a greater variety of creative and fun messages could be made, leading to more positive communication among participants.

4.4 Prototyping

After deciding the message layout style, we started implementing a prototype version of the Communicator, the software for pictographic message exchange. Since the PangaeaNet system was in its design phase, this prototype was implemented as a stand-alone Windows client application using the .Net framework. This prototype allowed users to compose pictographic messages by placing pictograms on the screen using drag-and-drop operations from the pictogram dictionary palette to the Pictogram Message Canvas. In this prototype, users could send and receive pictographic messages as email which is implemented using the SMTP/POP3 protocols. The header pane of the Message Canvas always shows the facial portraits of both sender and recipient to facilitate mutual awareness of the other.

4.5 User Testing

The prototype of the Communicator was introduced as a Pangaea Activity menu, and we observed participants who exchanged pictogram messages internationally by using the software. The subjects were approximately 30 children, aged between eight and seventeen, in Japan and Korea. This Pangaea Activity was carried out with a combination of other Pangaea Activity menus, which facilitates participants to learn about each other by sharing their vital information as well as their daily life.

The pictographic message logs, which were exchanged between Japan and Korea, were reviewed. In comparison with the paper based pictographic messages, the participants appeared more satisfied with the Communicator. Some subjects overlapped some pictograms to express a new meaning. For example, as shown in Fig. 3 right, a pictogram of a "Palm" was put over a pictogram of the "arm" to express "bye-bye." Besides, some subjects used lots of copies of a specific pictogram to express a new meaning. For example, many "tree" pictograms were laid out over the Message Canvas to express "forest." These expressions reflect very unique aspects of utilizing digital images, which can be made transparent and can also be superimposed over any number of images. In addition, we observed some empathic cases. For example, a participant who laid out pictograms vertically from top to bottom, as shown in Fig. 3 left, and said that he knows that there exist some languages written from right to left in the world such as Arabic. But he has never heard of languages written from the bottom up. He would like to show his message to everybody so he laid out pictograms from top to bottom." This subject was aware and considerate of the existence of others who had different languages, cultures, and social backgrounds, which is essential to intercultural collaboration. These observation results and comments convinced us to use a "Canvas-style" layout for the Communicator.

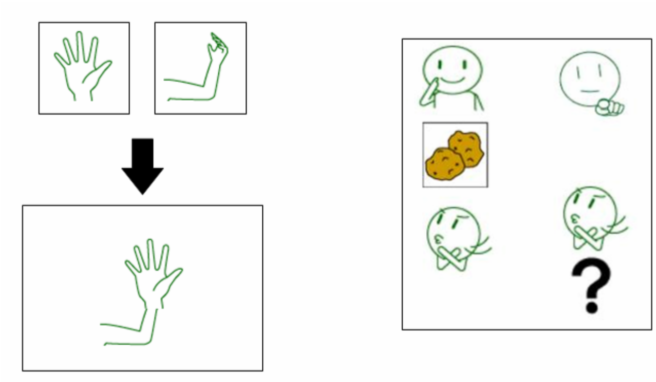


Fig. 3. Left: This is an example of overlapping pictograms. A "Palm" pictogram was put over an "arm" pictogram to express "bye-bye." Right: This is an example of empathic process of composing a message. This participant was aware of the right-to-left languages and laid out pictograms from top to bottom vertically.

4.6 Lessons Learned

We decided to use the Canvas style layout to the message pane of the Communicator based on data gathered from a paper-based activity of pictographic message composition, prototyping the Communicator, and from user experiments. We are still actively getting feedback from users. For example, some users requested that they be able to rotate pictograms freely on the Canvas, and others requested that they be able to scale pictograms freely. We are working on implementing these ideas. On the other hand, research partnerships were established with a number of academic institutions to conduct quantitative usability testing. This testing has already been conducted with respect to exploring the optimal size of pictograms and the canvas on screen, and the usable shape of the Canvas [7]. By merging results from qualitative observation of Pangaea and quantitative evaluation done by other research, Communicator is continuously evolving.

5 Case3: Categorization of Pictograms

This chapter explains the decisions made when creating the pictogram dictionary.

5.1 Motivation

The prototype version of the Communicator described earlier had a Pictogram palette which showed a list of pictograms. A user could drag a pictogram from the palette and drop it onto the message canvas. However, the Communicator at that time did not have any search capacities such as natural language queries. The pictogram palette only showed 200 pictograms from top to bottom with a vertical scroll bar (Fig. 4). As we were planning to increase the number of pictograms, we realized that it may become more difficult for a user to find a specific pictogram. Therefore, we decided to categorize the pictograms and implement more accessible palette.

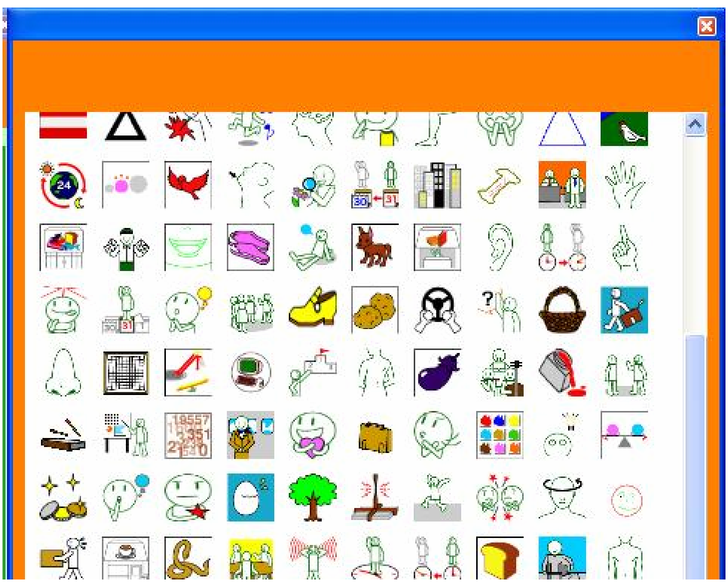


Fig. 4. Pictogram palette of the Communicator prototype

5.2 Preliminary Research of Categorization for Children

First, we surveyed existing classification systems tailored to both general audiences and children with advice from educators, usability researchers, AI researchers, and pictogram designers. We reviewed the aforementioned PIC, portal websites for children such as Yahoo! Kids, encyclopedias for children, categorization theories in cognitive science as well as classic categorization systems such as the Dewey Decimal system, thesauri, and ontology. This survey showed that a categorization system geared for children should prioritize popular topics among children such as entertainment and food.

5.3 Specifying Categorization

Although by this survey and by listening to advisors, we found abundant references to existing categorization methods, standardized categorization systems did not exactly fit our goals. Therefore, we set out the following guidelines characteristic of an ideal categorization system for children's pictograms:

(1) The categorization system should be defined from a children's point of view.
(2) The number of the primary categorizations should be between five and nine, based on the magical number seven theory [14].
(3) The fact should be kept in mind that a categorization system may depend on culture specific subtleties.

The third directive differs from the first two cases mentioned above in a way that this decision was not between the two choices such as "deciding if pictogram communication works" and "deciding horizontal-layout or two-dimensional-layout." Furthermore, there is no standardized method. Therefore, we decided to take more empirical method than in the other cases, by finalizing the first draft of the categorization system, implementing it at Pangaea activity, reviewing the feedback from children, and modifying it. We reviewed advices and ideas of advisors and those who had interacted with Pangaea children internationally for the longest time finalized the categorization system at Pangaea Activities as the following:

(1) Basic: Frequently-used pictograms such as “I”, “you”, “OK”, “like”, and so on
(2) Feeling: Facial expressions and emotional pictograms
(3) Entertainment: Sports, game, music, shopping and so on
(4) People: Family, friend, body part, and occupations
(5) Place: Countries, building structures, and stores
(6) Relation: Adjective and adverb pictograms as well as time and spatial pictograms
(7) Food: Tastes, ingredient, utensils
(8) Action: Motion and action pictograms
(9) Encyclopedia: Symbols of Pangaea Activity and others

5.4 Implementation

We applied this categorization to the pictogram palette of the Communicator as shown in Fig. 5. After completing the design and review process, 450 pictograms were placed on the pictogram palette. Pictograms with a clearly defined relationship, such as "small" and "big" or "brother" and "sister", were placed next to each other. The Communicator was implemented as an Adobe Flash based web application. Not only did we develop the software, we also provided laminated-cards which display a list of pictograms and their categorization.

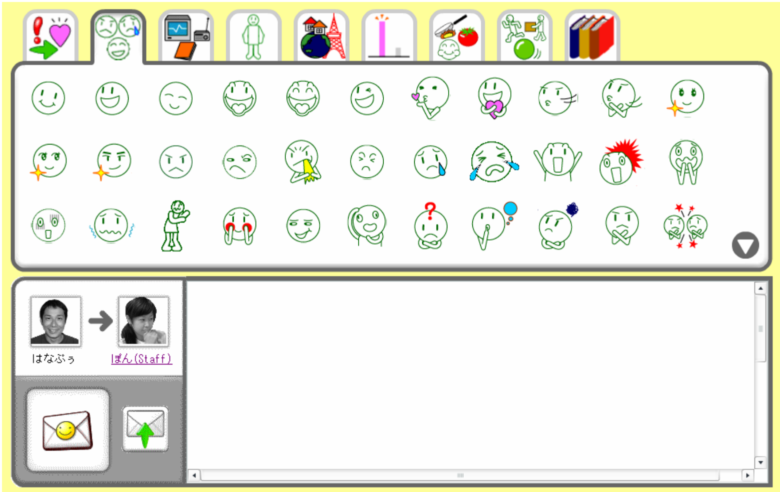


Fig. 5. This is a screenshot of updated Communicator, which consists of the pictogram categories on the top and a message canvas at the bottom. In this particular case, the “Feeling” category is selected.

The Communicator has been used by 120 children in Japan, Korea, Austria, and Kenya during Pangaea Activities. At every Pangaea Activity, we have collected feedback about the Communicator from participating children through feedback forms. We have also collected feedback from facilitators and technical staff. The feedback showed three main points. First, the current palette enables users to see and find pictograms more easily than previous versions. Second, as children spend more time using the program, they become more comfortable with the categorization method, and ultimately better at finding pictograms. Third, categorization of the pictograms helps to identify pictograms that are essential, yet missing from the dictionary. In case users cannot find an appropriate pictogram, we made a pictogram request form with which the user may request new pictograms from the NPO Pangaea R&D center.

5.5 Lessons Learned

We have described the processes of categorizing pictograms and implementing the pictogram palette. The method of categorization is still in the development stage. We also have been conducting academic research on the text search of pictograms, as well as on the categorization system with our research partners.

6 Conclusion

The first case showed the decision making process of selecting pictograms as the shared communication method in Pangaea Activity, which was during the beginning phase of pictogram project.

We have described the process by which we selected pictograms to be the communication method in the Pangaea Activity. We surveyed possible communication methods other than English, such as artificial language, machine translation and pictograms. By setting guidelines and priorities, we determined that pictograms are better than the other systems. Paper-based pictogram activities and observations revealed the positive feasibility of pictogram communication when participants share common generational traits, tasks such as Pangaea Activity, and curiosity toward others. Therefore, we decided to implement a system using pictograms. At the same time, we learned that the Pangaea Activity reflects an abundance of communication and interaction among participants.

We have also described how we established the rules of pictogram messaging. We surveyed phenomenon such as the rules of cell phone emoticons and existing pictogram system such as Blissymbolics. After prototyping the Communicator and conducting user experiments, we decided on a Canvas style that defines almost no rules. This is because empathic messages along with fun and creative messages were considered essential to intercultural collaboration.

We have also described the reasoning behind the categorization of the pictogram palette of Communicator. Compared to the other two cases above, this decision was the most empirical. This is because, though there were clear differences in each possible solution in the other two cases, as in a multiple-choice question, defining categorization was more of an open-question. Also, we found that a standardized categorization system which exactly fit to our purpose does not exist. Therefore, we gathered ideas and advice from advisors, prioritized the Pangaea participants' point of view, and made up our mind to decide on our own categorization system. In addition, we prepared a feedback form for participants and a Pangaea Activity Report form for facilitators and technical staff. After implementing the Communicator with a categorized pictogram palette, we gathered feedback from Pangaea participants. In fact, our development process was turned out to be an effective and practical UI design method with children according to usability researches [2, 5].

As NPO Pangaea, with limited resources, has on-going field-activities for children worldwide, we have developed and delivered each Activity content and software and have received feedback from participants without forgetting the overall mission of the entire NPO Pangaea operation. Therefore, we evaluate feasibility qualitatively through an empirical approach. At the same time, as described above, we have kept partnerships with academic research institutes to conduct qualitative researches deeply on specific topics such as usability of Communicator and a search system for pictograms [3]. In the middle of our development, the Language Grid Project, started in NICT (Japan). It aims to create a global language platform by connecting language resources, such as machine translation and dictionaries [6, 9]. We developed a partnership and started testing a multilingual system for the communication among international Pangaea facilitator staffs. In the near future, we may utilize this system by combining it with the pictogram communication system. This may lead to further possibilities of future communication in intercultural collaboration.

This is on-going project. As we have tested with the Japanese and Korean children, as well as some Austrian and Kenyan children, we should observe more users' interaction cases among various cultures and refine the user interface as well as pictogram design. One of the most important practices of ours is that we have developed not only software and an ICT system but also the Pangaea Activity menu, facilitation know-how, and operation flow to make the most use of them in pursuing our mission. These non-ICT elements are indispensable in implementing the Universal Playground.

Acknowledgments. We would like to thank Prof.Toru Ishida of Kyoto University, Dr. Kazuhiko Nishi, Prof. Mitchel Resnick of the MIT Media Lab, and Walter Bender of the former Director of the MIT Media Lab for their early advice and support. We would also like to thank Heeryon Cho and members of Kyoto University, Dr. Rieko Inaba and other members of NICT in Japan for their academic help, Alexander Patrikalakis and Tom Scharfeld for voluntary English correction, volunteers in Tama Art University in Japan for their pictogram designs, Tomoya Terasaki and Junko Kouta for their system and GUI design, Yoko Yamada and Hayato Aoshima for their help of surveys, and international Pangaea volunteers and staffs for their great support.

References

1. Bliss, C.K.: Semantography (Blissymbolics). Sydney Semantography Publications (1965)
2. Cooper, A.: About Face The Essential of User Interface Design. IDG Books (1995)
3. Cho, H., Ishida, T., Yamashita, N., Inaba, R., Mori, Y., Koda, T.: Culturally-Situated Pictogram Retrieval. In: Ishida, T., Fussell, S.R., Vossen, P.T.J.M. (eds.) IWIC 2007. LNCS, vol. 4568, Springer, Heidelberg (2007)
4. Dhande, S., Roy, D.: Visual Representations for Learning: Doodle project (2002), http://www.media.mit.edu/cogmac/projects/doodle.html
5. Druin, A.: Cooperative Inquiry: Developing New Technology for Children with Children. In: Proceedings of ACM CHI 99, pp. 592–599. ACM Press, New York (1999)
6. Inaba, R., Murakami, Y., Nadamoto, A., Ishida, T.: Multilingual Communication Support Using the Language Grid. In: Ishida, T., Fussell, S.R., Vossen, P.T.J.M. (eds.) IWIC 2007. LNCS, vol. 4568, Springer, Heidelberg (2007)
7. Inaba, R., Takasaki, T., Mori, Y.: How do kids use pictograms? In: Conference on Computer Supported Cooperative Work (CSCW2006) (suppl.), pp. 179–180 (2006)
8. Ingen-Housz, T.: The Elephant's Memory. In: Search of a Pictorial Language. Learning Technology Review, 2-25 (Spring/Summer 1999)
9. Ishida, T.: Language Grid: An Infrastructure for Intercultural Collaboration. In: IEEE/IPSJ Symposium on Applications and the Internet, keynote address (SAINT-06), pp. 96–100 (2006)
10. Koda, T.: Cross-cultural Study of Avatars' Facial Expressions and Design Considerations within Asian Countries. In: Ishida, T., Fussell, S.R., Vossen, P.T.J.M. (eds.) IWIC 2007. LNCS, vol. 4568, Springer, Heidelberg (2007)
11. Koda, T., Ishida, T.: Cross-cultural Comparison of Interpretation of Avatars' Facial Expressions. In: IEEE/IPSJ Symposium on Applications and the Internet (SAINT-06), pp. 130–136 (2006)
12. Marcus, A.: Icons, symbols, and signs: visible languages to facilitate communication. Interactions 10(3), 37–43 (2003)
13. Maharaj, S.: Pictogram ideogram communication. Saskatoon, SK. The Pictogram Centre, Saskatchewan Association of Rehabilitation Centres (1980)
14. Miller, A.G.: The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. The Psychological Review 63, 81–97 (1956)

15. Mori, Y.: Atoms of Bonding: Communication Components Bridging Children Worldwide. In: Ishida, T., Fussell, S.R., Vossen, P.T.J.M. (eds.) IWIC 2007. LNCS, vol. 4568, Springer, Heidelberg (2007)
16. Olaszi, P., Koutny, I., Kálmán, S.L.: From Bliss Symbols to Grammatically Correct Voice Out-put: A Communication Tool for People with Disabilities. In: International Journal of Speech Technology, January 2002, vol. 5(1), Springer, Netherlands (2002)
17. Rivera, K., Cooke, N., Bauhs, J.: The Effects of Emotional Icons on Remote Communication. In: CHI Conference Companion, pp. 99–100 (1996)
18. Takasaki, T.: PictNet: Semantic Infrastructure for Pictogram Communication. In: International WordNet Conference (GWC-06), pp. 279–284 (2006)

Cross-Cultural Study of Avatars' Facial Expressions and Design Considerations Within Asian Countries

Tomoko Koda

Faculty of Information Science and Technology, Osaka Institute of Technology, 1-79-1 Kitayama, Hirakata city, Osaka, 573-0196, Japan
koda@is.oit.ac.jp

Abstract. Avatars are increasingly used to express our emotions in our online communications. Such avatars are used based on the assumption that avatar expressions are interpreted universally among any cultures. However, our former study showed there are cultural differences in interpreting avatar facial expressions. This paper summarizes the results of cross cultural evaluations of avatar expressions among five Asian countries. The goals of this study are: 1) to investigate cultural differences in avatar expression evaluation and apply findings from Psychological study in human facial expression recognition, 2) to identify design features that cause cultural differences in avatar facial expression interpretation. The results confirmed that 1) there are cultural differences in interpreting avatars' facial expressions among Asian countries, and the psychological theory that suggests physical proximity affects facial expression recognition accuracy is also applicable to avatar facial expressions, 2) use of gestures and gesture marks may sometimes cause counter-effects in recognizing avatar facial expressions.

Keywords: avatar, character, facial expression, cross-culture, network communication.

1 Introduction

Since instant messenger and chat services are frequently used in our daily communication beyond nationality and languages, emoticons and expressive avatars are widely used to provide nonverbal cues to text-only messages [1, 2, 3]. Recent growth of virtual world [4] attracts worldwide attention to avatar mediated communication both from entertainment and businesses. Studies on emoticons and avatars report positive effects on computer-mediated communication. Those studies indicate that emoticons and avatars improve user experiences and interactions among participants [5, 6, 7] and build enthusiasm toward participation and friendliness in intercultural communication [8, 9].

However, these avatars are used based on an implicit assumption that avatar expressions are interpreted universally across cultures. Since avatars work as graphical representations of our underlying emotions in online communication, those expressions should be carefully designed so that they are recognized universally. We need to

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 207–220, 2007.

closely examine cultural differences in the interpretation of expressive avatars to avoid misunderstandings in using them.

However, few studies have compared the cultural differences in interpreting avatars. One of those studies compared interpretations of avatars' animated gestures between the Netherlands and Japan [10]. Their results showed that there are cultural differences in perceived valance in animated characters between the two countries. Japanese women perceived stronger emotions in some animated gestures of an avatar, i.e., bowing, than the Dutch subjects, although there were no overall differences in interpretation of the presented gestures. In our former study, we conducted a cross-cultural experiment in the form of a series of discussions on a multilingual BBS with expressive avatars between China and Japan [8]. The results show some facial expressions used in the experiment were interpreted completely differently and used for different purposes between Chinese and Japanese. Those "misinterpreted" expressions are "sweat-on-the-face," "wide-eyed," and "closed-eyes." For example, the "wide-eyed" expression was interpreted as "surprised" by the Japanese subjects, while the Chinese subjects interpreted it as "intelligent" and used it when presenting a novel idea or asking questions. We observed that the Japanese subjects tried to confirm the meaning of the Chinese subject's message with the "wide-eyed" expression. This is one example of communication gaps caused by different interpretations of avatar expressions between the two countries.

The above two studies were each conducted between only two countries. We need to conduct an evaluation experiment among multiple countries in order to investigate cultural differences in avatar expression interpretation and what kinds of expressions are interpreted universally and what kinds are not. We believe the results would serve as a design guideline for universal avatar expression that would not lead to miscommunication.

In our previous study [11], we applied findings from psychological studies on human facial expressions, since there have been a much wider variety of studies in psychology on human expressions than on avatar expressions. The most widely accepted findings come from the work of Ekman. He states that seven emotions, namely, anger, fear, disgust, surprise, sadness, happiness and contempt, are universally expressed by all cultures. However, he also argues the implications and connotations of those facial expressions are culturally dependent, and the degree of allowance in showing or perceiving those expressions socially differs across cultures [12]. Recent psychological research found evidence for an "in-group advantage" in emotion recognition. That is, recognition accuracy is higher for emotions both expressed and recognized by members of the same cultural group [13]. Elfenbein et al. state, "This in-group advantage, defined as extent to which emotions are recognized less accurately across cultural boundaries, was smaller for cultural groups with greater exposure to one another, for example with greater physical proximity to each other [13]." Also, the decoding rule implies that we concentrate on recognition of negative expressions, since misinterpretation of negative expressions leads to more serious social problems than misinterpretation of positive expressions would cause [15].

We conducted an open web experiment to compare interpretations of avatars' facial expressions among 8 countries, namely, Japan, South Korea, China, United States, United Kingdom, France, Germany, and Mexico in our previous study [11]. The results indicated that there are cultural differences in interpreting avatars' facial expressions, and the in-group advantage was found in interpreting avatar expressions. The next step is to validate each avatar's graphical design and find the design features that would lead to cultural difference in interpretation.

In this paper we summarize the results of a further experiment within 5 Asian countries using a more controlled experiment set than the one used in the previous experiment. The reason for conducting the experiment within Asia is to validate the cultural differences found across Asia, Europe, and North America in the previous experiment are again applicable within Asian countries, which have less geographical distances.

The goal of this experiment is: 1) to investigate cultural differences in avatar expression evaluation and apply findings from Psychological study in human facial expression recognition, namely the "in-group advantage", within Asian countries, 2) to identify design features that might cause cultural differences in avatar facial expression interpretation.

2 Experiment Overview

2.1 Experimental Procedure

The experiment was conducted in 2005 as part of the Intercultural Collaboration Experiment (ICE2005), jointly hosted by Chinese, Japanese, South Korean, Malay, and Thai universities and research institutes. The experiment was set on WWW, which was accessible only to pre-registered participants from the participating countries.

The experiment was developed using the application of Macromedia Flash. Participants first answer a brief questionnaire on their background profile such as their nationality and mother tongue. The main experiment starts after the questionnaire, which is presented as a matching puzzle game as shown in Fig. 1. Participants are requested to match 12 facial expressions to 12 adjectives. The 12 facial expressions are displayed in a 4 x 3 matrix and the 12 adjectives as buttons below the matrix. As shown in Fig. 1, participants can drag/drop the adjective buttons to/on the 12 expressions and continue changing the location of each button until they are satisfied with their answer. One avatar representation is chosen randomly from 10 avatars, and facial expression images are randomly placed in the 4 x 3 matrix. The adjective buttons are always displayed in the same order, and the 12 adjectives are always the same (see sec. 2-2 for the adjectives used in the experiment).

Participants' answers to the puzzle game and questionnaire, as well as their background profile including gender, age, county of origin, and native language, are logged in the server for later analysis. Participants are required to continue the experiment until they finish evaluating all the 10 avatar designs.

The adjectives can be shown in English, Chinese, Korean, and Japanese (all validated by native speakers). Participants from countries where the above languages are primarily spoken can see the adjective selections in their native language according to the background profile. Japanese, Chinese, Korean paticipants are shown the adjectives in their native language, and Thai and Malay participants in English. The participants from Thailand and Malaysia are fluent in English.

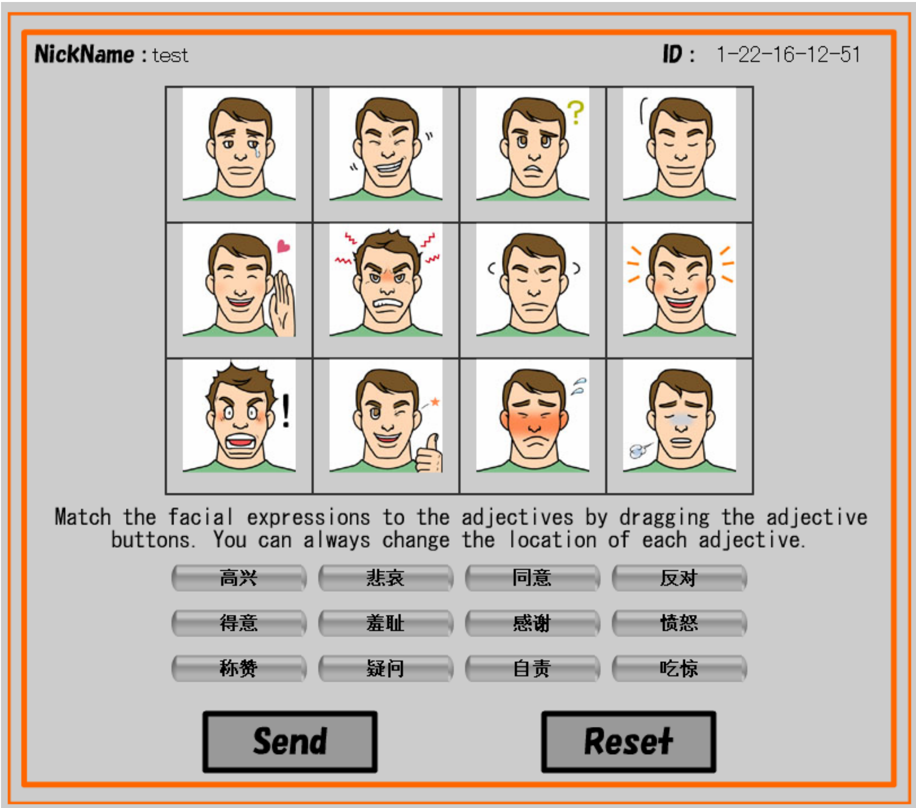


Fig. 1. Experiment screen shown as a matching puzzle game (Example screen for Chinese participants)

2.2 Avatar and Expression Design

Commercially used avatars are represented not by photo-realistic images but as caricatures or comic figures. We prepared 10 avatar representations drawn by three Japanese designers using Japanese comic/anime drawing style. By using avatars drawn with techniques from one culture, we can use those avatars as "expressers" and subjects as "recognizers" as in [15]. Accordingly, comparing the answers between Japanese users and those of other countries made it easier to validate the in-group advantage. Fig. 2 shows three examples from the 10 avatar representations.

The 12 expressions used in the experiment are "happy," "sad," "approving," "disapproving," "proud," "ashamed," "grateful," "angry," "impressed," "confused," "remorseful," and "surprised" as shown in Fig. 3. Those expressions are selected from Ortony, Clore and Collins' global structure of emotion types, known as the OCC model [16]. These are commonly used expressions in chat and instant messenger systems [1, 2, 3], and they reflect those emotions desired by the subjects for intercultural communication in [11].

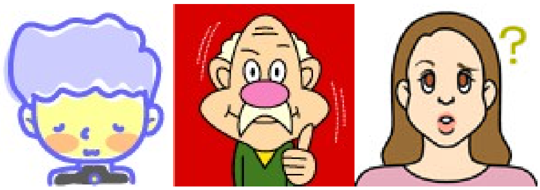
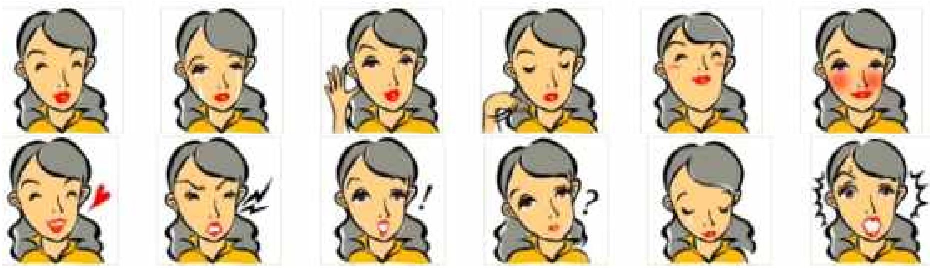


Fig. 2. Examples of avatar representation



Notes: From top left, happy, sad, approving, disapproving, proud, ashamed, grateful, angry, impressed, confused, remorseful, and surprised, drawn in Japanese comic style.

Fig. 3. Twelve facial expressions of one of the avatars

These 12 expressions are paired as valanced expressions as defined in the OCC model, that is, negative/positive emotions that arise in reacting to an event or person. "Happy," "approving," "proud," "grateful," and "impressed" are positive expressions, while "sad," "disapproving," "ashamed," "angry," "confused," and "remorseful" are negative expressions, leaving "surprised" as a neutral expression.

The experiment procedure and the matching puzzle game was the same as the one conducted in the previously except the following. These changes are made to control the experimental conditions more strictly.

1) Only the pre-registered participants can access the experiment site, while the participants in the previous experiment were freely access the experiment site.

2) The number of avatar design used in this experiment is limited to 10 instead of 40. The 10 avatar designs are selected according to the design features to express emotions in order to clarify difference in interpretations. The designs are categorized into three groups, namely, expression only, expression with a gesture mark, and expression with gesture.

3) Participants evaluate all the 10 avatar designs in this experiment, while the participants could stop evaluating the avatar designs any time in the previous experiment. Thus, the avatar designs and the number of avatars each participant evaluates are the same across participants in this experiment.

3 Results

The web experiment was conducted in July, 2005. The participants are gathered through research collaboration members among Japan, South Korea, China, Malaysia, and Thailand. Participation was not mandatory, but strongly encouraged among participants. The numbers of questionnaire answers are 19 from Japan, 12 from South Korea, 30 from China, 16 from Malaysia, and 15 from Thailand. The participants who answered the questionnaire are in their 20's and 30's, and the ratio of male and female was 1:1.

3.1 Difference in Interpretation of Avatar Facial Expression

This section aims to investigate whether there are cultural differences in interpreting avatar facial expressions.

The participants' answers to the puzzle game are analyzed by calculating matching rates between expressions and adjectives. There is no correct answer to the matching puzzle, but the avatar designers' original intention can be used as an expresser's "standard" answer. Each expression and adjective is assigned a number (1-12) within the system. The designer's intended pairs are described as (1,1), (2,2), (3,3), (4,4) reflecting (expression number, adjective number). We calculated each country's number of "expression-adjective" pairs that are the same as the designers' pairs. Consequently, here, "matching rate" means the percentage of pairs of expressions and adjectives that match the avatar designer's intentional pairs. For example, the matching rate of answer pairs

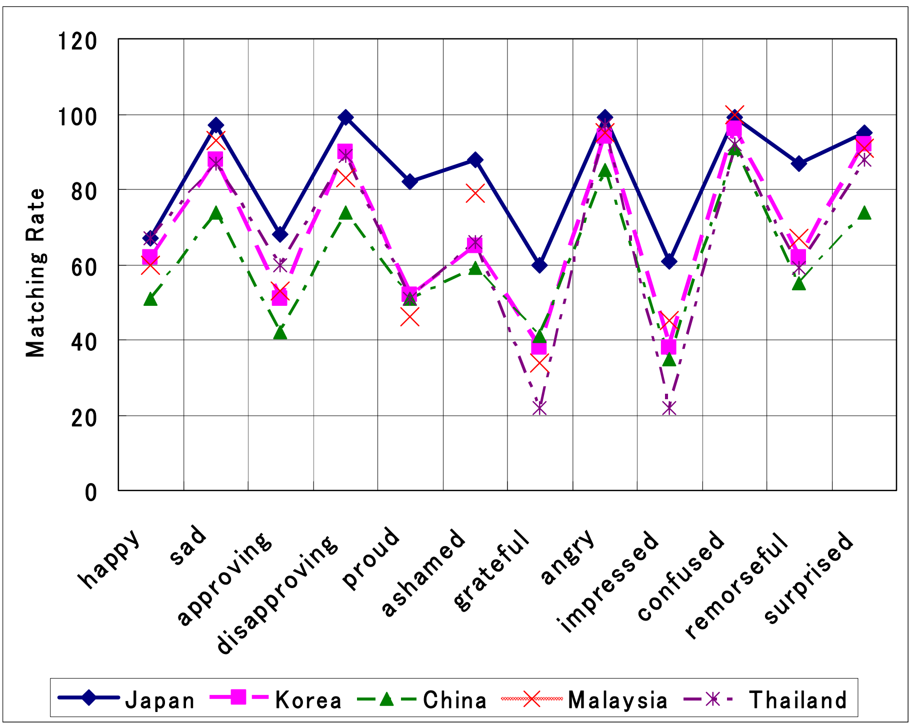


Fig. 4. Matching rate of each expression by country within Asia

(1,5), (2,1), (3,3), (4,9) is 25%. We use this "expression-adjective matching rate" in comparing the answers to the 12 facial expressions.

Fig. 4 shows the matching rates shown by expression and country. When we focus on the matching rates by country, Japan's matching rates are the highest in the five countries in all expressions. This means the degree of matching the expresser (avatar designer)'s intention and the answers of the recognizers (participants) is high. Hence the in-group advantage within the same country is identified in this experiment. This result further confirms the results of the previous experiment, in which Japan's answers had significantly highest matching rates among the eight countries, namely, Japan, South Korea, China, the United States, the United Kingdom, Germany, France, and Mexico. Thus, this result suggests that there are cultural differences among the five Asian countries which geographical distance is smaller.

When we focus on the matching rates by facial expression in Fig. 4, we again observe that the negative expressions have higher matching rates than the positive ones. Thus, as found in [11], the result suggests that the decoding rule is applicable to the answers to the five Asian countries.

3.2 Analysis of Recognition Accuracy by Facial Expression Design

This section analyzes the design features that would cause cultural differences in interpretation of avatar facial expressions. Among the facial expressions that have lower matching rats than others, we analyze the answers to the "proud", "grateful", and "impressed" expression by country.

The design features that are used in the above three expressions are categorized into three groups, namely, "facial expression only", "facial expression with a gesture mark", "facial expression with gesture."

Analysis of the design that uses facial expression only. Firstly, we analyze the answers to the avatar design that uses facial expression only, by using the "proud" expression. As in Fig. 4, the difference in the matching rate of Japan and other countries is the biggest in the "proud" expression. Fig. 5 shows the design examples of the "proud" expression, and Fig. 6 shows the answers to the "proud" expression shown by country. As shown in Fig. 5, "chin-up" expression is used to express "proud" in the avatar design. The detailed answers by country to the proud expression in Fig. 6 shows that more than 90% of answers to the "proud" expression were "proud" in Japan, while other countries' answers has only about 50% accuracy. The expressions that are mixed up with "proud" are "grateful", "impressed", and "happy", which all categorized as positive expressions.

Fig. 5. Design samples of "Proud" expression

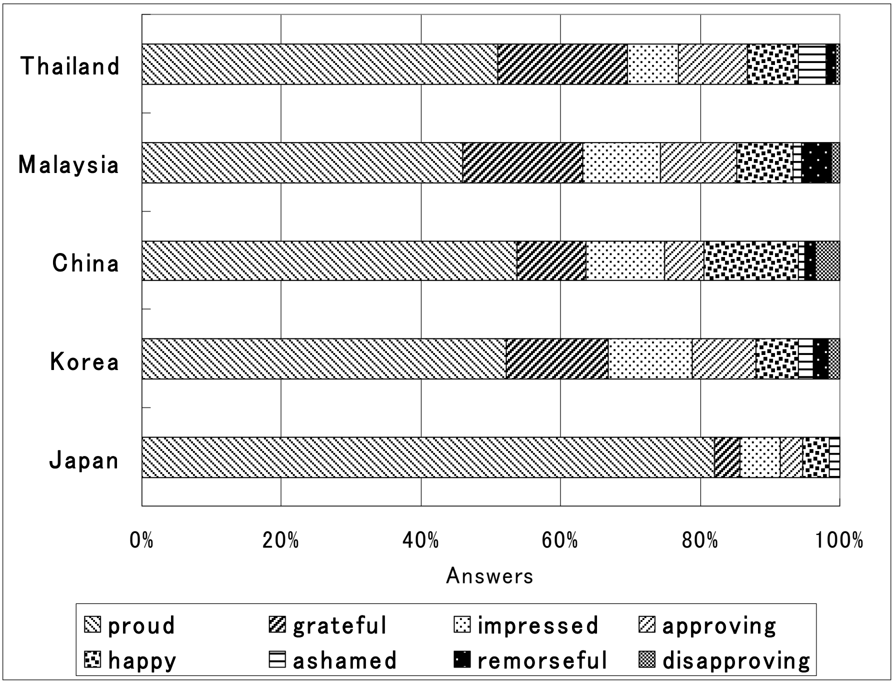


Fig. 6. Detail answers to “proud” expression shown by country

Analysis of the design that uses facial expression only and facial expression with a gesture mark. Secondly, we analyze the answers to the avatar design that uses both “facial expression only”, and “facial expression with a gesture mark”, using “grateful” expression. As in Fig. 4, the “grateful” expression has one of the lowest matching rates among positive expressions.

Fig. 7 shows the design examples used to express “grateful” expression. There are two designs used to express “grateful” expression. The first used facial expression only to express “grateful”, and the latter used facial expression with a gesture mark (heart mark) as shown in Fig. 7.

Fig. 7. Design samples of "Grateful" expression with a heart mark

Fig. 8 shows the detailed answers to the “grateful” expression presented by facial expression only, and the ones with a gesture mark. The answers of Japan to the design with a heart mark has higher matching rate (the percentage of answers that answered

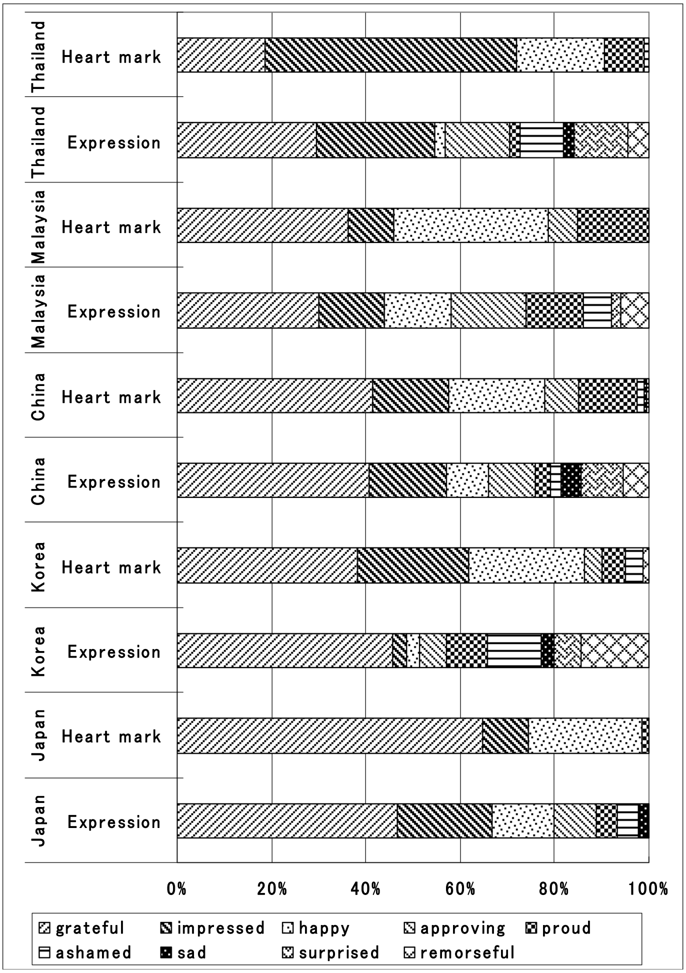


Fig. 8. Detail answers to the "grateful" expression (comparison of answers to the design that used facial expression only and with a heart mark)

"grateful") than the ones to the design that used facial expression only. While in other countries, using a heart mark does not necessarily result in higher matching rate. Especially in South Korea and Thailand, answers to the design used "facial expression only" have higher matching rate (the percentage of answers that answered "grateful") than the ones to the design that used "a heart mark". Adding a heart mark to the "grateful" expression design increased the number of answers that answered "impressed" in South Korea and Thailand.

Analysis of the design that uses facial expression with a gesture mark, and facial expression with gesture. Lastly, we analyze the answers to the avatar design that uses" facial expression with a gesture mark", and "facial expression with gesture", by using the "impressed" expression. As in Fig. 4, the "impressed" expression has one of the lowest matching rates among positive expressions.

Fig. 9 shows the design examples used to express "impressed" expression. There are two designs used to express "impressed" expression. The first used "facial expression with a gesture" ("clapping hands" gesture) to express "impressed", and the latter used "facial expression with a gesture mark" (exclamation mark "!") as shown in Fig. 9.

Fig. 10 shows the detailed answers to the design that used facial expression with a clapping hand gesture, and Fig. 11 shows the detailed answers to the design that used facial expression with "!". The detailed answer of Japan shows the highest matching rate, about 80% (percentage of answers that answered "impressed") among the five countries. While in other countries, detailed answers vary according to the design used to express "impressed". Especially in China, the "impressed" expression with clapping hands gesture is interpreted as "approving" more often as "impressed". In Thailand, the "impressed" expression with a "!" mark is interpreted as "grateful" rather than "impressed".

Note: Right: with "clapping hands" gesture, Left: with "!" mark)

Fig. 9. Example of the designs for "impressed" facial expression

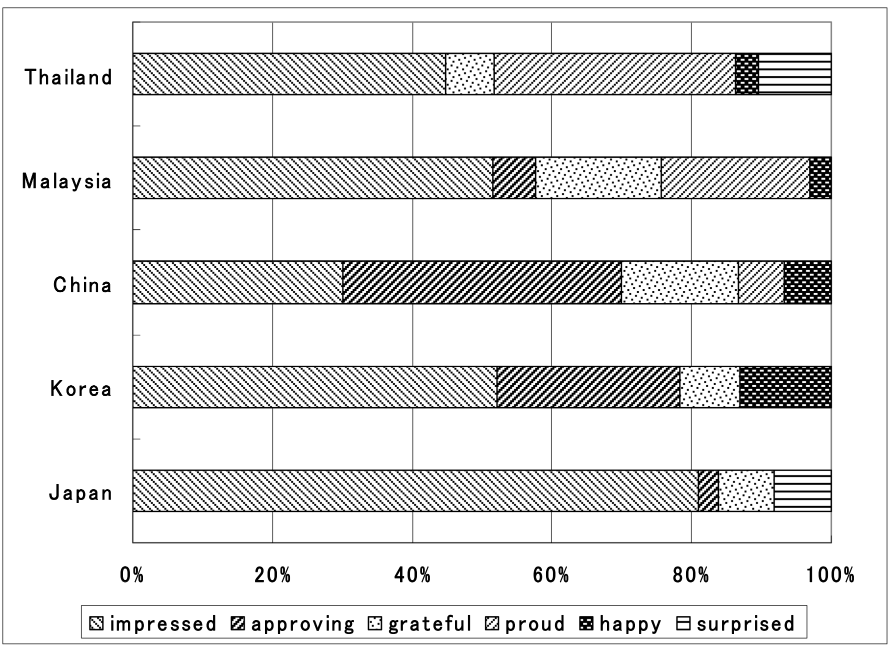


Fig. 10. Detail answers to "Impressed" expression with a clapping gesture shown by country

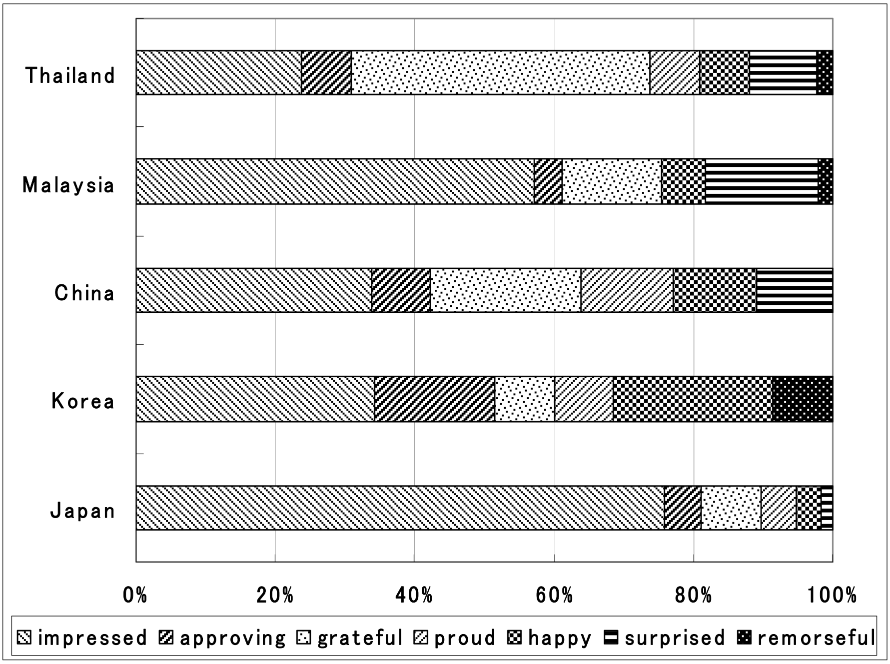


Fig. 11. Detail answers to "Impressed" expression with "!" shown by country

4 Discussion

The results shows the avatar facial expressions designed using the Japanese comic drawing techniques have higher recognition accuracy within Japanese participants than the participants in other countries within Asia. Thus this is another indication that the in-group advantage within the same country, which is found in recognizing human facial expressions, is also applicable to avatar expressions, when compared to the answers among the five Asian countries. We again found results that support the decoding rule, that negative expressions are more accurately recognized than positive ones in human facial expression, in avatar expressions.

Next analyses were made using different avatar expressions designs in the four facial expressions that had the lowest recognition accuracy. There are three different designs used in the experiment, namely, 1) facial expression only (i.e., "proud" expression), 2) facial expression with a gesture mark (i.e., a heart mark in the "grateful" expression, an exclamation mark in the "impressed" expression), 3) facial expression with a gesture (i.e., a clapping hand gesture in the "impressed" expression).

The results showed that Japan's recognition accuracy is the highest in all the three designs (facial expression only, facial expression with a gesture mark, and facial expression with a gesture) among the five Asian countries, and adding a gesture mark increases the recognition accuracy of Japan compared to the one to the designs that use facial expression only. While other countries than Japan, using a gesture mark tend to decrease the recognition accuracy compared to the one to the design that use facial expression only. Thus using a gesture mark does not necessarily improve the recognition accuracy in other countries than the expresser's country. Using a gesture with facial expression caused varied answers in other countries than Japan. Thus avatar expression designs with gesture do not lead to a better interpretation in this experiment.

Similar cultural differences in interpreting gestures in pictograms are reported in [17]. The survey was conducted using pictograms developed and used in NPO Pangaea's communication software, which allows children all over the world to communicate online regardless of their mother tongues using pictograms [18, 19]. The survey was conducted between the United States and Japan to ask meanings of 120 pictograms used in the Pangaea's communication software. The results suggested interpretations of gestures in pictograms vary according to culture [17]. Cho states the reason for these cultural differences can be explained by psychological studies by Efron [20] and Ekman [21]. [20] finds evidences for a human gesture to have different meanings according to culture, and [21] categorizes human gestures, among which "emblem" gestures (symbolic gestures) are cultural dependent. Both [17] and this study on avatar interpretation find "emblem" gesture (crossing arms to indicate "NO" in the former, clapping hands in the latter) has cultural differences in their interpretation. Thus, cultural differences in interpreting human gesture may be applicable to gestures in graphical representations such as avatars and pictograms.

Another reason for decreased recognition accuracy of the use of gestures and gesture marks in other countries may result in the unique comic culture in Japan, given the fact that avatars were drawn using Japanese comic drawing style. The comic culture has been developed in relative isolation in Japan [22], has grown in a different way than other countries, by creating new drawing techniques, i.e., slashing gesture line style to express motion and dynamics, collages of faces and symbolic expressionistic effects, word-picture linkage, and iconic characters instead of caricatures) [23]. Example of such unique Japanese iconic expressions are "frames to express anger", and "balloon from a nostril to express sleeping". The avatar designs used in the experiment were designed by three Japanese designers to limit the expressers from one country, thus have unique comic drawing styles that is interpreted accurately only by Japanese, and lead to various interpretations from other countries. Further study should be done to evaluate avatars designed by artists of other cultures, e.g., European or American.

5 Summary

This paper compared the interpretations of avatars' facial expressions among five Asian countries and analyzed the design features that would lead to misinterpretation of avatars' facial expression. The goal of this experiment is: 1) to investigate cultural differences in avatar expression evaluation and apply findings from psychological study in human facial expression recognition within Asian countries, 2) to identify design features that cause cultural differences in avatar facial expression interpretation.

The results suggested that 1) there are cultural differences in interpreting avatars' facial expressions among Asian countries, and the psychological theory that suggests physical proximity affects facial expression recognition accuracy is also applicable to avatar facial expressions, 2) use of gestures and gesture marks may sometimes cause counter-effects in recognizing avatar facial expression. Using gesture and gesture marks increase the recognition accuracy of the expresser's country, while other countries' recognition accuracy was not increased when compared to the one to the designs that uses facial expression only. This indicates that we have to be careful in adding a gesture or mark in designing avatars and avatar facial expression.

Acknowledgements. This research was supported by a Grant-in-Aid for Scientific Research (A) (15200012, 2003-2005) from the Japan Society for the Promotion of Science (JSPS). The research was conducted under a supervision of Professor Toru Ishida at the Department of Informatics, Kyoto University. The experiment was conducted as a part of ICE2005 (Intercultural Collaboration Experiment 2005). I am truly grateful to Professor Ishida for giving me the opportunity to participate in ICE2005, and giving me sheer insights and suggestions. The author truly appreciates the effort of Dr. Naomi Yamashita at NTT Communication Science Laboratories for coordinating the experiments. We are grateful to the participants from Shanghai Jiao Tong University (China), Southwest China Normal University (China), Thai Computational Linguistics Laboratory, NICT (Thailand), Kyung Hee University (Korea), University of Malaya (Malaysia), Wakayama University (Japan) and Kyoto University (Japan).

References

1. MSN Messenger: http://www.msn.com
2. Yahoo! Messenger: http://messenger.yahoo.com/
3. Kurlander, D., Skelly, T., Salesin, D.: Comic Chat. In: Proceedings of Computer Graphics and Interactive Techniques, pp. 225–236. ACM Press, New York (1996)
4. Second life: http://secondlife.com/
5. Damer, B.: Avatars: Exploring and Building Virtual Worlds on the Internet. Peachpit Press, Berkeley (1997)
6. Smith, M.A., Farnham, S.D., Drucker, S.M.: The Social Life of Small Graphical Chat Spaces. In: Proceedings of CHI, pp. 462–469. ACM Press, New York (2000)
7. Pesson, P.: ExMS: an Animated and Avatar-based Messaging System for Expressive peer Communication. In: Proceedings of GROUP, pp. 31–39. ACM Press, New York (2003)
8. Koda, T.: Interpretation of Expressive Characters in an Intercultural Communication. In: Negoita, M.G., Howlett, R.J., Jain, L.C. (eds.) KES 2004. LNCS (LNAI), vol. 3214, pp. 862–868. Springer, Heidelberg (2004)
9. Isbister, K., Nakanishi, H., Ishida, T.: Helper Agent: Designing and Assistant for Human-Human Interaction in a Virtual Meeting Space. In: Proceedings of Human Factors in Computing Systems (CHI2000), pp. 57–64. ACM Press, New York (2000)
10. Bartneck, C., Takahashi, T., Katagiri, Y.: Cross Cultural Study of Expressive?Avatars. In: Proceedings of the Social Intelligence Design (2004)
11. Koda, T., Ishida, T.: Cross-cultural study of avatar expression?interpretation. In: International Symposium on Applications and the Internet (SAINT) (2006)
12. Ekman, P.: Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life. Henry Holt and Company (2003)
13. Elfenbein, H.A., Ambady, N.A: Cultural similarity's consequences: A distance perspective on cross-cultural differences in emotion recognition. Journal of Cross-Cultural Psychology 34, 92–110 (2003)

14. Ekman, P.: About Brows: Emotional and Conversational Signals. In: Cranach, M.V., Foppa, K., Lepenies, W., Plog, D. (eds.) Human Ethology: Claims and Limits of a New Dicipline: Contributions to the Colloquium, pp. 163–202. Cambridge University Press, Cambridge (1979)
15. Elfenbein, H.A., Ambady, N.: On the Universality and Cultural Specificity of Emotion Recognition: A Meta-Analysis. Psychological Bulletin, American Psychological Association, Inc. 128(2), 203–235 (2002)
16. Ortony, A., Clore, G.L., Collins, A.: The Cognitive Structure of Emotions. Cambridge Univ. Press, Cambridge (1998)
17. Cho, H., Inaba, R., Takasaki, T., Mori, Y., Ishida, T.: Culturally-Situated Pictogram Retrieval. In: Fellbaum, C., Vossen, P.T.J.M.: Connecting the Universal to the Specific: Towards the Global Grid. In: Ishida, T., Fussell, S.R., Vossen, P.T.J.M. (eds.) IWIC 2007. LNCS, vol. 4568, Springer, Heidelberg (2007)
18. Mori, Y.: Atoms of bonding: Communication Components Bridging Children Worldwide. In: Fellbaum, C., Vossen, P.T.J.M.: Connecting the Universal to the Specific: Towards the Global Grid. In: Ishida, T., Fussell, S.R., Vossen, P.T.J.M. (eds.) IWIC 2007. LNCS, vol. 4568, Springer, Heidelberg (2007)
19. Takasaki, T., Moir, Y.: Design and Development of a Pictogram Communication System for Children around the World. In: Fellbaum, C., Vossen, P.T.J.M.: Connecting the Universal to the Specific: Towards the Global Grid. In: Ishida, T., Fussell, S.R., Vossen, P.T.J.M. (eds.) IWIC 2007. LNCS, vol. 4568, Springer, Heidelberg (2007)
20. Efron, D.: Gesture and Environment. King's Crown Press, New York (1941)
21. Ekman, P., Friesen, W.: The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding. Semiotica 1, 48–98 (1969)
22. McLoud, S.: Reinventing Comics: How Imagination and Technology Are Revolutionizing an Art Form, vol. 118, pp. 122–123. Perennial, New York (2000)
23. McLoud, S.: Understanding Comics: The Invisible Art, vol. 210, pp. 131–133. Harper Perennial, New York (1993)

Culturally-Situated Pictogram Retrieval

Heeryon Cho[1], Toru Ishida[1], Naomi Yamashita[2], Rieko Inaba[3], Yumiko Mori[4], and Tomoko Koda[5]

[1] Department of Social Informatics, Kyoto University, Kyoto 606-8501, Japan
cho@ai.soc.i.kyoto-u.ac.jp, ishida@i.kyoto-u.ac.jp
[2] Media Interaction Principle Open Laboratory,
NTT Communication Science Laboratories, Kyoto 619-0237, Japan
naomi@cslab.kecl.ntt.co.jp
[3] Language Grid Project, National Institute of Information and Communication Technology (NICT), Kyoto 619-0289, Japan
rieko.inaba@nict.go.jp
[4] Kyoto R&D Center, NPO Pangaea, Kyoto 600-8411, Japan
yumi@pangaean.org
[5] Faculty of Information Science and Technology,
Osaka Institute of Technology, Osaka 573-0196, Japan
koda@is.oit.ac.jp

Abstract. This paper studies the patterns of cultural differences observed in pictogram interpretation. We conducted a 14-month online survey in the U.S. and Japan to ask the meaning of 120 pictograms used in a pictogram communication system. A total of 935 respondents in the U.S. and 543 respondents in Japan participated in the survey to submit pictogram interpretations which added up to compose an average of 147 English interpretations and 97 Japanese interpretations per pictogram. Three human judges independently analyzed the English–Japanese pictogram interpretation words, and as a result, 19 pictograms were found to have culturally different interpretations by two or more judges. The following patterns of cultural differences in pictogram interpretation were observed: (1) two cultures share the same underlying concept, but have different perspectives on the concept, (2) two cultures only partially share the same underlying concept, and (3) two cultures do not share any common underlying concept.

Keywords: pictogram, interpretation, analysis, cultural difference.

1 Introduction

Hand drawn images have long been used to convey messages, and are still being used as an effective iconic medium of representation. For instance, prehistoric drawings inside the Altamira Cave[1] tell us what wild animals lived during the ice age. Outlines of walking or standing human figures on the surface of a pedestrian traffic light alert us when to proceed or to stop.

[1] http://whc.unesco.org/en/list/310

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 221–235, 2007.

Hand drawn images are in essence iconic representation carrying semantic interpretation. We will call such images *pictograms* in this paper. One of the most familiar pictograms used nowadays are universal signs such as road signs, direction boards at the airports, and symbols of sports played in the Olympics. These pictograms are intended to convey particular information to a wide range of audiences. Much effort has been put on developing pictograms for AAC (Augmentative and Alternative Communication). AAC assists people with severe communication disabilities to be more socially active in interpersonal interaction, education, employment, and community activities. Sign language and Braille are good examples of AAC. Blissymbolics[1] and PIC[2] are some pictogram communication systems used in AAC.

In this paper, we look at a new kind of communication involving pictograms, one that exchanges pictogram messages via a network system[3,4,5]. A participant involved in pictogram message exchange creates a pictogram message by selecting and combining one or more pictograms which are registered to the system. Note that these registered pictograms are created by art major students who are novices at pictogram design.

Because pictograms have clear pictorial similarities with some object[6], pictogram communication has the potential to establish communication between participants speaking different languages. Successful pictogram communication, however, can be realized when the two participants share common pictogram interpretation. In the case of differing interpretation, misunderstanding may arise. In an intercultural communication setting where multilingual, multicultural users are involved, it would be beneficial if some intermediating system automatically detects and notifies the users of possible misunderstanding that might arise during message exchange. Such automatic detection, especially the detection of misconception attributable to users' linguistic or cultural differences, could help to establish mutual understanding and to facilitate communication among multilingual, multicultural users.

Various studies to support intercultural communication have been reported to date. [7] analyzed a large volume of multilingual BBS message log, and discovered that misunderstanding is likely to arise among different language speakers when there is a gap between the BBS message thread structure and the words used in the BBS messages. [8,9] conducted a large scale web experiment to reveal cultural differences in the interpretation of avatars' facial expressions. [10,11] proposed an infrastructure which supports composition of language services.

Here, we focus on culturally-situated pictogram retrieval, where a pictogram communication system user, situated in an intercultural communication setting, searches for relevant pictograms to compose a pictogram message. Retrieved pictograms are included in the pictogram message, and this message is sent to the conversational partner with different cultural background. Since pictograms used here are created by novices at pictogram design, each pictogram does not guarantee a single, clear interpretation: their interpretations may be various[12]. Consequently, multicultural users participating in pictogram communication may have varying, culture-specific interpretations of these pictograms.

Our goal is to notify the users of information regarding pictogram interpretation so that users creating pictogram messages can know in advance how certain pictograms are interpreted by members of different cultures. When this kind of notification is done during pictogram retrieval, it will allow the message creator to choose pictograms with discretion. This in turn will lead to the composition of more understandable pictogram messages.

To enable the notification of culture-specific pictograms, we first need to understand what kind of culture-specific pictogram interpretations exist. We do this by conducting an online survey, which asks the meaning of pictograms, to members of two different cultures: U.S. and Japan. Section 2 summarizes the U.S.–Japan online pictogram survey and reports the details of culture-specific pictograms found in the two countries. Section 3 discusses the findings, and Section 4 concludes this paper.

2 Cultural Ambiguity in Pictogram Interpretation

To understand how different cultures interpret pictograms, we conducted an online survey in the U.S. and Japan. The selection of the two countries is based on the fact that chances of finding cultural differences in pictogram interpretation would be higher if we choose cultures that have greater cultural differences. Since existing literatures on cross-cultural studies have found the two countries' cultures to be distinct in many aspects[13,14,15], we proceed with our survey in the two countries.

2.1 Pictogram Web Survey

Objective. An online pictogram survey was conducted to understand whether differences in pictogram interpretation exist in two countries, U.S. and Japan, and if so, what they are.

Method. A pictogram survey, which asks the meaning of 120 pictograms used in the system, was conducted to respondents in the U.S. and Japan via the WWW from October 1, 2005 to November 30, 2006.[2] Human respondents were shown a webpage containing 10 pictograms, and were asked to write the meaning of each pictogram inside the textbox provided below the pictogram. Each time a set of 10 pictograms was shown at random, and respondents could choose and answer as many question sets they liked. The maximum question sets a respondent could answer were 12 sets which contain a total of 120 pictograms.

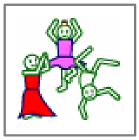
Data. A total of 543 respondents in Japan and 935 respondents in the U.S. participated in the survey. An average of 97 interpretations consisting of Japanese words or phrases (duplicate expressions included) and an average of 147 interpretations consisting of English words or phrases (duplicate expressions included) were collected for each pictogram. For each pictogram, unique interpretation

[2] The URL of the online pictogram survey is http://www.pangaean.org/iconsurvey/

words or phrases were listed for each language, and the occurrences of those unique words were counted to calculate the frequency.

An example of U.S.–Japanese word count result for one of the surveyed pictogram is shown in Table 1. The left two columns show interpretation words and frequencies collected from the U.S. respondents. The right two columns show interpretation words and frequencies collected from the Japanese respondents.

Table 1. U.S.–Japan interpretation words and frequencies for the below pictogram



Interpretations in U.S.	Freq.	Interpretations in Japan	Freq.
dancing	51	dance (*dansu*: kt)	45
dance	25	dance (*odori*: kj+hr)	13
gymnastics	7	dance (*odori*: hr)	6
dancers	6	dance (*dansu*: hr)	2
ballet	5	fun (*tanoshii*: kj+hr)	2
play	5	dance (*odoru*: hr)	1
cheerleaders	4	circus (*sa–kasu*: kt)	1
danceing	3	dancer (*dansa–*: kt)	1
playing	3	performance (*pafo–mansu*: kt)	1
family	2	clown (*piero*: kt)	1
friends	2	theatrical play (*engeki*: kj)	1
acrobatics	1	hobby (*shumi*: kj)	1
ballerina show	1	battle (*tatakai*: kj+hr)	1
cheerleading	1	gymnastics (*taisou*: kj)	1
cherrleaders	1	dance (*odoru*: kj+hr)	1
dance class	1	dance (*odori*: kj+hr), dance (*dansu*: kt)	1
dancing triplets	1	everyone getting along well (*minnanakayoku*: hr+kj+hr)	1
exercise	1	rhythmic sports gymnastics (*shintaisou*: kj)	1
flexable	1		
girls playing	1		
hurting eachother	1		
i like to dance	1		
play time	1		
playin	1		
Total Frequency	126	Total Frequency	81

For example, U.S. interpretation word "*dancing*" placed at the top has a frequency of "51". This means that fifty-one U.S. respondents wrote "*dancing*" as the meaning of the pictogram displayed at the top of the table. Comparative charts that were created for analyses contain the original Japanese words as they are, but in this paper we translate all Japanese words into English for readability. A Japanese-English dictionary, EDICT[3], was used for translation. Words and phrases which were not listed in the dictionary (including colloquial expressions) were translated by humans. Parentheses following each English translation of the Japanese word in Table 1 contain the original Japanese word expressed in alphabet (in *italics*) and the Japanese character construction of the original term: "hr" denotes *hiragana*, "kt" denotes *katakana*, and "kj" denotes *kanji*. Italicized Japanese term and its character construction are delimited by a colon(:).

Analysis. Tables comparing English and Japanese pictogram interpretation words and frequencies (similar to Table 1, but containing the original Japanese words and phrases) were created for each of the 120 surveyed pictograms. To determine whether culture-specific interpretations were present, three human judges independently analyzed the 120 English–Japanese pictogram interpretations for cultural differences. Two judges were Japanese and one judge was Korean. All three judges had college level Japanese and English proficiency. After reviewing the 120 pictogram interpretation words, each of the three judges found 8, 21(Korean judge), and 28 pictograms to have culturally different interpretations. 19 pictograms were found to have culturally different interpretations by two or more judges. 7 pictograms were found to have culturally different interpretations by all three judges.

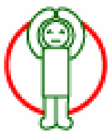
2.2 Result

We give details of the 19 pictograms which were judged by two or more judges to have culturally different interpretations by the U.S. respondents and Japanese respondents. To guide our explanation, we divide the pictograms into the following groups: (i) *gesture*, (ii) *gender and color*, (iii) *time*, (iv) *space*, (v) *familiar scene*, and (vi) *facial expression*. The top five frequently occurring U.S.–Japan pictogram interpretation words are listed for each pictogram along with their percentages. The percentage(Pct, %) of each word or phrase is calculated by dividing the interpretation word frequency with the total frequency. For example, the percentage of the word "*dancing*" in Table 1 can be calculated as $(51/126) * 100 = 40.48\%$. For all Japanese interpretation words, English translations, alphabetical expressions of the Japanese terms (in *italics*), and Japanese character constructions are provided as those shown in Table 1.

Gesture. Pictogram of a person holding up one's hands above one's head to form a circle-like shape (Table 2 top) was interpreted as "*exercise, jump rope, exercising, yoga, dance, stretch*" by a majority of the U.S. respondents whereas a majority of the Japanese respondents interpreted it as "*OK, circle, correct, all right, bingo.*" U.S. interpretations center on exercise-related concept while Japanese interpretations center on agreement-related concept.

[3] http://www.csse.monash.edu.au/~jwb/j_edict.html

Table 2. Gesture

Pictogram	U.S.	Pct(%)	Japan	Pct(%)
	exercise	19.47	OK (originally submitted in alphabet)	18.81
	jump rope	5.31	circle (*maru*: hr)	9.90
	happy	4.43	correct (*seikai*: kj)	8.91
	exercising	3.54	O.K. (*okke–*: kt)	7.92
	yoga	3.54	O.K. (*iiyo*: hr)	5.94
	mad	31.90	no good (*dame*: hr)	26.73
	angry	30.17	no good (*dame*: kt)	11.88
	no	4.31	wrong* (*batsu*: hr)	5.94
	stubborn	3.45	wrong* (*batsu*: kt)	3.96
	anger	1.72	no (*iie*: hr)	2.97
	talking	10.19	thank you (*arigatou*: hr)	6.33
	praying	9.55	please (*onegai*: hr+kj+hr)	6.33
	thinking	8.28	to speak (*hanasu*: kj+hr)	5.06
	speaking	5.10	soliloquy (*hitorigoto*: hr)	3.80
	lonely	3.19	soliloquy (*hitorigoto*: kj+hr+kj)	3.80

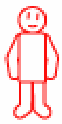
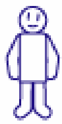
Likewise, pictogram of a person holding up one's arms to form an "X" (Table 2 middle) was interpreted as "*mad, angry, anger, frustrated, upset*" by a majority of the U.S. respondents whereas a majority of the Japanese respondents interpreted it as "*no good, wrong*[4], *no, miss, don't*." U.S. interpretations revolve around the concept which deals with negative emotions while Japanese interpretations revolve around the concept which deals with prohibition or criticism.

As for the pictogram that shows a standing person placing hands together while a speech balloon hangs next to the head (Table 2 bottom), approximately 40% of both the U.S. and Japanese respondents interpreted it as some kind of a speech act ("*talking, speaking*" and "*to speak, soliloquy*" respectively). At the same time, however, 14.6% of U.S. respondents interpreted it as "*praying, pray, prayer*" while 17.7% of Japanese respondents interpreted it as "*thank you, please*." These differences in the interpretations of the three pictograms, we think, are due to the differences in how gestures are interpreted in the U.S. and Japan. The body gestures expressing a circle or a cross are gestures well-recognized in Japan which respectively indicate that something is correct or wrong. However, such gesture is not recognized in the United States: therefore we suppose that the circle depicted in the pictogram was perceived as an expression of motion while the "X" was perceived as crossing of one's arms (hence, the stubborn or angry gesture) by the U.S. respondents.

The important thing to note is that while the two countries' overall interpretations of the top and middle pictogram differ greatly, the bottom pictogram contains a mixture of both the differing interpretations ("*praying*" vs. "*thank you*") and the common interpretation shared by the two countries ("*talking*" and "*to speak*").

[4] Although EDICT lists three entries to the Japanese term "*batsu*," English translation fitting the context of the pictogram (the "X" gesture) could not be found so a more appropriate human translation is given.

Table 3. Gender and color (top in red, bottom in blue)

Pictogram	U.S.	Pct(%)	Japan	Pct(%)
	woman	29.05	woman (*onnanohito*: kj+hr+kj)	28.00
	man	11.49	woman (*josei*: kj)	27.00
	mom	8.78	woman (*onna*: kj)	10.00
	dad	7.43	mother (*okaasan*: hr)	5.00
	adult	5.41	mother (*okaasan*: hr+kj+hr)	3.00
	man	34.23	man (*otokonohito*: kj+hr+kj)	27.72
	dad	10.07	male (*dansei*: kj)	26.73
	woman	8.05	man (*otoko*: kj)	9.90
	adult	6.04	father (*otousan*: hr)	3.96
	mom	5.37	father (*otousan*: hr+kj+hr)	3.96

Gender and Color. "The color red denotes women and the color blue denotes men" is a prevalent notion in Japan, but it is not so in the U.S. as indicated by the pictogram interpretations.

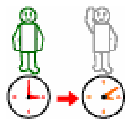
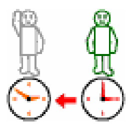
While 92% of the Japanese respondents interpreted the red human figure (Table 3 top) as "*woman, mother, adult female, sister, girl*" which all contain the female gender concept, 31.8% of the U.S. respondents interpreted it as "*man, dad, father, boy, male*" which all contain the male gender concept. Strong agreement in interpretation was reached by the Japanese respondents, but not by the U.S. respondents. The remaining 8% of the Japanese interpretations consisted of "*adult*" and "*person*" which lacks any gender concept, and "*boy*" that was answered by one Japanese respondent. As for the remaining U.S. respondents, most of them interpreted the red human figure similarly as the majority of Japanese did as some kind of a female person. Small portion of the U.S. respondents interpreted it as "*adult, person, teenager, grown up, parent.*"

Likewise, the blue human figure (Table 3 bottom) was interpreted by 93% of the Japanese respondents as "*man, male, father, adult male, brother, boy*" which all contain the male gender concept. In contrast, 20.8% of the U.S. respondents interpreted it as some person with the female gender, i.e. "*woman, mom, big girl, female, old women.*" Only one Japanese respondent interpreted it as a "*girl.*" The remaining U.S. respondents interpreted the blue human figure similarly as the Japanese as some kind of a person with the male gender.

In sum, it can be concluded that the correlation of color and gender (red denotes female and blue denotes male) is evident in the Japanese interpretations, but not in the U.S. interpretations. The important thing to notice is that while the Japanese interpretations center on a single gender concept, i.e. the concept of either male or female, the U.S. interpretations include both gender concepts for each pictogram leading to a greater ambiguity in interpretation.

Time. In both countries, pictograms containing clock image(s) (Table 4) were interpreted as some kind of a concept relating to time, but the first ranking interpretations were different between the two countries.

Table 4. Time

Pictogram	U.S.	Pct(%)	Japan	Pct(%)
	late	11.38	the future (*mirai*: kj)	16.83
	time	10.18	10 minutes later (*juppungo*: num+kj)	9.90
	10 minutes	3.59	the future (*mirai*: hr)	5.94
	later	2.99	afterwards (*atode*: hr)	3.96
	future	2.40	time passes (*jikangasusumu*: kj+hr+kj+hr)	2.97
	on time	16.87	now (*genzai*: kj)	11.11
	time	12.65	now (*ima*: kj)	10.00
	now	3.61	time (*jikan*: kj)	6.67
	what time is it	3.61	now (*ima*: hr)	4.44
	clock	3.01	time (*jikan*: hr)	4.44
	early	12.27	the past (*kako*: kj)	18.09
	before	4.29	10 minutes ago (*juppunmae*: num+kj)	11.70
	past	3.68	the past (*kako*: hr)	7.45
	time	3.68	time is turned back (*jikangamodoru*: kj+hr+kj+hr)	4.26
	late	3.07	some time ago (*sakki*: hr)	3.19

Starting with the top pictogram in Table 4, the first ranking U.S. interpretation was "*late* (11.38%)" whereas the first raking Japanese interpretation was "*the future* (16.83%)." Since the third ranking Japanese interpretation shown in the table is also "*the future* (5.94%)," it can be combined with the first ranking interpretation to yield a total percentage of 22.77%. Similar interpretations to the first ranking U.S. interpretation ("*late*") also existed below the ranking, which include "*5 min. late, late or later, you are late*" which were each answered by one U.S. respondent. A total of 4.19% U.S. interpretations contained interpretations similar to the first ranking Japanese interpretation: they were "*future, forward in time, past to future*." A total of 5% of Japanese interpretations contained interpretations similar to the first ranking U.S. interpretation: they were "*lateness* (*chikoku*: hr, kj), *to be late* (*okureru*: kj+hr)." Common interpretations shared by the two countries included "*10 minutes later, time passes*."

As for the middle pictogram in Table 4, the first ranking U.S. and Japanese interpretations were "*on time* (16.87%)" and "*now* (11.11%)" respectively. Since the second and fourth ranking Japanese interpretations shown in Table 4 are also "*now* (10.00% and 4.44%)," they can be combined to yield a total percentage of 25.55%. Similar interpretations to the first ranking U.S. interpretation ("*on time*") existed below the ranking, which include "*be on time, you are on time*." A total of 6.63% U.S. interpretations contained interpretations similar to the first ranking Japanese interpretation: they were "*now, present, current time, present time*." A total of 5.56% Japanese interpretations contained interpretations similar to the first ranking U.S. interpretation: they were "*just* (*choudo*: hr), *on time* (*ontaimu*: kt, *jikandoori*: kj+kt)." Common interpretation shared by the two countries was "*time*."

For the bottom pictogram in Table 4, the first ranking U.S. and Japanese interpretations were "*early* (12.27%)" and "*the past* (18.09%)" respectively. Since the third ranking Japanese interpretation shown in the table is also "*the past* (7.45%)," it can be combined to yield a total percentage of 25.54%. Similar interpretations to the first ranking U.S. interpretation ("*early*") existed below the ranking which include "*10 minutes early, 5 min early, early or earlier, someone's early, you are early*." A total of 4.91% U.S. interpretations contained interpretations similar to the first ranking Japanese interpretation: they were "*past, backward in time, future to past*." A total of 2.11% Japanese interpretations contained interpretations similar to the first ranking U.S. interpretation: they were "*arrived early* (*hayakutsuichatta*: kj+hr+kj+hr), *arrived 10 minutes ago* (*juppunmaenikimashita*: num+kj+hr)." Common interpretation shared by the two countries was "*10 minutes ago*."

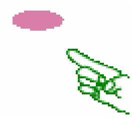
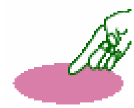
In sum, the three pictograms containing clock image(s) were interpreted by the U.S. respondents as "*late, on time, early*" whereas the Japanese respondents interpreted them as "*future, present, past*." It can be said that the U.S. interpretations deal with a concept of appointment in relation to time while the Japanese interpretations deal with temporal relations along the time axis. The important thing to notice is that the basic time concept is shared by the two countries, but the detailed interpretations that unfold around the time concept differs as manifested by the two countries' first ranking interpretations.

Space. Two pictograms portraying an index finger pointing to a specific place were interpreted differently by the two countries. Although both countries' interpretations revolved around the concept of space, the perspectives held by the respondents were different. We focus on the first ranking interpretations (as we did in the prior time related pictograms) to highlight the differences.

For the Table 5 top pictogram, 18.88% of the U.S. respondents interpreted the finger's direction to be pointing "*up*" whereas 30.63% of the Japanese respondents interpreted it as pointing to "*there*." Since the third ranking Japanese interpretation shown in the table is also "*there* (9.91%)," it can be combined with the first ranking Japanese interpretation to yield a total percentage of 40.54%. Similar interpretations to the first ranking U.S. interpretation ("*up*") existed below the ranking which add up to 11.89%: example interpretations include "*high, pointing up, up/above, above, look up, up high*." Combining the first ranking U.S. interpretations with the similar, below ranking interpretations, the percentage of the major U.S. interpretation "*up*" adds up to 30.77%.

For the Table 5 bottom pictogram, 18.88% of the U.S. respondents interpreted it as "*down*" whereas 36.04% of the Japanese respondents interpreted it as "*here*." Since the third ranking Japanese interpretation, "*this direction* (5.41%)," shown in the table contains similar meaning to the first ranking Japanese interpretation, it can be combined to yield a total percentage of 41.45%. Similar interpretations to the first ranking U.S. interpretation ("*down*") existed below the ranking which add up to 20.98%: example interpretations include "*low, pointing down, down/below, below, look down, down low*." Combining the first ranking

Table 5. Space

Pictogram	U.S.	Pct(%)	Japan	Pct(%)
	up	18.88	there (*asoko*: hr)	30.63
	there	14.69	that (*are*: hr)	27.03
	far	6.29	there (*acchi*: hr)	9.91
	over there	4.90	above (*ue*: kj)	5.41
	point	4.90	far (*tooi*: kj+hr)	3.60
	down	18.88	here (*koko*: hr)	36.04
	here	16.08	this (*kore*: hr)	26.13
	near	6.99	this direction (*kocchi*: hr)	5.41
	big	3.50	below (*shita*: kj)	5.41
	low	3.50	near (*chikai*: kj+hr)	3.60

U.S. interpretations with the similar, below ranking interpretations, the percentage of the major U.S. interpretation "*down*" adds up to 39.86%.

Major interpretation observed in one country was also observed in the other country, but with a lower percentage. The first ranking Japanese interpretations "*there*" and "*here*" for the top and bottom pictogram (Table 5) were also observed within the U.S. interpretations (totaled 24.48% and 25.87% respectively): example interpretations include "*there, over there, go there, look there, spot there*" and "*here, right here, come here, look here, spot here.*" On the other hand, Japanese interpretations similar to the first ranking U.S. interpretations "*up*" and "*down*" (top and bottom pictogram in Table 5) were totaled 9% and 8.11% respectively: example interpretations include "*above* (*ue*: kj, hr), *high* (*takai*: kj+hr)" and "*below* (*shita*: kj, hr), *low* (*hikui*: kj+hr)." Common interpretations shared by the two countries were "*far*" and "*near*" respectively for the top and bottom pictogram in Table 5.

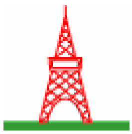
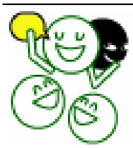
In sum, the two pictograms depicting a finger pointing to a certain direction were interpreted as "*up*, *down*" by the U.S. respondents whereas the Japanese respondents interpreted them as "*there*, *here*." It can be said that the U.S. interpretations contain a vertical perspective of space while the Japanese interpretations contain a horizontal perspective of space. The important thing to notice is that while the basic concept of space is shared by the two countries, the major (or the first ranking) interpretations vary as evidenced by "*up* vs. *there*" and "*down* vs. *here*."

Familiar Scene. In some cases, the U.S. and Japanese respondents recalled familiar scenes from the visual scenery depicted in the pictograms. These recalled scenes varied according to culture.

In the case of the top pictogram in Table 6, nearly half (43.08%)[5] of the U.S. respondents interpreted the red tower as the "*Eiffel Tower*" while nearly half (47.83%) of the Japanese respondents interpreted it as the "*Tokyo Tower.*"

[5] Twelve misspelled versions of the "Eiffel" were observed in the U.S. interpretations including the fourth ranking "*Eifel Tower.*"

Table 6. Familiar scene

Pictogram	U.S.	Pct(%)	Japan	Pct(%)
	Eiffel Tower	19.23	Tokyo Tower (*toukyoutawa–*: kj+kt)	44.57
	paris	19.23	tower (*tawa–*: kt)	23.91
	tower	15.38	tower (*tou*: kj)	7.61
	Eifel Tower	4.62	Eiffel Tower (*efferutou*: kt+kj)	6.52
	france	4.62	tower (*denpatou*: kj)	3.26
	winner	30.63	athletic meet (*undoukai*: kj)	36.59
	winning	6.88	number one (*ichiban*: kj)	8.54
	champion	5.63	overall victory (*yuushou*: kj)	6.88
	first place	5.00	number one (*ichiban*: hr)	3.66
	cheering	3.13	first place prize (*ittoushou*: kj)	3.66
	friends	9.38	liar (*usotsuki*: hr)	7.89
	party	8.13	to tell a lie (*usowotsuku*: hr)	5.26
	gossip	3.75	lie (*uso*: kj)	3.95
	happy	3.13	lie (*uso*: hr)	2.63
	happy group	3.13	malicious gossip (*kageguchi*: hr)	2.63

Apparently, the respondents recalled specific instances of the tower they were familiar with. None of the U.S. respondents submitted "*Tokyo Tower*" as the interpretation, but 7.6% of the Japanese respondents submitted "*Eiffel Tower*" as the interpretation. Common interpretation shared by the two countries was "*tower*."

In the case of the middle pictogram in Table 6, the first ranking interpretations given by the U.S. and Japanese respondents were "*winner* (30.63%)" and "*athletic meet* (36.59%)" respectively. Note that such athletic meet depicted in the pictogram is a regularly held school event in Japan. Therefore, it is reasonable to assume that the Japanese respondents associated the pictogram's visual scenery to the school hosted athletic meet. Similar interpretations shared by the two countries (U.S. / Japan) were "*winning* / *overall victory*" and "*champion, first place* / *number one*."

The case with Table 6 bottom pictogram should be given greater attention since the two countries' interpretations vary greatly, almost going the opposite direction. While most of the U.S. respondents interpreted the pictogram to mean "*friends, party, happy, happy group, laughing, having fun*, etc." which all indicate a cheery, positive scene, most of the Japanese respondents interpreted it to mean "*liar, to tell a lie, lie, malicious gossip, split personality, vicious, to deceive, scheming*, etc." which all indicate a shadowy, negative image. We assume that the Japanese respondents have interpreted the black face on the upper right corner as a person having a malicious intent or an ulterior motive. Hence, the negative interpretation is given.

In contrast, we assume that the U.S. respondents interpreted the black face to be an African American, and as a result, interpreted the four faces as a group of people with varying ethnic background. Since people from diverse ethnic groups are portrayed as chatting together, it is a desirable scene, and thus positive interpretations are derived. Such interpretation, however, may be difficult to come

Table 7. Facial expression

Pictogram	U.S.	Pct(%)	Japan	Pct(%)
	whistling	13.21	feigning ignorance (*shiranpuri*: hr)	5.06
	whistle	10.06	to be peevish (*suneru*: hr)	5.06
	no	5.66	hmm (*hun*: hr)	5.06
	annoyed	2.52	acting rudely and suddenly (*pui*: hr)	5.06
	ignore	2.52	whistle (*kuchibue*: kj)	5.06
	scared	18.01	cold (*samui*: kj+hr)	27.18
	cold	10.56	cold (*samui*: hr)	23.30
	worried	10.56	scary (*kowai*: hr)	9.71
	nervous	9.94	scary (*kowai*: kj+hr)	4.85
	sad	9.94	trembling (*buruburu*: hr)	2.91
	happy	6.49	good-looking (*kakkoii*: hr)	31.7
	mean	5.19	handsome (*hansamu*: kt)	8.65
	smart	4.55	boast (*jiman*: kj)	2.88
	boy	3.90	nice man (*iiotoko*: hr+kj)	1.92
	mischievous	3.90	ahem (*ehhen*: hr)	1.92
	happy	25.64	cute (*kawaii*: hr)	42.72
	girl	3.85	pretty (*kirei*: hr)	5.83
	nice	3.85	cute (*kawaii*: kj+hr)	2.91
	pretty	3.85	beautiful person (*bijin*: kj)	2.91
	sweet	3.85	chuckling (*ufufu*: hr)	1.94
	happy	8.05	pretty (*kirei*: hr)	16.49
	in love	4.70	beautiful person (*bijin*: kj)	13.40
	cute	4.03	cute (*kawaii*: hr)	8.25
	pretty	3.36	beautiful (*utsukushii*: kj+hr)	4.12
	sweet	3.36	a prim girl (*osumashi*: hr)	2.06
	sly	11.95	to make fun of (*bakanisuru*: hr)	3.00
	sneaky	11.32	bitter smile (*nigawarai*: kj+hr)	3.00
	happy	6.92	doubt (*utagai*: hr)	2.00
	cool	2.52	grinning (*niyaniya*: hr)	2.00
	shy	2.52	broadly grinning (*niyari*: hr)	2.00

out from Japanese respondents, since Japan is an ethnically homogeneous country, and almost all people (excluding the foreigners) belong to the same ethnic group. Therefore, it is more natural to interpret a different face color as signifying the person's state of mind. The important thing to mention with regard to the three pictograms dealing with familiar scenery is that they contain a mixture of different interpretation patterns: while the top and middle pictogram respectively contains a common underlying concept such as "*tower*" and "*winning*," the bottom pictogram contains vastly varying interpretations.

Facial Expression. Facial expressions were interpreted differently not only between the two countries, but also among the respondents within the same country. Starting with the top pictogram in Table 7, the greatest common U.S. interpretation was "*whistling, whistle* (25.16%)" whereas the greatest common Japanese interpretation was "*feigning ignorance, pretending not to know* (30.38%)." Other varying interpretations were given by the members of each country. For instance, U.S. respondents interpreted as "*curious, kiss, relieved, sad, startled, embarrassed, snobby*" while Japanese respondents interpreted as "*to pout, to deceive, boring, to jeer, to get angry, to tell a lie, to bluff.*"

As for the second pictogram in Table 7, the top two interpretations in the U.S. and Japan were "*scared* (24.84%), *cold* (11.8%)" and "*cold* (60.95%), *scared* (24.27%)" respectively. Notice that although both countries share the same two interpretations "*cold, scared*," the first and second commonly shared interpretations are reversed between the two countries.

Moving to the third, fourth, and fifth pictogram in Table 7, the first ranking interpretations for each of the three pictograms were "*happy, happy, happy*" by the U.S. respondents, and "*good-looking, cute, pretty*" by the Japanese respondents. Japanese respondents tend to interpret the outer appearance of the face while U.S. respondents interpreted the state of the mind projected through the face. The fourth pictogram had, compared to the other two pictograms, a relatively high agreement in interpretation within each country with 25.64% answering "*happy*" in the U.S. and 42.72% answering "*cute*" in Japan. As for the remaining two pictograms (Table 7 third and fifth), low agreement on interpretation was reached especially among the U.S. respondents.

The last pictogram shown at the bottom of Table 7 consists of widely varying interpretations not only between the two countries, but also among the members within each country. It can mean "*sly, sneaky, happy, cool, shy*" in the U.S. while "*to make fun of, bitter, doubt, grinning, broadly grinning*" in Japan. However, most of the Japanese interpretations contained negative connotations whereas the U.S. interpretations contained both negative and positive connotations. For example, U.S. interpretations such as "*pleased, clever, glad, proud, smart*" were positive interpretations that never appeared in the Japanese interpretations.

In sum, pictograms containing facial expressions can have varying interpretations not only between the two countries, but also among the members of the same country. Note that although there were other pictograms that depicted facial expressions, for example, pictograms depicting a crying face or an angry face, these pictograms with negative facial expressions were interpreted similarly by the two countries. There were no cultural differences in the interpretations.

3 Discussion

We looked at the details of culturally different interpretations in 19 pictograms. Although each pictogram contained a specific cultural difference in interpretation, we think that these cultural differences can be categorized into the following three patterns. We give examples of each pattern.

- The basic concept captured by the two cultures are the same, but the perspectives on that concept are different.

  `E.g. The concept related to time, space, tower, face are captured by`

`both cultures (U.S. and Japan), but how they are perceived vary.`
`- [Table 4]` *late* `vs.` *future*, *on time* `vs.` *now*, *early* `vs.` *past*
`- [Table 5]` *up* `vs.` *there*, *down* `vs.` *here*
`- [Table 6 top]` *Eiffel Tower* `vs.` *Tokyo Tower*
`- [Table 7 third, fourth, fifth]` *happy* `vs.` *good-looking, cute, pretty*

– The basic concept(s) are only partially shared by the two cultures.
`- [Table 2 bottom]` *talking, to speak* `is shared, but not` *praying*, *thank you*
`- [Table 3]` *woman* `is shared, but not` *man* `and vice versa`
`- [Table 6 middle]` *winning, overall victory* `is shared, but not` *athletic meet*

– There is no common concept captured by the two cultures.
`E.g. A gesture is recognized by one culture, but not by the other.`
`- [Table 2 top & middle]` *exercise* `vs.` *O.K.*, *mad* `vs.` *no good*
`- [Table 7 top]` *whistle* `vs.` *feigning ignorance*
`E.g. Specific environment leads to specific recognition.`
`- [Table 6 bottom]` *friends* `vs.` *liar*

4 Conclusion

As a first step to understanding how pictograms are interpreted in different cultures, a pictogram web survey asking the meaning of pictograms was conducted in the U.S. and Japan. Three human judges independently analyzed the survey results containing English–Japanese pictogram interpretations to see whether cultural differences in pictogram interpretation exist between the two countries. As a result, 19 out of the 120 surveyed pictograms were judged by two or more human judges to have culturally different interpretations.

Analysis of the 19 culturally different pictogram interpretations confirmed the following three patterns of cultural differences in pictogram interpretation: (1) two cultures share the same underlying concept, but have different perspectives on the concept, (2) two cultures only partially share the same underlying concept, and (3) two cultures do not share any common underlying concept. These findings can be utilized in designing a pictogram retrieval system which can detect and notify the cultural differences in pictogram interpretation.

Acknowledgements. We are grateful to Satoshi Oyama (Department of Social Informatics, Kyoto University), Toshiyuki Takasaki (NPO Pangaea) and members of Ishida Laboratory for valuable discussions and comments.

References

1. Bliss, C.K.: Semantography (Blissymbolics). Semantography Press, Sydney (1965)
2. Maharaj, S.: Pictogram Ideogram Communication. In: Saskatoon, S.K.: The Pictogram Centre, Saskatchewan Association of Rehabilitation Centres (1980)

3. Mori, Y.: Atoms of Bonding: Communication Components Bridging Children Worldwide. In: Ishida, T., Fussell, S.R., Vossen, P.T.J.M. (eds.) IWIC 2007. LNCS, vol. 4568, Springer, Heidelberg (2007)
4. Takasaki, T.: PictNet: Semantic Infrastructure for Pictogram Communication. In: The 3rd International WordNet Conference (GWC-06), pp. 279–284 (2006)
5. Takasaki, T.: Design and Development of a Pictogram Communication System for Children around the World. In: Ishida, T., Fussell, S.R., Vossen, P.T.J.M. (eds.) IWIC 2007. LNCS, vol. 4568, Springer, Heidelberg (2007)
6. Marcus, A.: Icons, Symbols, and Signs: Visible Languages to Facilitate Communication. Interactions 10(3), 37–43 (2003)
7. Yamashita, N., Ishida, T.: Automatic Prediction of Misconceptions in Multilingual Computer-Mediated Communication. In: International Conference on Intelligent User Interfaces (IUI-06), pp. 62–69 (2006)
8. Koda, T., Ishida, T.: Cross-cultural Comparison of Interpretation of Avatars' Facial Expressions. In: IEEE/IPSJ Symposium on Applications and the Internet (SAINT-06), pp. 130–136 (2006)
9. Koda, T.: Cross-cultural Study of Avatars' Facial Expressions and Design Considerations within Asian Countries. In: Ishida, T., Fussell, S.R., Vossen, P.T.J.M. (eds.) IWIC 2007. LNCS, vol. 4568, Springer, Heidelberg (2007)
10. Ishida, T.: Language Grid: An Infrastructure for Intercultural Collaboration. In: IEEE/IPSJ Symposium on Applications and the Internet (SAINT-06), pp. 96–100 (2006)
11. Inaba, R., Murakami, Y., Nadamoto, A., Ishida, T.: Multilingual Communication Support Using the Language Grid. In: Ishida, T., Fussell, S.R., Vossen, P.T.J.M. (eds.) IWIC 2007. LNCS, vol. 4568, Springer, Heidelberg (2007)
12. Cho, H., Ishida, T., Inaba, R., Takasaki, T., Mori, Y.: Pictogram Retrieval Based on Collective Semantics. In: The 12th International Conference on Human-Computer Interaction (2007)
13. Hall, E.: Beyond Culture, pp. 85–103. Doubleday & Company, New York (1976)
14. Hofstede, G., Hofstede, G.J.: Cultures and Organizations: Software of the Mind. McGraw Hill, New York (2005)
15. Vatrapu, R., Suthers, D.: Culture and Computers: A Review of the Concept of Culture and Implications for Intercultural Collaborative Online Learning. In: Ishida, T., Fussell, S.R., Vossen, P.T.J.M. (eds.) IWIC 2007. LNCS, vol. 4568, Springer, Heidelberg (2007)

Emotion Eliciting Events in the Workplace: An Intercultural Comparison

Peter Schneider[1] and Axel Mattenklott[2]

[1] IAAEG, Universität Trier,
54286 Trier, Germany
peterschneider@iaaeg.de
[2] Psychologisches Institut der Johannes Gutenberg-Universität,
55099 Mainz, Germany
mattenklott@uni-mainz.de

Abstract. Different emotional experiences at the work place are evaluated in respect to their influence on job satisfaction. A sample of 75 Japanese employees and 169 German employees rated their emotional level following daily hassles in the work place that were attributed on the two dimensions: locus of causality and controllability. It was predicted that the same attribution pattern of daily hassles leads to different emotional responses and different levels of job satisfaction between employees with an interdependent and independent cultural background. Results indicate that equal attribution patterns of job related daily hassles lead to different emotional experiences between the two cultural groups and different levels of job satisfaction. It is argued that while emotions have a major influence on job satisfaction, this influence is culturally dependent. One element hereby is the explanation of the work related daily hassle.

Keywords: Daily hassles, attributions, emotions, job satisfaction, German and Japanese employees.

1 Introduction

The literature on organizational behavior is increasingly aware of different affective experiences in the multicultural workforce of internationally operating companies.

Although the affective experiences of employees receive more attention in the literature on organizational behavior [43, 7, 4], research of discrete work-related emotions and the nature of eliciting events is still scarce [3, 2]. For a long time job satisfaction has been considered as the main emotional reaction to the workplace [25]. Therefore interest in individual experiences in the work place is governed by research of job satisfaction. Though job satisfaction is considered to be an affective reaction to the work environment, little is known about the affective elements of job satisfaction, their antecedents and processing mechanisms and even less about their cultural tinge.

Possible cultural differences should come to the fore however, since employees collaborate more and more across national and cultural borders leaving a variety of challenges. The complexity of social relations for example can lead to a strain on

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 236–248, 2007.

human interactions far beyond language differences 22, 40]. Kitayama, Markus and Kurokawa (2000) for example have shown cultural dependency in the emotions experienced by interdependent (Japanese) and independent (American) students.

This paper contributes to the literature of job satisfaction and extends the existing literature on two points. Presuming cultural dependency for the elicitation of emotions, it examines the character of work events to elicit discrete emotions and in a second step the influence of these emotions on job satisfaction.

1.1 Emotions in the Concepts of Job Satisfaction

One of the concepts most often studied in the human resource literature is job satisfaction. Though there was much discussion of its predictive power [14] recent analyses have again highlighted its relationship to work related behaviors [17].

Research also indicates that job satisfaction is a global but yet culturally tinged concept, usually considered as an aggregated statement towards certain facets of the work situation [39]. Depending on their cultural background different patterns of job satisfaction emerge for different employees [10]. However up to this point only a few studies have examined cultural divergences in job satisfaction statements [13].

Recently more attention has been given to extended concepts of job satisfaction. Weiss and Cropanzano (1996) introduced "Affective Events Theory" (AET) as an alternative to the cognitive focus in the workplace. Based on their position that "...very little is known about the causes and consequences of true affective experiences in work settings" AET proposes that job satisfaction not only represents cognitive evaluations about certain facets of the work environment but that the work context is also saturated with feelings. Although already in 1992 Cranny, Smith and Stone refer to job satisfaction as: "…an affective (that is, emotional) reaction to a job, that results from the incumbent's comparison of actual outcomes with those that are desired (expected, deserved, and so on)", the emotional reactions to the work environment in respect to job satisfaction are still vaguely scrutinized. Job satisfaction is either itself considered an emotion or the emotional dimension of job satisfaction is considered to be reflected through affective personality traits [5, 16].

Yet the cognitive and the affective viewpoint are not contradictory but compliment each other. Different objects or situations in the work environment are evaluated on different levels. Some aspects, such as one's salary cause a cognitive examination of the job situation rather than an emotional response. Other aspects, such as being bullied by a colleague, elicit affective reactions rather than just cognitive reflections.

An additional advantage to separate between cognitive and affective levels are the behavioral consequences connected to it. Like intention to leave as a cognitive driven behavior or voluntarily helping colleagues as a affective driven behavior [9, 33].

1.2 Cultural Aspects of Work Related Emotions

Only a few studies assess job satisfaction in non-Western countries and describe different patterns of antecedents and consequences of job satisfaction usually known for western countries [38, 13].

Money and Graham (1999) report that changing one work facet important to American employees leads to a steep decline in work satisfaction when realized among Japanese employees. This decline is due to a different appraisal of the subjective work situation between the two groups. Hagihara, Tarumi, Babazono, Nobutomo and Morimoto, (1998) proposed that non-western cultures also show concern about certain job facets but that the weightings of the facets are different.

Distinguishing people for its cultural background is always a sensitive subject (for a good overall view see Vatrapu and Suthers, 2007). In our study we characterize different cultures for their concern for others and follow the principle by Markus and Kitayama (1991) who refer to cultures as interdependent and independent. In interdependent cultures, like Japan, people show a high concern and awareness for others. They reflect the individual in relation to their social environment. On the other hand side members of independent cultures like America or Germany are concerned about their own and see themselves as unique entities. Additionally in interdependent cultures the sustainment of harmony within groups is emphasized. They are concerned about avoiding negative situations, whereas for members of independent cultures the main goal is to strive for positive end states [12, 29].

For employees with an interdependent background the social group they belong to manifests cognitions and affective experiences in the workplace more than stable working conditions. Thus the concern for the behavior of others and the sustainment of harmony within the group leads to different emotional experiences and subsequently job satisfaction, depending on the cultural group [32].

1.3 Emotions and Their Elicitation

According to Morris (1989) the role of -especially negative- emotions in the work context deserves further attention since employers and organizations don't accept their expressions and even oppress them in the work environment.

However work life is riddled with small unpleasant events which are exogenous stimuli for emotional experiences in the work situation and therefore one major cause in the process of the formation of job satisfaction [37]. These events are then decoded according to the cultural background one belongs to [26, 28]. During the decoding procedure, the events are evaluated on only a few basic dimensions, particularly the explanations of their elicitation: who caused it and was it under volitional control [42, 36].

In contrast to the broad categories of work facets such as working conditions or pay, small unpleasant events are highly specific to the actual workplace and might not be alike even within the same occupation at the same employer [11, 8], nonetheless they are manifested through underlying dimensions which resemble each other with the main focus on the elicitor of the event and the event’s controllability [19, 38].

There is widespread agreement that many emotions are essentially universal [15] but cross-cultural differences of its origin remain. They are due to different explanations of the antecedents and consequences depending on the social embedding of the individual [27].

In independent cultures negative events caused by someone else (external) will lead to anger and frustration [42, 34] and negative events caused by oneself (internal) will lead to shame and guilt [35].

On the other hand side in interdependent cultures different attribution patterns exist. Externally caused negative events might lead to shame and guilt due to norm violation of maintaining harmony [24, 20].

1.4 Research Model and Hypotheses

In accordance with the Affective Events Theory (AET) by Weiss and Cropanzano (1996) we argue that within stable work conditions many random or arbitrary situations and events occur. Our model presumes that these work events will be evaluated according to the originator and the controllability. The same attribution pattern then leads to cultural dependent emotions and subsequently impacts job satisfaction (figure 1). In AET the nature of work events is not specified. Therefore in this study, daily hassles as a specific type of events is tested. The concept of daily hassles became introduced by Kanner, Coyne, Schaefer and Lazarus (1981) and refers to “irritating, frustrating, distressing demands that to some degree characterize everyday transaction with the environment”.

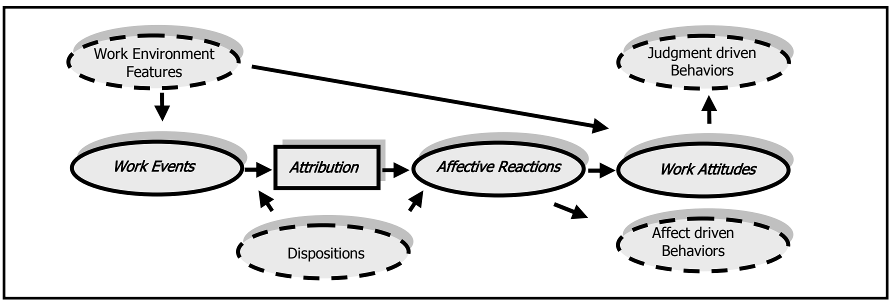


Fig. 1. Research model in accordance to AET [43]. The box “Attribution” refers to the study’s extension. The theoretical chain with solid line figures is tested. “Work events” are daily hassles; “Attribution” are attribution patterns; “Affective Reactions” are emotions; “Work Attitudes” is job satisfaction.

In our study we consider them as regularly appearing mild unpleasant events in the job. They are presumed to affect job related emotional experiences to a much higher degree than positive work events. Due to its high job dependence we did not rely on a set of pre-constructed events but had our participants construct daily hassles themselves according to their personal job experiences.

In contrast to former studies with job specific daily hassles as direct determinants of job satisfaction [11, 1], we consider their influence mediated through cultural specific emotions.

We had our participants construct daily hassles according to the locus of causality: oneself (internal) or another person (external) and controllability: under volitional control (controllable) or not under volitional control (not controllable). This leads to a 2x2 matrix of four different attribution patterns of daily hassles.

The hypotheses are as follows:

- The level of any emotion will be higher in case of a daily hassle than in a neutral situation (H1);

In case of work related daily hassles the same attribution patterns lead to different levels of job satisfaction.

- We predict that if an unpleasant event is caused externally and is not controllable, then frustration will be higher in independent cultures; shame and guilt will be higher in interdependent cultures (H1.1). Furthermore, the effect of frustration on job satisfaction will be higher for independent cultures while the effect of shame on job satisfaction will be higher for interdependent cultures due to its uncontrollability (H1.2).
- If an unpleasant event is caused externally and is controllable, then anger will be higher in independent cultures; shame and guilt will be higher in interdependent cultures (H2.1). The effect of anger on job satisfaction will be higher for independent cultures, while the effect of guilt on job satisfaction will be higher for interdependent cultures due to its controllability (H2.2).
- When an unpleasant event is caused internally and is controllable, shame and guilt will be the strongest emotion in both cultures with guilt being the strongest in interdependent cultures (H3.1). Only in interdependent cultures there will be an effect on job satisfaction due to its threat to group harmony (H3.2).
- When an unpleasant event is caused internally and is not controllable, guilt and shame will be the strongest emotions in both cultures with shame being the strongest in interdependent cultures (H4.1). Only in interdependent cultures there will be an effect on job satisfaction due to its threat to group harmony (H4.2).

2 Method

2.1 Sample

We recruited 169 German participants from three German service oriented medium-sized companies and 75 participants from one medium-sized electronic company located in Tokyo. The mean tenure was 6.8 years (SD=5.24 years) for the German sample, the mean for the Japanese sample was 11.4 years (SD=7.44 years). The mean tenure differed between the two groups ($t(239)=5.5$, $p<.01$).

46.2% in the German sample were male, 52.1% were female and 1.7% did not answer. In the Japanese sample 60% of the participants were male, 40% female. A chi-square test didn't show gender differences between the two samples. Participants' age was assessed through intervals: A chi-square test indicated a skewness towards older employees in the Japanese sample.

2.2 Questionnaire

We constructed a questionnaire and asked our respondents to remember modest unpleasant situations recently experienced in their work environment. We did not present identical work events or vignettes as emotional stimuli since they can easily be emotionally misinterpreted by people with different cultural backgrounds [23].

Rather situations had to be remembered by our participants according to one out of four variations: caused either by oneself (internal) or someone else (external) and the emergence of the event being under control or not. As a fifth variation we constructed a baseline and asked people to assess their usual work life without unpleasant situations. Finally our respondents replied to one of the four different attribution patterns in the work context or one neutral situation. Though they were free to decide about the daily hassle, all Japanese and all but eight German employees evaluated daily hassles elicited by colleagues or supervisors.

According to our hypotheses we assessed “frustration” and “anger” as well as "shame" and "guilt". Respondents rated the intensity of each specific emotion on a five-point rating scale ranging from "not at all" to "extremely". The English expressions were translated from English into German and back into English by two researchers. Discrepancies were resolved through discussion.

We assessed overall job satisfaction with a seven-point rating scale ranging from "very unsatisfied" to "very satisfied".

The questionnaires were all alike but each had one of four different attribution patterns or one neutral situation.

Questionnaires were translated from German into Japanese and back by two bilingual translators, one with German, one with Japanese mother tongue. Discrepancies were resolved after discussion among them.

Questionnaires were applied in two ways: one as an internet-based questionnaire and one as a paper-and pencil version. Paper and pencil versions of the questionnaires were administered in the Japanese sample by the supervisor and afterwards anonymously collected. The same procedure was carried out in one German company. In the two remaining companies the questionnaire was administered through the internet.

3 Results

First we assessed the impact of attributions on emotional reactions (table 1).

As posed in H1 a univariate analysis of variance yielded differences for the experience of emotions in unpleasant situations (versions 1 to 4) in relation to the neutral situation (version 5) for the independent sample.

There were no main effects for different emotions between unpleasant situations and the neutral situation in the Japanese sample. All German employees indicated mean differences between unpleasant situations and neutral situations except for the experience of guilt and shame in versions 1 and 2. Also there were mean differences for all emotions in the neutral situation between German and Japanese employees: T-Tests for independent samples in the neutral version indicate significantly higher

Table 1. Mean values and standard deviations of emotional reactions

Attri-bution		Emotion Frustration	Anger	Shame	Guilt
Version 1	G (N=52)	M=3.31* (1.13)	M=3.94* (0.96)	M=1.88 (1.16)	M=1.87 (1.22)
	J (N=15)	M=3.53 (0.74)	M=3.27 (0.96)	M=2.80 (1.08)	M=2.80 (1.21)
Version 2	G (N=35)	M=3.17* (0.96)	M=3.77* (1.40)	M=1.66 (1.11)	M=1.64 (0.95)
	J (N=15)	M=3.47 (0.83)	M=3.60 (0.91)	M=2.47 (1.47)	M=2.73 (1.33)
Version 3	G (N=25)	M=3.22* (1.28)	M=3.35* (1.27)	M=2.78* (1.51)	M=2.40* (1.35)
	J (N=15)	M=3.40 (1.18)	M=3.4 (0.99)	M=2.60 (1.24)	M=2.67 (1.29)
Version 4	G (N=25)	M=3.24* (1.13)	M=3.54* (0.98)	M=2.72* (1.43)	M=2.50* (1.58)
	J (N=15)	M=3.20 (1.01)	M=3.13 (0.99)	M=2.60 (1.24)	M=2.67 (1.29)
Version 5	G (N=30)	M=1.57 (1.07)	M=1.63 (1.10)	M=1.41 (0.95)	M=1.33 (0.92)
	J (N=15)	M=3.00 (1.00)	M=3.33 (0.98)	M=2.93 (1.33)	M=2.57 (1.40)

Mean values (M) and standard deviations (in parenthesis) for culture (G=Germany; J=Japan) and attribution; Version 1: external originator and not controllable; Version 2: external originator and controllable; Version 3: internal originator and controllable; Version 4: internal originator and not controllable; Version 5: neutral; *=difference towards neutral situations significant (p<.01).

levels in the Japanese group for all emotions: frustration (t(43)=4.32, p<.001); anger (t(43)=5.07, p<.001); shame (t(41)=4.30, (p<.001) and guilt (t(36)=3.30, p<.01).

Hypothesis 1.1: Consistent with our hypothesis in version 1 Japanese employees feel significantly higher shame (t(65)=2.97, p<.01) and guilt (t(36)=2.61, p<.01) than German employees. No differences were found for frustration yet German employees experienced higher anger (t(65)=-2.40, p<.05).

Hypothesis 2.1: Consistent with our hypothesis Japanese employees feel significantly higher shame (t(48)=2.18, p<.05) and guilt (t(38)=3.02, p<.01) than German employees though there is no difference in anger between the two samples.

Hypothesis 3.1 and 4.1: Contrary to our hypothesis we did not find guilt (H3.1) or shame (H4.1) to be the strongest emotional reaction. In both versions (3 and 4) the experience of frustration and anger is stronger for Japanese as well as German employees.

We then analyzed the emotional effect on job satisfaction regarding the cultural background and the attribution pattern (table 2).

Table 2. Effect of emotions on job satisfaction

Attri-bution		Emotion Frustration	Anger	Shame	Guilt
Version 1 Df (6,31)	G (N=52)	β = -1.086**		-	-
	J (N=15)	β = -.721*		-	-
Version 2 Df (6,33)	G (N=35)		β = -.788***	-	-
	J (N=15)		β = -.719**	β = 1.314**	β = -1.286**
Version 3 Df (4,25)	G (N=25)			-	-
	J (N=15)			-	β = -2.997*
Version 4 Df (4,28)	G (N=25)			-	β = -.929**
	J (N=15)			-	-

Df= degrees of freedom; G=Germany; J=Japan; β=standardized beta value; *=p<.1, **=p<.05; ***=p<.01; Version 1: external originator and not controllable; Version 2: external originator and controllable; Version 3: internal originator and controllable; Version 4: internal originator and not controllable.

To test H1.2 we analyzed the cultural pattern of frustration, shame and guilt on job satisfaction for externally caused, not controllable events. We conducted a linear regression analysis to estimate cultural effects of emotions on job satisfaction through interaction terms for culture and emotion. Culture: {1, if Germany; 0, if Japan}; β1 = (culture x frustration) = frustration, if Germany; 0, if Japan; β2 = ((1-culture) x frustration) = 0, if Germany; frustration, if Japan.

The interaction terms for shame and guilt were constructed in the same manner. Significant effects of frustration on job satisfaction were found for the German sample (T=-3.79, p<.001) and a modest effect in the Japanese sample (T=-1.82, p<.07). No effect for shame or guilt on job satisfaction was found for the Japanese sample.

Hypothesis 2.2: As for hypothesis 1.2 we tested the hypothesized effects through interaction terms for the relevant emotions. We found significant effects on job satisfaction for anger in the German sample (T=-2.92, p=<.01) and a modest effect in the Japanese sample (T=-2.18, p=<.05). Consistent with our hypothesis there was also a significant effect for guilt on job satisfaction (T=-1.98, p<.05) in the Japanese sample. Contrary to prediction and theory the Japanese sample exhibited a positive effect for shame (T=2.15, p<.05).

Hypothesis 3.2: Again we constructed interaction terms and tested the hypothesized effects for the relevant two emotions. According to our predictions results reveal a strong but modestly significant effect in the Japanese sample for guilt on job satisfaction (T=-1.80, p<.08).

Hypothesis 4.2: We tested the hypothesis through interaction terms as mentioned above and found an effect of guilt on job satisfaction for the German sample (T=-2.71, p<.05).

4 Discussion

The study compared discrete emotion experiences in the work place and their distinct influence on job satisfaction between representatives of interdependent and independent cultures. The results underline the presumptions by AET that the work context leads to emotional experiences, even after small unpleasant events like daily hassles and not only after major detrimental situations. Our results show that emotions do not lead deterministically to job satisfaction but that they depend on the preceding event and the cultural context in which they are appraised. These findings push the model further by emphasizing the cultural dependency of emotions on the structure of unpleasant work related events.

We confirmed our first assumption that unpleasant job events are elicitors for unpleasant emotions but that the emotion depends on the cultural background whether independent or interdependent. Yet the study pushed forward the discussion about the nature of events. The results indicate that even mild unpleasant social situations, often just regarded as regular parts of work life, determine job satisfaction through its link via emotions. But the events do not cause similar emotions. The experience of emotions rather depends on the way employees give meaning to the situation and this also has a cultural tone.

Consistent with our predictions Japanese employees experienced higher levels of self-conscious emotions shame and guilt even in situations where somebody else was the originator of the daily hassle (H1.1 and H2.1). Contrary to our expectations the level of neither frustration in the first version nor anger in the second version was higher in the German group.

Kitayama and Uchida (2003) already suggested that some emotional experiences are also quite universal. The results indicate that frustration and anger for example are present for Japanese employees as well but are accompanied by cultural dependent emotions of guilt and shame. No cultural differences for any of the four emotions were found when the respondent himself was the originator of a daily hassle. These results differ from our hypotheses H3.1 and H4.1 where we assumed that shame or guilt will be the strongest emotions in both cultures and should have been experienced to a much higher degree by employees from an interdependent background.

Effects of emotions on job satisfaction were tested in regard to cultural backgrounds and attribution pattern. The hypotheses were partly supported. For externally caused and not controllable daily hassles a clear effect for frustration on job satisfaction could be demonstrated for German and also Japanese employees but no effect of shame for the Japanese sample (H1.2).

Also for externally caused and controllable daily hassles anger had an effect on job satisfaction for German and also Japanese employees. As proposed guilt had an effect on job satisfaction in the Japanese sample although shame revealed a positive effect on job satisfaction.

For internally caused daily hassles which were controllable we found a small effect of guilt on job satisfaction among Japanese employees (H3.2). Although not posed for internally caused and not controllable events (H4.2) an effect of guilt on job satisfaction can be seen for the German sample.

In the independent sample job satisfaction was influenced to a great degree by frustration when daily hassles were elicited by colleagues or supervisors and when these people were not expected to have had control over the situation. According to our predictions job satisfaction was influenced by anger in the German sample when other people caused daily hassles and were supposed to have had control over its elicitation. Although not predicted self caused daily hassles under no control cause the experience of guilt which subsequently influence job satisfaction.

Also in the interdependent sample our predictions were partly supported. Guilt influences job satisfaction when daily hassles were controllable externally. On the other hand side as in the German sample frustration and anger also influence job satisfaction when daily hassles were attributed to someone else.

In sum our results indicate that job satisfaction is influenced by emotional experiences but that the pattern depends on the cultural background of employees and the attribution to the causes of the job related events. It seems that for Japanese employees the attribution on controllability over an event is more important for the emotional response of guilt and its subsequent influence on job satisfaction than the attribution on the originator. For the German sample on the other hand side the originator seems more important for the emotional impact on job satisfaction. Yet externally caused daily hassles seem to influence job satisfaction through frustration and anger in both cultures.

The results so far support the proposed chains of AET but give them a new perspective about the structure of the events their subsequent emotional reactions and their influence on job satisfaction.

Yet one finding challenges the propositions of AET. For the independent sample our results confirm stronger experience of emotions after daily hassles compared to neutral situations. As proposed in AET this will lead to direct detrimental emotion driven behavior in the workplace and at the same time emotions will also influence cognitive driven behavior, mediated through job satisfaction.

But for the interdependent sample this proposed chain is not so clear. The results indicate that their emotional level is the same in neutral situations as well as after unpleasant events. There is reason to assume that cognitive driven behavior is influenced to a much higher degree by emotions because they are experienced during regular working conditions as well. According to our results that also means that employees from interdependent cultures will conduct more emotion driven behaviors than employees from an independent background, even during presumed regular working conditions and their cognitive driven behavior is actually saturated with feelings.

That also means that Japanese employees generally demonstrate a higher level of emotional reactions during their job than German employees with all the -detrimental- consequences. Regarding the high standard deviations for shame and guilt in the neutral version one can argue that this reflects the emotional work of interdependent employees for harmony in the work group.

Although the last argument remains speculative at this point and the results leave room for further studies, the findings shed further light on emotional experiences in the work place and their role as cultural dependent determinants of job satisfaction.

Although unpleasant events can be very different even within the same occupation further studies should find culturally different samples with the same occupational background to further standardize results.

References

1. Basch, J., Fisher, C.D.: Affective events-emotions matrix: A classification of work events and associated emotions. In: Ashkanasy, N.M., Härtel, C.E.J., Zerbe, W. J. (Hrsg.) Emotions in the workplace: Research, theory and practice, pp. S. 36–48. Quorum Books, Westport, CT (2000)
2. Brief, A.P., Weiss, H.W.: Organizational Behavior: Affect in the Workplace. Annual Review of Psychology 53, 279–307 (2002)
3. Briner, R.B.: The Neglect and Importance of Emotion at Work. European Journal of Work and Organizational Psychology 8(3), 323–346 (1999)
4. Briner, R.B.: What can research into emotion at work tell us about researching well-being at work? International Journal of Work Organisation and Emotion 1(1), 67–73 (2005)
5. Connolly, J.J., Viswesvaran, C.: The role of affectivity in job satisfaction: A meta-analysis. Personality and Individual Differences 29(2), 265–281 (2000)
6. Cranny, C.J., Smith, P.C., Stone, E.F.: Job Satisfaction: How people feel about their jobs and how it affects their performance, Lexington, New York (1992)
7. Daniels, K.: Measures of five aspects of affective well-being at work. Human Relations 53(2), 275–294 (2000)
8. Dormann, C., Zijlstra, F.R.H.: Call centres: High on technology - high on emotions. European Journal of Work and Organizational Psychology 12(4), 305–310 (2003)
9. Fisher, C.D.: Antecedents and consequences of real-time affective reactions at work. Motivation and Emotion 26, 3–30 (2002)
10. Hagihara, A., Tarumi, K., Babazono, A., Nobutomo, K., Morimoto, K.: Work versus non-work predictors of job satisfaction among japanese white-collar workers. Journal of Occupational Health 40, 285–292 (1998)
11. Hart, P.M.: Predicting employee life satisfaction: A coherent model of personality, work and nonwork experiences, and domain satisfactions. Journal of Applied Psychology 84, 564–584 (1999)
12. Heine, S.J., Lehman, D.R., Markus, H.R., Kitayama, S.: Is There a Universal Need for Positive Self-Regard? Psychological Review 106(4), 766–794 (1999)
13. Hui, C.H., Yee, C.: Culture, control beliefs, and job satisfaction. In: Adamopoulos, J., Kashima, Y. (Hrsg.) Social psychology and cultural context, pp. S. 231–241. Sage, Thousand Oaks (1999)
14. Iaffaldano, M.T., Muchinsky, P.M.: Job satisfaction and job performance: A meta-analysis. Psychological Bulletin 97, 251–273 (1985)
15. Izard, C.E.: The Psychology of Emotions. Springer, Heidelberg (1991)
16. Judge, T.A., Heller, D., Mount, M.K.: Five-factors model of personality and job-satisfaction: A meta-analysis. Journal of Applied Psychology 87, 530–541 (2002)
17. Judge, T.A., Thoresen, C.J., Bono, J.E., Patton, G.K.: The job satisfaction-job performance relationship: A qualitative and quantitative review. Psychological Bulletin 127, 376–407 (2001)
18. Kanner, A.D., Coyne, J.C., Schaefer, C., Lazarus, R.S.: Comparison of Two Modes of Stress Measurement: Daily Hassles and Uplift Versus Major Life Events. Journal of Behavioral Medicine 4(1), 1–39 (1980)
19. Karasawa, K.: Cognitive antecedents of emotions: Findings and future directions. Japanese Psychological Research 37, 40–55 (1995)
20. Kitayama, S., Markus, H.R., Kurokawa, M.: Culture, Emotion, and Well-being: Good Feelings in Japan and the United States. Cognition and Emotion 14(1), 93–124 (2000)

21. Kitayama, S., Uchida, Y.: Explicit self-criticism and implicit self-regard: Evaluating self and friend in two cultures. Journal of Experimental Social Psychology (39), 476–482 (2003)
22. Knapp, K.: Interpersonale und interkulturelle Kommunikation. In: Bergemann, N., Sourisseaux, A.L.J. (Hrsg.) Interkulturelles Management (2. Aufl., S. 59-79). Physica-Verlag, Heidelberg (1996)
23. Koda, T.: Cross-Cultural Study of Avatars' Facial Expressions and Design Consideration within Asian Countries. In: Ishida, T., Fussell, S.R., Vossen, P.T.J.M. (eds.) IWIC 2007. LNCS, vol. 4568, Springer, Heidelberg (2007)
24. Lebra, T.S.: Shame and guilt: A psychocultural view of the japanese self. Ethos 11, 192–209 (1983)
25. Locke, E.A.: The nature and causes of job satisfaction. In: Dunette, M.D. (ed.) Handbook of industrial and organizational psychology, pp. 1297–1349. Rand McNally, Chicago (1976)
26. Markus, H.R., Kitayama, S.: Culture and the self: Implications for cognition, emotion and motivation. Psychological Review 98, 224–253 (1991)
27. Markus, H.R., Kitayama, S.: Culture, Self, and the Reality of the Social. Psychological Inquiry 14, 277–283 (2003)
28. Mesquita, B., Frijda, N.H.: Cultural variations in emotions: A review. Psychological Bulletin 112, 179–204 (1992)
29. Mesquita, B., Karasawa, M.: Different emotional lives. Cognition and Emotion 16, 127–141 (2002)
30. Money, R.B., Graham, J.L.: Salesperson Performance, Pay, and Job Satisfaction: Tests of a Model Using Data Collected in the United States and Japan. Journal of International Business Studies 30(1), 149–172 (1999)
31. Morris, W.N.: Mood: The frame of mind. Springer, New York (1989)
32. Ohbuchi, K.I., Suzuki, M., Hayashi, Y.: Conflict management and organizational attitudes among Japanese: Individual and group goals and justice. Asian Journal of Social Psychology 4, 93–101 (2001)
33. Organ, D.W., Near, J.P.: Cognition vs affect in measures of job satisfaction. International Journal of Psychology 20, 241–253 (1985)
34. Perrewé, P.L., Zellars, K.L.: An examination of attributions and emotions in the transactional approach to the organizational stress process. Journal of Organizational Behavior 20, 739–752 (1999)
35. Roos, J.: Peinlichkeit, Scham und Schuld. In: Otto, J.H., Euler, H.A., Mandl, H. (Hrsg.) Emotionspsychologie, pp. S. 264–271. Psychologische Verlags Union, Weinheim (2000)
36. Scherer, K.R.: Profiles of Emotion-antecedent Appraisal: Testing Theoretical Predictions across cultures. Cognition and Emotion 11(2), 113–150 (1997)
37. Six, B., Kleinbeck, U.: Arbeitsmotivation und Arbeitszufriedenheit. In: Roth, E (ed.) Enzyklopädie der Psychologie: Organisationspsychologie. Verlag für Psychologie Hogrefe, Göttingen, Toronto, Zürich (1995)
38. Spector, P.E., Cooper, C.L., Sanchez, J.I., O'Driscoll, M., Sparks, K., Bernin, P., Büssing, A., Dewe, P., Hart, P., Lu, L., Miller, K., de Moraes, L.R., Ostrognay, G.M., Pagon, M., Pitariu, H.D., Poelmans, S.A.Y., Radhakrishnan, P., Russinova, V., Salamatov, V., Salgado, J.F., Shima, S., Siu, O.-L., Stora, J.B., Teichmann, M., Theorell, T., Vlerick, P., Westmann, M., Widerszal-Bazyl, M., Wong, P.T.P., Yu, S.: Locus of control and well-being at work: How generalizable are western findings. Academy of Management Journal 45, 453–466 (2002)

39. Spector, P.E., Wimalasiri, J.: A cross-cultural comparison of job satisfaction dimensions in the United States and Singapore. International Review of Applied Psychology 35, 147–158 (1986)
40. Sugitani, M.: Kontextualismus als Verhaltensprinzip: Kritisch erlebte Interaktionssituationen in der japanisch-deutschen Begegnung. In: Thomas, A. (Hrsg.) Psychologie interkulturellen Handelns, pp. S. 227–245. Hogrefe, Göttingen (1996)
41. Vatrapu, R., Suthers, D.: Culture and Computers: A Review of the Concept of Culture and Implications for Intercultural Collaborative Online Learning. In: Ishida, T., Fussell, S.R., Vossen, P.T.J.M. (eds.) IWIC 2007. LNCS, vol. 4568, Springer, Heidelberg (2007)
42. Weiner, B.: An attributional theory of achievement motivation and emotion. Psychological Review 92, 548–573 (1985)
43. Weiss, H.M., Cropanzano, R.: Affective Events Theory: A theoretical discussion of the structure, causes and consequences of affective experiences at work. In: Staw, B.M., Cummings, L.L. (eds.) Research in organizational behavior, vol. Bd. 18, pp. S. 1–74. JAI Press, Greenwich (1996)

Intra- and Inter-cultural Collaboration in Science and Engineering

Gary M. Olson and Airong Luo

School of Information, University of Michigan,
1075 Beal Avenue, Ann Arbor, MI 48109-2112
{gmo,airongl}@umich.edu

Abstract. Collaboratories that support science and engineering have become more and more common. Unfortunately, many of them experience serious difficulties. Those that involve inter-cultural collaboration are especially problematic. We have identified more than 200 such projects, and have formulated a series of working hypotheses about what factors are associated with success and failure. In this article we review these factors, focusing in particular on those aspects that arise in inter-cultural collaborations.

1 Introduction

A collaboratory is a geographically dispersed organization that brings together scientists, instrumentation, and data to facilitate scientific research[1, 2]. It supports rich and recurring human interaction oriented to a common research area, and provides access to the data sources, artifacts and tools required to accomplish research tasks. Collaboratories have been made possible by new communication and computational tools that enable more flexible and ambitious collaborations. Such collaborations are increasingly necessary. As science progresses, the unsolved problems become more complex, the need for expensive instrumentation increases, larger data sets are required, and a wider range of expertise is needed. At least 200 collaboratories have appeared in the past fifteen years, representing almost all areas of science.

Since collaboratories support geographically distributed interactions, many of them involve participants from different countries or cultures. Studies have shown that collaboration in general is difficult, and geographical dispersion adds further complications to be overcome [3]. The addition of inter-cultural participants adds yet more risk to such projects. Our Science of Collaboratories (SOC) project (www.scienceofcollaboratories.org) has been examining why some collaboratories are successful and others fail. We will summarize our findings in this paper, highlighting the intercultural factors that we have seen.

2 Types of Collaboratories

One of the first things we noticed in studying collaboratories is that there are a number of different kinds. We have developed a taxonomy of collaboratory types based on the defining characteristics of a collaboratory that we provided earlier.

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 249–259, 2007.

Collaboratories vary on the extent to which they support all aspects of collaboration. Further, some are focused on research, others on the support of practice.

2.1 Collaboratories with a Research Focus

The following four types of collaboratories focus on enabling geographically distributed research.

Distributed Research Center. This type of collaboratory functions like a full-fledged research center or laboratory, but at a distance. It has a specific area of interest and a general mission, with a number of specific projects. A good example of a Distributed Research Center is the Alliance for Cellular Signaling (www.cellularsignaling.org), a large, complex distributed organization of people at universities whose goal is to understand how cells communicate with each other to make an organism work.

Shared Instrument. A Shared Instrument collaboratory provides access to specialized or geographically remote facilities. As the frontiers of science are pushed back, the instrumentation required for advances becomes more and more esoteric, and therefore usually more and more expensive. Alternatively, certain kinds of science need instrumentation that is in specific geographic settings, such as in an isolated or inhospitable area. A typical example is the Keck Observatory (www.keckobservatory.org), which provides access to an observatory on the summit of Mauna Kea in Hawaii to a consortium of people at California universities.

Community Data System. An especially common collaboratory type is one in which a geographically dispersed community agrees to share their data through a federated or centralized repository. The goal is to create a more extensive data set, for which more sophisticated or powerful analyses can be done than if the parts of the data set were kept separately. A typical example of a Community Data System is the Zebrafish Information Network (zfin.org), an on-line aggregation of genetic, anatomical, and methodological information for zebrafish researchers.

Open Community Contribution System. This is an instance of an emerging organizational type known as a “voluntary association” [4, 5]. Interested members of a community (defined usually quite broadly) are able to make small contributions (Sproull calls them “microcontributions”) to some larger enterprise. These contributions are judged by a central approval organization and placed into a growing repository. The classic example is open source software development, where hundreds or even thousands of contributors offer bug fixes or feature extensions to a software system. In science, such schemes are used to gather data from a large number of contributors. Two examples will help illustrate this. The NASA Ames Clickworkers project (clickworkers.arc.nasa.gov) invited members of the public to help with the identification of craters on images from a Viking mission to Mars. They received 1.9 million crater markings from over 85,000 contributors, and the averaged results of these community contributions were equivalent in quality to those of expert geologists. A second example is the Open Mind Common Sense Initiative (commonsense.media.mit.edu), which is collecting examples of common sense knowledge from the public for use in artificial intelligence programs that do reasoning or language understanding.

2.2 Collaboratories with a Practice Focus

The next two collaboratory types support the professional practice of science more broadly, as opposed to the conduct of research itself.

Virtual Community of Practice. This is a network of individuals who share a research area of interest and seek to share news of professional interest, advice, job opportunities, practical tips on methods, and the like. A good example of this kind of collaboratory is Ocean US (www.ocean.us), which supports a broad community of researchers interested in ocean observations.

Virtual Learning Community. This type of collaboratory focuses on learning that is relevant to research, but not research itself. A good example is the Ecological Circuitry Collaboratory (www.ecostudies.org/cc/), whose goal is to train PhD. students in ecology in quantitative modeling methods.

2.3 Evolution of Collaboratories

Collaboratories that last more than a year or two tend to evolve. For example, a collaboratory may start as a Shared Instrument collaboratory. Those who share the instrument may add a shared database component to it, moving the collaboratory toward a Community Data System. The users may add communication and collaboration tools so they can plan experiments or data analyses, making the collaboratory more like a Distributed Research Center. This is the path that the Upper Atmospheric Research Collaboratory (UARC) and the Space Physics and Aeronomy Research Collaboratory (SPARC) took over a ten-year period [6].

3 What Is Success?

The question of what makes for a successful collaboratory depends, of course, on what one means by success. This turns out to be a complicated question. Table 1 lists the primary criteria for success (a more complete list is available on the SOC web site mentioned above). Clearly the criterion of greatest interest is direct effects on the science. This could be getting things done more quickly or more accurately. It could mean making discoveries that otherwise would be difficult or impossible. While this is a highly desirable outcome, it is of course the most difficult to achieve. In the more than 200 collaboratories we have studied so far, this has been an extremely rare outcome up to this point.

However, many of the other criteria in Table 1 are also desirable. For instance, collaboratories sometimes have impacts on science careers, such as enhancing productivity, making for more collaboration, or helping the development of young scientists. Collaboratories can also affect the careers of scientists in the developing world, who can have more opportunities to participate in leading research projects.

Table 1. Criteria for success in collaboratories

Success Criterion
Direct effects on the science
Science careers
Effects on learning, science education
Broader participation in science
Effects on funding, public perception
Creation of software technology
Use of the collaboratory
Inspiration for other collaboratories

This is certainly true in the world of high energy physics [7]. A number of collaboratories have provided access to science for school children or university students, possibly playing a role in attracting young people into science careers. Collaboratories can make it easier for scientists at less prestigious institutions or in the developing world to participate in science. Several funding initiatives can be traced to early successes in certain highly visible collaboratories, having affected the public perception of the value of science in places like Congress

The last few criteria in Table 1 are probably the weakest measures of success. But in a number of projects, the tools created and distributed to scientists were never even used. Or they resulted in one-off software that was not reused in other projects. Some of the earliest collaboratories inspired others to attempt them, but of course this is a success criterion of largely passing historical interest.

In this brief paper we do not have space to elaborate further on these matters. Our forthcoming book [8] we discuss these issues in greater detail.

4 Factors That Affect Success

Our examination of more than 200 collaboratory projects has led us to formulate a series of working hypotheses as to what lies behind success or failure. We briefly review these here, stressing for purposes of this paper those aspects that affect collaboratories involving inter-cultural participants.

4.1 The Nature of the Work

Not all kinds of work can be done easily at a distance. In general, the more partitionable the work is, the easier it is to carry out at a distance [9]. Tasks that involve divide-and-conquer work best. Of course the tasks cannot be totally independent – that would not be collaboration. There needs to be some level of interaction to keep the tasks coordinated. In inter-cultural collaborations, there can sometimes be misunderstandings about what the tasks are or how they might be partitioned. In one of the collaboratories we studied, the hierarchical culture of the Chinese and Korean teams decides that the group leaders tended to shoulder the tasks that they think important, such as video or teleconferencing with their remotely located collaborators, attending international conferences or visiting their collaborators' labs. However, sometimes their collaborators from US or Europe preferred to communicate with the junior scientists or students who involve in the tasks directly.

4.2 Common Ground

Communication effectiveness, a key element of any collaboration, requires that the participants have shared knowledge, described by Clark [10] as *common ground*. This can consist of shared technical or domain knowledge, shared beliefs, and shared assumptions about social practices. It can be reflected in shared vocabulary. Further, it is important that the participants know that they have such knowledge in common, so it can influence how they communicate.

A good example of these issues is the ATLAS project based at CERN. This is a collaboration involving approximately 1500 scientists from all over the world. It has followed a divide-and-conquer strategy, in which the complex tasks have been divided into smaller ones that are handled by a more modest team size. But if course very large amounts of coordination are required, and the project has involved much travel and frequent audio or video conferences. Of course these are large costs for scientists from the developing world [7].

One complication with science is that much of the knowledge that is shared is *tacit*, that is, learned through participating in an interacting community of practice [7, 11]. Such tacit knowledge is extremely hard to acquire through distance communication technologies. For example, in an HIV/AIDS project that we studied the principals found it was necessary to travel frequently to southern Africa so the tacit knowledge involved in laboratory procedures could be shared. Though the participants used web conferencing and other communication tools, the tacit knowledge could only be shared in person. For scientists in the developing world, who are typically isolated and do not have access to substantial travel funds, this can be a huge obstacle.

There is also common ground about social practices. In an example we studied in the auto industry, engineers from the US, France, and Germany had a video conference on their work. It turns out that it was the last day of work for one of the French engineers. At the end of the working part of the conference, the US participants abruptly turned off their video connection. But the German and French engineers stayed on line to give the French engineer a friendly send-off. They all thought the Americans were quite rude to ignore this socially important occasion [9].In another instance, the US engineers set up a meeting for what in Europe was Friday afternoon, not understanding that the French engineers had a 35-hour work week and would ordinarily go home on Friday afternoon. Such cultural misunderstandings can make the work very difficult. As we pointed out elsewhere, the aphorism “when in Rome do as the Romans do” is difficult in distributed work. Where is Rome? [9]

It is well known that different communication media can affect how easily common ground can be established and supported [9, 12]. A laboratory study showed how this can affect intercultural communication [13]. Pairs of people who were in different locations and were either native speakers of English or were non-native speakers of English did a collaborative task. Native speakers were equally proficient at the task with either audio or video conferencing as their communication medium. However, non-native speakers were better with video conferencing than with audio conferencing. These latter collaborators, who had less common ground because of their different backgrounds, benefited from having an additional communication channel to support their conversation. The participants could more easily gauge the level of understanding of their partner, and could supplement speech with gestures. Indeed, adding video of the work objects as well even further enhances the establishment of common ground [14-16].We have seen evidence of similar phenomena in field studies of international collaboration [9]. In a HIV/AIDS collaboratory, where Chinese scientists collaborate with American scientists, they adopted teleconferencing and Centra, which enable them to share slides, view data, and have discussions. Scientists commented that compared to teleconferencing, the multiple channels provided by Centra helped to reduce misunderstandings in the communication process.

4.3 Collaboration Readiness

Collaboration readiness refers to the preparedness or willingness of participants to collaborate. Science is by its very nature a delicate balance of cooperation and competition. Successful collaborations require cooperation. But collaboration is also very difficult, and requires extra effort and motivation. Technologies that support collaboration will not be used if the participants are not ready or willing to collaborate. Various fields or user communities have quite different traditions of sharing. For instance, upper atmospheric physicists have had a long tradition of collaboration. Thus, the Upper Atmospheric Research Collaboratory (UARC) began with a collaborative set of users. On the other hand, several efforts to build collaboratories for biomedical research communities (for instance, HIV/AIDS, depression) have had difficulty in part because of their competitive atmosphere. Collaboration readiness can be an especially important factor when the collaboratory initiative comes from an external source, such as a funding agency.

Aligning individual and organizational incentives is an important element of collaboration readiness. A specific example that we have studied focuses on the incentives to participate in a Community Data System. What motivates a researcher to contribute data to a shared database, thus giving up exclusive access to the data they have collected? There are a variety of incentive schemes that have been tried, which we describe next.

Goodwill. The Zebrafish Information Network (ZFIN) that we mentioned earlier has relied so far on the goodwill of its members. Most of the members of this community had a connection to a specific senior researcher who both pioneered the use of Zebrafish as a model organism and also created in the community a spirit of generosity and collaboration. When ZFIN was begun, goodwill among the community of researchers was a sufficient incentive for participation. Now, as ZFIN expands its participation beyond its founders, it will be interesting to see whether the spirit of goodwill spreads to new researchers.

Goodwill plus Karma Points. Slashdot (slashdot.org) is a very large and active community of open source software developers who share and discuss news. Slashdot uses a system where the most informative contributions among the thousands of daily postings are brought more into the center of attention. Thus, high-quality contributions are noted and rewarded within a formal system of Karma points. Points are allocated to participants whose contributions are rated highly by others. These Karma points in turn give contributors some additional privileges on the site. But the main value of Karma points is as a tangible measure of community participation and status. Karma points are a formalization of good will, valuable primarily because the members of the community value them as an indicator of the quality of the sharing done by specific individuals.

Required Contribution for Doing Other Things. In order to get the details of gene sequences out of published articles in journals, a consortium of high prestige journals in biology require that the authors of a journal submission have a GenBank accession number indicating that they've stored their relevant gene sequences in the shared data base.

New Forms of Publication. The Alliance for Cellular Signaling explored a novel approach to providing an incentive to contribute molecule pages to their database. Because the molecules pages (i.e., a series of related data base entries) represent a lot of work, the Alliance worked out an agreement with *Nature*, one of the high prestige journals in the field, to count a molecule page as a publication in *Nature*. *Nature* coordinates the peer reviews, and although molecule page reviews do not appear in print they are published online with the prestige of the Nature Publishing Group. While this specific experiment has been terminated, *Nature* is exploring other similar arrangements with a variety of scientific fields.

Incentives to collaborate and to share resources can vary considerably when inter-cultural participants are involved. For instance, in an international HIV/AIDS collaboration that we studied, the scientists in North America and Europe had quite different motivations to collaborate than their colleagues in southern Africa. The overseas participants were eager to access patients and data about them, whereas the African participants wanted to enhance their skill level, build up their laboratories, and increase their standing in the field. Issues of trust and alignment of goals required constant attention, including lots of travel for face-to-face contact.

4.4 Management, Planning and Decision Making

Any large organization faces difficult management issues, and practicing scientists may not always have the time or the skills to properly manage a complex enterprise. And of course the management issues get even more complicated when the organization is geographically distributed. Many large collaboratories have faced key management issues. For instance, the two physics collaboratories -, Grid Physics Network (www.griphyn.org) and International Virtual Data Grid Laboratory (www.ivdgl.org), found that it was necessary to hire a full-time project manager for each one in order to help the science project directors manage the day-by-day activities of the projects. The Alliance for Cellular Signaling has benefited from a charismatic leader who has excellent management skills, and who has set up a rich management structure to oversee the project. The Biomedical Informatics Research Network (BIRN) has a project manager, an explicit Governance manual that contains guidance for a host of tricky management issues, and a steering committee that is responsible for implementing these management guidelines.

Data are a central component of all collaborations. There are numerous issues about how data are represented and managed that can contribute to collaboratory success. For example, good metadata are critical as databases increase in size and complexity. Metadata are data about data, such as a library catalog or an index to a file system. Metadata are key to navigation and search through such databases. Information about the provenance or origins of the data is also important. Data have often been highly processed, and researchers will want to know what was done between the original raw data and the processed data currently in the database. Two related collaboratories in high energy physics, GriPhyN (www.griphyn.org) and iVDGL (www.ivdgl.org), we just introduced, are developing schemes for showing investigators the path of transformations that led to the data in the database. This will help in understanding the data, and also in correcting any errors in the transformations that are discovered later.

An interesting new issue raised by some kinds of collaboratories is the complexity of jurisdictional issues when data are combined into a large database. The BIRN project (www.nbirn.net) is facing just such issues as it builds up a database of brain images. The original brain images were collected in different physical sites, at different universities or hospitals with different Institutional Review Boards (IRBs) that must approve any human data collection and preservation. The stipulations under which the original images were collected may be incompatible.

These issues are of course complex in international collaborations. Participants from different cultures can have different styles of managing and making decisions. For instance, a Japanese researcher pointed out that the kinds of decision making assumptions embedded in decision support software from the U.S. would not work in Japan [17]. Similarly, different countries may have quite different regulatory systems involving, for instance, different rules about informed consent in things like clinical trials. Further, in places like China there are very different regulations on the sharing of research materials like biological samples, seriously affecting the ability of HIV/AIDS researchers to collaborate across borders [7].

4.5 Technical Readiness

This refers to the technical capacity of the participants, the supporting infrastructure, and the design of the tools. Some communities are sufficiently collaborative to be good candidates for a successful collaboratory, but their experience with collaborative technologies or the supporting infrastructure is not sufficient. Technical readiness can be of three kinds.

Individual Technical Readiness. People in various organizations or fields have different levels of experience with collaboration tools. A specific new technology like application sharing may be a leap for some and an easy step for others. It is important to take account of users' specific experience when introducing new tools.

Infrastructure Readiness. Collaborative technologies require good infrastructure, both technical and social. Poor networks, incompatible workstations, or the lack of version control can cause major problems. It is also very important to have good technical support personnel, especially in the early phases of a collaboratory. The Worm Community System (WCS) was a very early collaboratory project, intended to support a community of researchers who studied the model organism *c. elegans*. Sophisticated software was developed for the WCS on a technical platform (UNIX) that was not commonly used in the laboratories of the scientists. Since these tools were thus not integrated with everyday practice, they were seldom used. Furthermore, the necessary technical support for such systems was usually not present in the lab, so when there were problems they were showstoppers [18].

Scientists in the developing world often have quite different infrastructures at their disposal. For example, we found in southern Africa that universities often had reasonable infrastructures on campus or within a lab, but that connections to the larger Internet were severely limited, in part because of very high costs for access to higher bandwidth. Some initiatives were being planned which would give the universities substantially reduced tariffs for high bandwidth access. In a high energy physics laboratory, scientists from Morocco reported that because of the network speed and high costs of telecommunication, they could never participate in any video conferences or teleconferences, which are important for coordination of the whole project. They were also worried that they would encounter great difficulty to obtain data from the detector located in Europe, which would seriously affect their participation in the project.

Social Ergonomics of Tools. The social interactions that take place in teams are affected both by the characteristics of its participants as well as the tools that are used. We refer to the study of the impact of technology characteristics on this process as social ergonomics (ergonomics is the application of knowledge about humans to the design of things). For example, video conferencing systems often ignore such details as screen size, display arrangement in relation to participants, camera angle, and sound volume. But it turns out these details can have social effects. A study found that the apparent height of videoconference participants, as conveyed via camera angle, influenced a negotiation task [19]. The apparently taller person was more influential in shaping the final outcome than the apparently shorter person.

5 Conclusions

Geographically dispersed collaborations are quite difficult, and when inter-cultural dimensions are added to the mix, can be even more problematic. However, we have seen that there are interventions that can minimize the risks associated with the success factors we have listed. These interventions often require extra work or better tools. For instance, frequent face-to-face contact can make a big difference in communication effectiveness, trust, and motivation. Better communication tools like video can help as well. Of course, such interventions are not always easy to do,

especially with collaborations that involve developing nations, very large distance, and very large time zone differences. Nonetheless, in many cases the needs of science and engineering research require such collaborations, and efforts made to ensure their success can be worthwhile.

Acknowledgments. The work reported in this paper has been supported by a grant from the National Science Foundation (IIS 0085951) and a contract from the Army Research Institute (W74V8H-06-P-0518). Many people have been involved in this research, but primary collaborators include Judy Olson, Nathan Bos, Ann Zimmerman, and Dan Cooney.

References

1. Wulf, W.A.: The collaboratory opportunity. Science 261, 854–855 (1993)
2. Finholt, T.A., Olson, G.M.: From laboratories to collaboratories: A new organizational form for scientific collaboration. Psychological Science 8, 28–36 (1997)
3. Cummings, J., Kiesler, S.: Collaborative research across disciplinary and institutional boundaries. Social Studies of Science 35, 703–722 (2005)
4. Sproull, L., Kiesler, S.: Public volunteer work on the Internet. In: Kahin, B., Dutten, W. (eds.) Transforming enterprise. MIT Press, Cambridge, MA (2004)
5. Sproull, L., Conley, C., Moon, J.Y.: Pro-social behavior on the net. In: Amichai-Hamburger, Y. (ed.) The social net: the social psychology of the Internet. Oxford University Press, New York (2004)
6. Olson, G.M., Finholt, T.A., Killeen, T.: The Upper Atmospheric Research Collaboratory (UARC) and Space Physics and Aeronomy Research Collaboratory (SPARC). In: Olson, G.M., Zimmerman, A., Bos, N. (eds.) Science on the Internet. MIT Press, Cambridge, MA (2008)
7. Luo, A., Olson, J.S.: How collaboratories affect scientists from developing countries. In: Olson, G.M., Zimmerman, A., Bos, N. (eds.) Science on the Internet. MIT Press, Cambridge, MA (2008)
8. Olson, G.M., Zimmerman, A., Bos, N. (eds.): Science on the Internet. MIT Press, Cambridge, MA (2008)
9. Olson, G.M., Olson, J.S.: Distance matters. Human-Computer Interaction 15, 139–179 (2000)
10. Clark, H.H.: Using language. Cambridge University Press, New York (1996)
11. Knorr Cetina, K.: Epistemic cultures: How the sciences make knowledge. Harvard University Press, Cambridge, MA (1999)
12. Clark, H.H., Brennan, S.E.: Grounding in communication. In: Resnick, L., Levine, J.M., Teasley, S.D. (eds.) Perspectives on socially shared cognition, pp. 127–149. APA, Washington, DC (1991)
13. Veinott, E., Olson, G.M., Olson, J.S., Fu, X.: Video helps remote work: Speakers who need to negotiate common ground benefit from seeing each other. In: CHI 99, pp. 302–309. ACM Press, New York (1999)

14. Karsenty, L.: Cooperative work and shared visual context: An empirical study of comprehension problems in side-by-side and remote help dialogues. Human-Computer Interaction 14, 283–315 (1999)
15. Fussell, S.R., Kraut, R.E., Siegel, J.: Coordination of communication: Effects of shared visual context on collaborative work. In: CSCW 2000, pp. 21–30. ACM Press, New York (2000)
16. Kraut, R.E., Fussell, S.R., Siegel, J.: Visual information as a conversational resource in collaborative physical tasks. Human-Computer Interaction 18, 13–39 (2003)
17. Ishii, H.: Cross-cultural communication & CSCW. Whole Earth Review, 48–52 (1990)
18. Star, S.L., Ruhleder, K.: Steps toward an ecology of infrastructure: Complex problems in design and access for large-scale collaborative systems. In: CSCW 94, pp. 253–264. ACM Press, New York (1994)
19. Huang, W., Olson, J.S., Olson, G.M.: Camera angle affects dominance in video-mediated communication. In: CHI 2002, pp. 716–717. ACM Press, New York (2002)

Culture and Computers: A Review of the Concept of Culture and Implications for Intercultural Collaborative Online Learning

Ravi Vatrapu and Dan Suthers

Laboratory for Interactive Learning Technologies (LILT), Department of Information and Computer Sciences, University of Hawaii at Manoa, POST Bldg 317, 1680 East-West RD, Honolulu, HI, USA, 96822
vatrapu@hawaii.edu, suthers@hawaii.edu

Abstract. Our research is aimed at a systematic investigation of phenomena in the nexus of culture, technology and learning. The basic premise of our research is that social affordances of technologies might vary along cultural dimensions. In this paper we present a brief overview of the concept of culture. We then discuss empirical findings demonstrating cultural effects on social behavior, communication and cognition and draw implications to online collaborative learning. In the last part of this paper, we present a selective review of research in cross-cultural human computer interaction.

1 Introduction

Our research is aimed at a systematic investigation of phenomena in the nexus of culture, technology and learning. *The basic premise of our research is that social affordances of technologies might vary along cultural dimensions.* The challenge for technological learning environments is that interacting through technology is not unproblematic. First, it makes interaction more difficult [7], [52]. Second, it may not mean, feel and afford the same thing to students from different cultures. Designers assume that the online environment is the same for everyone, but do users perceive something different? One of our research objectives is to explore alternative representational contexts for interaction because they may provide resources that are appropriate for different cultural members.

In this paper we will present a brief history of the problematics associated with the notion of culture. A summary discussion of the intellectual history of the concept of culture follows. Hofstede's [27] definition of culture is mentioned next. The section on culture, behavior, communication and cognition discusses Geert Hofstede's cultural dimensions model [27], cultural differences in traditional learning settings; Edward Hall's [25] communicative context cultural dimension and cultural psychology findings of Nisbett & colleagues [47], [48]. This paper concludes with a selective summary of the growing literature on culture and computers.

2 Concepts of Culture

The concept of culture has a checkered intellectual history. Raymond Williams [66] has termed culture "*one of the two or three most complicated words in the English*

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 260–275, 2007.

language.” Williams attributes this complexity of the concept to its complicated historical development in many European languages and its subsequent adoption in to a plurality of academic disciplines. The further complications with the concept of culture arise from the slippage of meaning between the academic usages of the term and the popular usages of the term like in “high culture” vs. “pop culture.” As Williams [67] puts it “*the concept at once fuses and confuses the radically different experiences and tendencies of its formation*.”

2.1 Definitions of Culture

The Victorian Ethnologist, Edward Tylor, is generally credited for providing the first definition of culture in anthropology. According to Tylor [60], “*culture or civilization, taken in its wide ethnographic sense, is that complex whole which includes knowledge, belief, art, morals, law, custom, and other capabilities and habits acquired by man as a member of society*.”

A compiled list of over 200 different definitions of culture can be found in Kroeber and Kluckhohn’s [35] critical review of the concept of culture. Tracing the intellectual evolution of notion of culture from the German “*kultur*” and its tensions with the notion of civilization, Kroeber and Kluckhohn categorize the different definitions of cultures into 6 groups: *descriptive, historical, normative, psychological, structural and genetic*. The psychological definitions of culture have “emphasis on adjustment, on culture as a problem-solving device” [35] and also have an emphasis on culture as learning. Two other definitions of culture are listed below:

"Culture is learned and shared human patterns or models for living; day-to-day living patterns. These patterns and models pervade all aspects of human social interaction. Culture is mankind's primary adaptive mechanism" [11].

"Culture is the shared knowledge and schemes created by a set of people for perceiving, interpreting, expressing, and responding to the social realities around them" [37].

In the above two definitions, culture is defined as patterns or schemas. Culture is shared and is a factor in an individual’s social interactions. The definitions also emphasize the adaptiveness implicit in the notion of culture and the importance of culture as an active notion and not a passive one. The conception of cultures as instrumental adaptations to environmental demands has been a feature of cultural ecology strands of anthropology. For example, according to White [65], “*all life is a struggle for free energy*” and culture is a means in that struggle toward the ends of survival.

Geert Hofstede’s definition of culture has been influential in the organizational communication and business literatures. According to Hofstede [27] culture is “*the collective programming of the mind which distinguishes the members of one group or category of people from another.”* The emphasis in Hofstede’s definition is that culture is learnt in nurture and not inherited by human nature. The “*collective programming of the mind*” highlights culture as a collective activity that is to be conceived as a dynamic process rather than a passive state. The other part of the definition “*which distinguishes the members of one group or category of people from another”* points out the individual and group identity formation and sustenance aspects of enculturation in social institutions like family, school and work. Culture comes from similarity of individuals within a cohort group (be it a linguistic community, an ethnic group or a scientific community) and in that sense it is collective. This similarity is not intended to be exact; neither does it imply essentialist homogeneity. In a multicultural society culture is about collective particularity.

Hofstede's definition is best interpreted from the Vygotskian socio-cultural perspective of the "*social formation of the mind*" [64]. "*Collective programming*" is not to be understood as an external imposition but an active social composition in which the particular individual plays the protagonist. Taken together, "social formation" and "collective programming" of the mind indicate a cognitive schema. Culture is operationalized as a cognitive schema in our research.

In our own research the concept of culture is treated as an independent variable. Cognitive sciences have highlighted the role of schemas and models in the mundane activities of everyday life. Culture in this proposal is operationalized in this cognitive scientific sense. Cultural schemas are putative structures; they are properties of an individual's mind. Cultural schemas are not things; neither are they substances. They are formed and forged from an individual's biography. This biography includes the interactive effects of the geography of that individual's upbringing (ecology) and the formative experiences of his/her life (history).

3 Culture, Behavior, Communication and Cognition

In the next three sections we will discuss three separate lines of empirical findings from the fields of social behavior, communication and cognition. We will attempt to concretize our discussion of cultural effects by drawing implications to *intercultural* online collaborative learning environments.

3.1 Culture and Behavior

Cultural models can be used to identify the differences in cultures that affect how computer supported collaborative learning environments (CSCL) are used. There are two kinds of cultural models: models that use typologies and models that use dimensions. Typologies describe a number of ideal types each easy to imagine. Dimensional models group together a number of phenomena in a society that were empirically found to occur in combination into dimensions. Typologies are difficult to adopt in empirical research as real cases very rarely correspond to one single ideal type.

In our own research the concept of culture is treated as an in Hofstede's seminal work on cultures in organizations formulated a framework of four dimensions of culture identified across nations. Each dimension groups together phenomena in a society that were empirically found to occur in combination. In this section, Hofstede's definitions for these original four cultural dimensions are listed followed by a discussion of each dimension with respect to online learning. Hofstede's cultural dimensions model indicates what reactions are likely and understandable given one's cultural background.

Low Power Distance vs. High Power Distance. Power distance is the "extent to which the less powerful members of institutions and organizations within a country expect and accept the power that is distributed unequally" [27]. People in large power distance cultures are much more comfortable with a larger power/status differential than small power distance cultures. Table 1, adapted from Hofstede [26], outlines the effects of power dimension that have implications for online learning environments. It is important to note that Hofstede's conception of power distance is not a bi-directional one; it is conceived as a subordinate's expectation and acceptance of unequal distributions of power in a social setting.

Table 1. Power distance dimension in traditional classrooms

Small Power Distance Societies	Large Power Distance Societies
Student-centered education (premium on initiative)	Teacher-centered education (premium on order)
Teacher expects students to initiate communication	Students expect teacher to initiate communication
Students may speak up spontaneously in class	Students speak up in class only when invited by the teacher
Students allowed to contradict or criticize teacher	Teacher is never contradicted nor publicly criticized
Effectiveness of learning related to amount of two-way communication in class	Effectiveness of learning related to excellence of the teacher
Outside class, teachers are treated as equals	Respect for teachers is also shown outside class

If online education is offered as an alternative to traditional schooling then it is important to investigate how students perceive the social affordances of the virtual learning institutions. For example, in our ongoing research study, set within a context of collaborative problem solving, students co-constructing concept maps are provided information attributed to scientists who have authority by virtue of their expertise and experience. Arguments from authority are valued in the scientific enterprise if those authorities themselves adhere to the scientific method. The point here is not whether the issues of power distance will show up in online classrooms but rather how does this dimension help understand the interactional behavior in an online learning setting.

Individualism vs. Collectivism. "Individualism pertains to societies in which the ties between individuals are loose: everyone is expected to look after himself or herself and his or her immediate family. Collectivism as its opposite pertains to societies in which people from birth onwards are integrated into strong, cohesive in-groups, which throughout people's lifetime continue to protect them in exchange for unquestioning loyalty" [27]. This dimension describes the degree to which a culture emphasizes an individual's reliance on the self or the group. Table 2, adapted from Hofstede [26], outlines the effects this dimension that have implications to online environments.

This dimension is of particular interest to the social constructivist theories of learning given the small group size emphasis of CSCL. In inter-cultural online learning groups, dynamics of in-group and out-group memberships might affect how

Table 2. Collectivism vs. individualism dimension in traditional classrooms

Collectivist Societies	Individualist Societies
Students expect to learn how to do	Students expect to learn how to learn
Individual students will only speak up in class when called upon personally by the teacher	Individual students will speak up in class in response to a general invitation by the teacher
Individuals will only speak up in small groups	Individuals will speak up in large groups
Large classes split socially into smaller cohesive subgroups based on particularistic criteria (e.g. ethnic affiliation)	Sub-groupings in class vary from one situation to the next based on universalistic criteria (e.g. the task "at hand")
Formal harmony in learning situations should be maintained at all times	Confrontation in learning situations can be salutary; conflicts can be brought into the open
Neither the teacher nor any student should ever be made to lose face	Face-consciousness is weak

certain technology affordances are appropriated as social affordances. They might also affect the perception of other students in the online learning environment. The notion of face-saving is of important when it comes to subjective perceptions and evaluation of the user interface, online interaction and instructional elements of an online course.

Based on socio–cognitive conflict theory [17] collaborative learning effectiveness is thought to be influenced by the extent that students jointly identify and discuss conflicts in their knowledge beliefs [8]. This works well in an individualist culture but in collectivist cultures consensual forms of intersubjective meaning making processes may be more prevalent.

Femininity vs. Masculinity. "Masculinity pertains to societies in which the gender roles are clearly distinct; femininity pertains to societies in which the gender roles overlap" [27]. This dimension refers to expected gender based division of labor in a culture. The cultures that score towards what Hofstede refers to, in a confusing choice of category labels, as "masculine" tend to have very distinct expectations of male and female roles in society. The more "feminine" cultures have a greater ambiguity in what is expected of each gender. Table 3, adapted from Hofstede [26], summarizes the implications of this dimension for online learning environments.

Table 3. Femininity vs. masculinity dimension in traditional classrooms

Feminine Societies	Masculine Societies
Teachers avoid openly praising students	Teachers openly praise good students
Teachers use average student as the norm	Teachers use best students as the norm
System rewards students' social adaptation	System rewards students' academic performance
Students practice mutual solidarity	Students compete with each other in class
Students try to behave modestly	Students try to make themselves visible

Collaborative learning is often distinguished from cooperative learning by the argument that collaboration involves joint activity or an effort to maintain a joint conception [55] whereas cooperation involves a mere joining of individual activities [15]. Collaboration is often conceived of as being beyond a basic division of labor and more of an enterprise involving parties with equal stakes. Division of labor is one of the most important concepts in Anthropology [29] and in Sociology [19]. The effects of culture on the division of interactional labor and the distribution of cognitive work in collaborative learning can be understood by investigating how participants appropriate affordances.

High Uncertainty Avoidance vs. Low Uncertainty Avoidance. *"The extent to which the members of the culture feel threatened by uncertain or unknown situations"* [27]. This dimension refers to how comfortable people feel towards ambiguity. Low uncertainty avoidance cultures feel much more comfortable with the unknown. High uncertainty avoidance cultures prefer formal rules and any uncertainty can express itself in higher anxiety. Table 4, adapted from Hofstede [26], summarizes the effects this dimension that have implications to online learning environments.

Table 4. Uncertainty avoidance dimension in traditional classrooms

Weak Uncertainty Avoidance Societies	Strong Uncertainty Avoidance Societies
Students feel comfortable in unstructured learning situations: vague objectives, broad assignments, no timetables	Students feel comfortable in structured learning situations: precise objectives, detailed assignments, strict timetables
Teachers are allowed to say "I don't know"	Teachers are expected to have all the answers
Students are rewarded for innovative approaches to problem solving	Students are rewarded for accuracy in problem-solving
Teachers are expected to suppress emotions (and so are students)	Teacher are allowed to behave emotionally (and so are students)
Teachers interpret intellectual disagreement as a stimulating exercise	Teachers interpret intellectual disagreement as personal disloyalty

The dimension of uncertainty avoidance can affect how students perceive social affordances of the online learning environment. Also of importance are the effects of culture on the interpretation or an acknowledgement of the ambiguous data and judgment of the relevance of data in the unfolding interactional sequence.

3.2 Culture and Communication

E. Hall's Low Context vs. High-Context Communication Dimension. Besides Hofstede's cultural dimensions model the dimension of *"low-context" vs. "high-context"* cultures introduced by Hall [24] is important in the contexts and situations of intercultural communication. According to Hall [24], in high-context cultures, usually the cultures with high power distance and high collectivism, a member needs to be explicitly asked to respond to elicit behavior that is a deviation from the norm. In low-context cultures, by contrast, members expect to influence others to act by explicitly pointing out pertinent information. The information provided implicitly enables the communicating other to take the desirable decision. Table 5 lists patterns of Hall's cultural communication context dimension. Hall characterizes speaking as an art in high-context cultures, with an emphasis on the emotional aspect. High-context cultures privilege social motivation. Low-context cultures privilege rational information.

Table 5. Low-context vs. high-context cultural communication styles

High-context communication	Low-context communication
Communication is aimed at emotions and rhetorical persuasion	Communicative focus is rational information
Speech is unhurried and long, as persuasion takes time	Information is desired in quantity and expected to be delivered at once
The main emphasis is not laid on the transmission of information, as most of it lies in the context	Decisions are taken on the basis of information
Both speaking and listening are something to be thoroughly enjoyed	Speaker errors carry social costs as they blur information
Ambiguous interpretation is allowed	Unequivocal Interpretation is desired

If the communicative context varies across cultures than it becomes a variable of interest in the learners' interactional accomplishment in online learning situations.

3.3 Culture and Cognition

According to Nisbett and Norenzayan [48], mainstream psychology in general had made four basic assumptions about cognition. Adapting from them, the four foundational psychological assumptions regarding human cognition are:

- *Universality*: Basic cognitive processes of sensation, perception, attention and memory are universal. In other words, basic cognitive processes are invariant across cultures and communities.
- *Content Independence*: Basic cognitive processes are invariant across contents. In other words, cultural differences in content do not affect the nature and structure of the basic cognitive processes.
- *Environmental Sufficiency*: Cognitive processes of general learning and interference operate upon environmental contents to equip the child for functional survival. The environment provides content to cognitive processes without the need for cultural or social interventions. In other words, cultural differences in cognitive processes are due to different environmental influences and not social influences.
- *Infinite Cultural Variance*: Since the universal basic cognitive substrate is content independent and environmentally-sufficient, the range of cultures is a function of the variance in environmental conditions. In other words, cognition places no constraints on the possible evolutionary design space of cultures.

All in all, these four assumptions have led to a belief in a fundamental dissociation between cognition and culture. One consequence of this was that psychology and anthropology evolved into independent academic disciplines with mostly non-overlapping research agendas. However there were some exceptions to this dualist view of cognition and culture. These exceptions include in psychology, Lev Vygotsky and colleagues [62], [64]; in cognitive anthropology, most notably D'Andrade [10]; and in cognitive sociology, Dimaggio [16]. For work on connecting the universal and the specific in language and culture, see [22].

Nisbett and Colleagues' Cross-Cultural Psychology Findings. Table 6, compiled from Nisbett [47] and Nisbett and Norenzayan [48], presents a concise summary of above discussion along with empirical evidence from the literature.

The cultural difference in attention to field vs. object might be relevant to collaborative "knowledge map" learning environments. East-Asian learners might pay attention to a meaningful group of interrelated knowledge map objects whereas Western learners might attend to individual objects and evidential relational links. The cultural difference in attention might vary the ways in which referencing and deixis are carried out in collaborative discourse. East-Asian learners might make more references to regions of the concept maps and groups of related concept map objects (i.e., to fields of interest), whereas Western learners might reference individual objects in their collaborative discourse. This can be translated into a tentative socio-technical design hypothesis that given the choice of referencing regions of concept map areas and individual objects in the concept map, East Asian learners may prefer to appropriate the affordances *for referencing fields*. On the other hand, Western learners may prefer to appropriate the affordances *for referencing individual objects*.

The implications from the cultural difference in perception is that Western learners by virtue of being more susceptible to "primacy effect" might favor earlier perceptions of information related to a collaborative learning task. East-Asian learners

Table 6. Cognitive differences between East-Asians and Westerners

Cultural Profile → **↓Cognitive Process**	**Westerners**	**East-Asians**	**Empirical Evidence**
Attention	*Object* Westerners tend to attend to individual objects in the perceptual field	*Field* East-Asians tend to attend to whole fields rather than individual objects	[43] Field Dependence [68]
Perception	*Object-oriented* Westerners have lesser difficulty in detaching objects from their perceptual contextual fields	*Relation-oriented* East-Asians have difficulty in disentangling an object from its perceptual surroundings	[31] [42]
Causal Inference	*Dispositional* Westerners susceptible to overlooking of situational factors on observed behavior	*Situational* East-Asians susceptible to "hind-sight bias"	[50] [4]
Knowledge Organization	*Categorical* Reliance on categories of objects/events	*Relational* Reliance on relationships between events/objects	[5]
Reasoning	*Analytical* Application of formal logical rules and analytical procedures that emphasize non-contradiction in hypothesis	*Holistic* Willingness to simultaneously entertain several contradicting hypotheses	[49] [48]

might perceive more relationships between the information in concept maps and instructional materials leading to a greater number of evidential relation links in the concept map.

The cultural difference in causal inferences implies that East Asian learners might be more inclined to reason-giving that prioritizes situational factors when compared to the dispositional attributions of Western learners. One particular implication would be the cultural affect on collaborative argumentation. Also, East-Asian learners' perception of their collaborative partners might follow this same trajectory. This might manifest as East-Asians' giving higher ratings for their collaborative peers due to situational attributions explaining any perceived unpleasant "performance." Western learners might perceive their collaborative peers for their dispositional "competence." This cultural difference in cognitive processes might manifest as East-Asian learners preferring a highly inclusive final conclusion in intercultural collaborative problem solving tasks. Western learners might argue for more differentiated analytical hypothesis that seems logically the most viable.

In the next section, a selective review of research into cultural issues in user interface design, usability evaluation, World Wide Web (WWW), information systems (IS), and computer supported cooperative work (CSCW) and online learning is presented.

4 Cultures and Computers

Research into social aspects of HCI [54] has shown that even computer-literate users tend to use social rules and display social behavior in routine interactions with computers. Social interaction is strongly grounded in culture as every person carries within himself/herself patterns of thinking, feeling, behaving and potential interacting. Much of this is learned during development processes in the childhood. As soon as certain patterns of thinking, feeling and acting have established themselves within a person's mind they reside there awaiting activation or inhibition in appropriate situations. To learn new patterns of thinking, feeling and acting one has to unlearn the old patterns, which is more difficult than learning them in the first place for the first time [27].

4.1 User Interface Design

In a recent article, Marcus [39] has expressed how little we know or understand culture as it relates to user-interface usability and design. He uses several examples to highlight the importance of culture and points to recent developments as evidence that culture does matter for user interface design and usability. Fernandes [23] has identified various cultural issues of nationalism, language, social context, time, currency, units of measure, cultural values, body positions, symbols and esthetics that need to be addressed during global interface design. The various solutions suggested include providing ample space for accommodating varying width of date formats and top-level menu design taking the languages with large words into account. Similarly, Russo and Boor [56] present a checklist of cross-cultural items to be considered in interface design. The checklist consists of text, number, date and time formats;

images; symbols; colors; flow and functionality. They discuss each item on the checklist, the problems encountered in practice and propose solutions to avoid them. Russo and Boor place high importance on testing with native users as they can help in identifying the subtler issues of the interface. Khaslavsky [33] describes the impact of culture on usability and design, presents variables useful for incorporating culture into design and mentions issues in localization of design. Elnahrawy has given culture specific recommendations and guidelines for website designers. Elnahrawy concludes that differences between cultures affect the understanding of the websites and calls for further research on cross-cultural effects in user interface design.

4.2 Usability Evaluation

Just as different cultures call for different versions of the same software and products, different usability methods might be needed for different cultures. International usability testing generally involves a usability expert from the product country and a local facilitator in the target country [46]. When differences in cultures exist between the usability test evaluators/facilitators and the test learners, usability assessment techniques employed in the usability testing may mask the usability problems instead of discovering them. Prior research has found that culture affects the usability evaluation process [2], [14], [20], [21], [28], [45], [57], [61], [69]. For example, culture affected the functioning of focus groups [2], the think-aloud protocol [69], questionnaires [14], understanding of metaphors and interface design [21] and efficacy of structured interviews [61]. The role of culture in usability evaluation of online learning environments in particular and HCI in general remains largely unexplored.

4.3 World Wide Web

K-P. Lee [38] conducted a study to identify how cultural characteristics influence people's interaction with products and to evaluate using the WWW as tool for multicultural study. The results of the study showed significant differences between Japan, Korea and USA in interaction styles. Depth of the interface was turned out to be most contributing factor for usability of Korea and USA whereas layout was the most significant factor for Japanese users. Lee concludes that the difference in the cultural characteristics of users accounted for the differences in their interaction styles. Sears et al [57] examined the international differences and effect of high-end graphical enhancements on the perceived usability of World Wide Web. They found significant differences between the users belonging to the two different cultures of United States of America and Switzerland. Sun discusses the localization strategies for building culturally-competent multilingual and transnational websites. Preliminary findings from the exploratory study confirm that contextual clues and cultural markers in website design influence usability and there is a need for integrated localization.

4.4 Information Systems

A research study has shown that culture had a limited impact on some specific aspects of IT decision making [59]. Chong, Yang and Wong [6] examined moderation aspects of cultural differences on the relationships among online trust, perceived value of the goods and services and online purchase intension. They identify the *individualism-collectivism* and *uncertainty avoidance* dimensions of Hofstede's cultural model with the potential to influence trust formation and purchase intension amongst consumers. Chau, Cole, Massey, Montoya-Weiss and O'Keefe [51] explored the cultural differences in the online behavior of consumers from the US and Hong Kong and found significant cultural differences in online consumer behavior across the two countries. Kumar and Bjorn-Andersen [36] did a cross-cultural comparison of IS designer values between business and governmental organizations in Canada and in Denmark. For both countries they found the dominance of technical and economic values and cross-cultural differences in these values between Canada and Denmark. Borchers [40] based on his experiences with software development projects in India, US and Japan notes that cultural differences had a large impact on software engineering. Borchers states that in several instances Hofstede's cultural model closely matched the observations.

4.5 Computer Supported Cooperative Work

Setlock, Fussell and Neuwirth's [58] findings suggest that experimental groups homogenous and heterogeneous cultural backgrounds had different perceptions of the study task. They also report that even though cultural differences were reduced in the computer-mediated communication condition of instant messaging (IM) they were not eliminated. The findings of Setlock et al., contradict those of Anderson and Hiltz [1] who found that culturally heterogenous experimental groups achieved greater consensus using asynchronous group support systems (GSS).

Olson and Olson [53] report a theoretical application of the cultural models to remote software development teams. They identify team composition and teamwork as the two classes of cultural differences that may influence multicultural teams independent of setting. According to Olson and Olson, culture influences both the process and the product of brainstorming, decision support systems, video as well as audio conferencing. Massey, Hung, Montoya-Weiss and Ramesh [41] studied the perceptions of task-technology fit in global virtual teams. They argue that different technologies may be better suited for conveyance of information against convergence to decisions. Their results from the post-study questionnaire showed significant differences in the perceptions of task-technology fit across the various cultures.

Krishna, Sahay and Walsham [34] explored cross-cultural issues in software outsourcing and state that challenges amount to the need to adapt to the cultural norms of social behavior, attitudes towards authority and language usage. The authors citing Brannen and Salk [3] state that "*an attempt to understand and move someway towards the other partner in a cross-cultural collaboration has been called a negotiated cultural perspective.*"

4.6 Online Learning

Kim and Bonk [13] report cross-cultural differences in online collaborative behaviors of the US, Finnish and Korean participants in their study. Daniels, Berglund and Petre

[12] found cultural differences in international projects in undergraduate CS education. McLoughlin [44] based on her experiences with developing web-based instruction for Australian Indigenous education calls for a culturally responsive technology. Iivonen, Sonnenwald, Parma, and Poole-Kober [30] found culturally influenced differences in language and communication styles in a library and information studies course taught over the Internet in Finland and US. Walton and Vukovic's [63] work with south African students from disadvantaged backgrounds found that cultural differences make it difficult for the students to make the transition to the web use.

Crump [9] explored the effects of computing learning environment on the newly arriving international students at universities in New Zealand. The author reports that the cooperative and collaborative learning environment was an issue of concern to the students. The author says it is likely due to the oversimplification of social structure of groups, individual and group goals and the diverse nature of knowledge construction in the collaborative learning environments. Duncker [18] conducted an ethnography of the usability of a library metaphor used in digital libraries in the cultural context of the Maori, who are the indigenous population of New Zealand. Duncker says that metaphors and metaphorical thinking are strongly rooted in culture. The Maori found the digital libraries interesting but difficult to use due to the breakdown of the library metaphor caused by a number of cultural misfits. Keller, Pérez-Quiñones and Vatrapu [32] outlined cultural issues and opportunities in computer science education.

5 Discussion

The cultural issues addressed in most of the prior cross-cultural HCI research studies are at the level of a direct manipulation single user interface and do not consider cultural issues in social interactional design. Symbols, rituals, norms, values and practices are amongst the most visible manifestations of culture, and taken together they cover important aspects of the concept of culture [27]. The cultural issues identified above consider only the symbols and rituals of different cultures, ignoring the basic question of the relationship between culture and affordances. Issues in interface design due to the differences in social and cognitive processes are left unidentified.

Taking the existing body of research on cultural effects on social behavior, cognitive processes, online pedagogies and HCI as the point of departure, we propose a new way of conceptualizing the relationship between cultures and affordances, and are initiating a systematic inquiry into the phenomena that emerge from the nexus of culture, technology and learning. The primary research problem addressed by our current research is the extent to which culture influences appropriation of technology affordances as social affordances and technological intersubjectivity of learners. A systematic investigation of this particular research problem can inform the design of technology affordances. This in turn can increase the efficacy of pedagogical design as measured by individual learning outcomes as well as student perceptions.

References

1. Anderson, W.N., Hiltz, S.R.: Culturally Heterogeneous vs. Culturally Homogeneous Groups in Distributed Group Support Systems: Effects on Group Process and Consensus. In: Proceedings of the 34th Hawaii International Conference on Systems Sciences.
2. Beu, A., Honold, P., Yuan, X.: How to Build Up an Infrastructure for Intercultural Usability Engineering. The International Journal of Human-Computer Interaction 12(3&4), 347–358
3. Campbell, A.: Cultural identity as a social construct. Intercultural Education 11(1), 31–39
4. Choi, I., Nisbett, R.: Situational Salience and Cultural Differences in the Correspondence Bias and Actor-Observer Bias. Personality and Social Psychology Bulletin 24(9), 949
5. Choi, I., Nisbett, R., Smith, E.: Culture, category salience, and inductive reasoning. Cognition 65(1), 15–32
6. Chong, B., Yang, Z., Wong, M.: Asymmetrical impact of trustworthiness attributes on trust, perceived value and purchase intention: a conceptual framework for cross-cultural study on consumer perception of online auction. In: Proceedings of the 5th international conference on Electronic commerce, pp. 213–219.
7. Clark, H.H., Brennan, S.E.: Grounding in communication. In: Resnick, L.B., Levine, J.M., Teasley, S.D. (eds.) Perspectives on Socially Shared Cognition, American Psychological Association, pp. 127–149 (1991)
8. Constantino-González, M., Suthers, D.: A Coached Collaborative Learning Environment for Entity-Relationship Modeling. In: Proceedings of the 5th International Conference on Intelligent Tutoring Systems, pp. 324–333.
9. Crump, B.J.: New Arrival students: mitigating factors on the culture of the computing learning environment. In: Proceedings of the sixth conference on Australian computing education, vol. 30, pp. 49–56.
10. D'Andrade, R.G.: The Development of Cognitive Anthropology. Cambridge University Press, New York (1995)
11. Damen, L.: Culture Learning: The Fifth Dimension on the Language Classroom, p. 367.
12. Daniels, M., Berglund, A., Petre, M.: Reflections on International Projects in Undergraduate CS Education. Computer Science Education 9(3), 256–267
13. Davison, R.: Technical opinion: Cultural complications of ERP. Communications of the ACM 45(7), 109–111
14. Day, D., Evers, V.: Questionnaire Development for Multicultural Data Collection. In: Proceedings of the International Workshop on Internationalisation of Products and Systems (1999)
15. Dillenbourg, P., Baker, M., Blaye, A., O'Malley, C.: The evolution of research on collaborative learning. Learning in Humans and Machines, 189–205
16. Dimaggio, P.: Culture and Cognition. Annual Review of Sociology 23
17. Doise, W., Mugny, G.: The Social Development of the Intellect. Pergamon Press, New York (1984)
18. Duncker, E.: Digital library communities and change: Cross-cultural usability of the library metaphor. In: Proceedings of the second ACM/IEEE-CS joint conference on Digital libraries, pp. 223–230. IEEE Computer Society Press, Los Alamitos
19. Durkheim, E.: The Division of Labor in Society. Free Press, Glencoe, Illinois (1893/1964)
20. Elnahrawy, E.M.: Users from Other Cultures than the U.S (2001)
21. Evers, V., Kukulska-Hulme, A., Jones, A.: Cross-Cultural Understanding of Interface Design: A Cross-Cultural Analysis of Icon Recognition. In: Proceedings of the International Workshop on Internationalization of Products and Systems (Rochester, 1999)

22. Fellbaum, C., Vossen, P.T.J.M.: Connecting the Universal to the Specific: Towards the Global Grid. In: Ishida, T., Fussell, S.R., Vossen, P.T.J.M. (eds.) Intercultural Collaboration I. LNCS, Springer, Heidelberg (2007)
23. Fernandes, T.: Global Interface Design. Academic Press, London (1995)
24. Hall, E.: Beyond Culture. Anchor Press, New York (1976)
25. Hall, E.: The Hidden Dimension. Anchor Press, Garden City, NY (1966)
26. Hofstede, G.: Cultural Differences in Teaching and Learning. International Journal of Intercultural Relations 10(3), 301–320
27. Hofstede, G.: Cultures and Organizations: Software of the Mind, Intercultural Cooperation and its Importance for Survival. McGraw-Hill, New York (1997)
28. Honold, P.: Culture and Context: An Empirical Study for the Development of a Framework for the Elicitation of Cultural Influence in Product Usage. Journal of Human-Computer Interaction 12(3&4), 327–345
29. Hutchins, E.: The social organization of distributed cognition. In: Resnick, L.B., Levine, J.M., Teasley, S.D. (eds.) Perspectives on socially shared cognition, American Psychological Association, Washington DC, pp. 283–307 (1991)
30. Iivonen, M., Sonnenwald, D.H., Parma, M., Poole-Kober, E.: Analyzing and understanding cultural differences: Experiences from education in library and information studies. In: Proceedings of the 64th IFLA General Conference, Amsterdam, Netherlands (1998)
31. Ji, L., Peng, K., Nisbett, R.: Culture, control, and perception of relationships in the environment. Journal of Personality and Social Psychology 78(5), 943–955
32. Keller, B., Pérez-Quiñones, M.A., Vatrapu, R.: Cultural Issues and Opportunities in Computing Education. In: 9th International Conference on Engineering Education, San Juan, PR, 2006 (July 23-28, 2006)
33. Khaslavsky, J.: Integrating Culture into Interface Design. In: CHI 98 conference summary on Human factors in computing systems, pp. 365–366.
34. Krishna, S., Sahay, S., Walsham, G.: Managing cross-cultural issues in global software outsourcing. Communications of the ACM 47(4), 62–66
35. Kroeber, A.L., Kluckhohn, C.: Culture: A Critical Review of Concepts and Definitions. Harvard University Peabody Museum of American Archeology and Ethnology (1952)
36. Kumar, K., Bjorn-Andersen, N.: A cross-cultural comparison of IS designer values. Communications of the ACM 33(5), 528–538
37. Lederach, J.P.: Preparing for peace: Conflict Transformation Across Cultures. Syracuse University Press, Syracuse, NY (1995)
38. Lee, K.-P.: A Study on the Cultural Effects on User Interface Design (2004)
39. Marcus, A.: Culture: Wanted? Alive or Dead? Journal of Usability Studies 1(2), 62–63
40. Marcus, A.: Fast forward: User-interface design and China: a great leap forward. Interactions 10(1), 21–25
41. Massey, A.P., Hung, Y.-T.C., Montoya-Weiss, M., Ramesh, V.: When culture and style aren't about clothes: perceptions of task-technology "fit" in global virtual teams. In: Proceedings of the 2001 International ACM SIGGROUP Conference on Supporting Group Work, pp. 207–213. ACM Press, New York (2001)
42. Masuda, T., Nisbett, R.: Culture and change blindness. Cognitive Science 30, 1–19
43. Masuda, T., Nisbett, R.E.: Attending holistically versus analytically: Comparing the context sensitivity of Japanese and Americans. Journal of Personality and Social Psychology 81(5), 922–934

44. McLoughlin, C.: Culturally responsive technology use: developing an on-line community of learners. British Journal of Educational Technology 30(3), 231–243
45. Moon, Y., Nass, C.: Adaptive Agents and Personality Change: Complementarity versus Similarity as Forms of Adaptation. In: Proceedings of the CHI '96 Conference Companion on Human Factors in Computing Systems: Common Ground, pp. 287–288.
46. Nielsen, J.: International Usability Engineering. In: Del Galdo, E.a.N.J. (ed.) International User Interfaces, John Wiley & Sons, Newyork (1996)
47. Nisbett, R.E.: The Geography of Thought: How Asians and Westerners think differently.. and Why. Free Press, New York (2003)
48. Nisbett, R.E., Norenzayan, A.: Culture and Cognition. In: Medin, D.L. (ed.) Stevens' Handbook of Experimental Psychology, pp. 561–597 (2002) (Page numbers used in the document refer to the manuscript retrieved on August 531, 2006 from http://www-personal.umich.edu/ nisbett/cultcog2002.pdf)
49. Nisbett, R.E., Peng, K., Choi, I., Norenzayan, A.: Culture and systems of thought: Holistic versus analytic cognition. Psychological Review 108(2), 291–310
50. Norenzayan, A., Nisbett, R.: Culture and Causal Cognition. Current Directions in Psychological Science 9(4), 132–135
51. Norton, D.: Development Consortium: Implementation of an electronic report viewing application for multi-cultural users. In: CHI '02 extended abstracts on Human factors in computing systems, pp. 534–535.
52. Olson, G.M., Olson, J.S.: Distance Matters. Human-Computer Interaction, 15 (2/3). (reprinted). In: Carroll, J.M. (ed.) Human-Computer Interaction in the New Millennum, pp. 397-417. ACM Press, New York.
53. Olson, J.S., Olson, G.M.: Culture Surprises in Remote Software Development Teams. Queue 1(9), 52–59
54. Reeves, B., Nass, C.: The media equation: how people treat computers, television, and new media like real people and places. Cambridge University Press, New York, NY, USA (1996)
55. Roschelle, J., Teasley, S.: The construction of shared knowledge in collaborative problem solving. In: O'Malley, C.E. (ed.) Computer supported collaborative learning, pp. 69–97. Springer, Heidelberg (1995)
56. Russo, P., Boor, S.: How Fluent is your Interface? Designing for International Users. Proceedings of the conference on Human factors in computing systems, 342–347
57. Sears, A., Jacko, J.A., Dubach, E.M.: International Aspects of World Wide Web Usability and the Role of High-End Graphical Enhancements. International Journal of Human-Computer Interaction 12(2), 241–261
58. Setlock, L., Fussell, S., Neuwirth, C.: Taking it out of context: collaborating within and across cultures in face-to-face settings and via instant messaging. In: Proceedings of the 2004 ACM conference on Computer supported cooperative work, pp. 604–613. ACM Press, New York
59. Teng, T.C.J., Calhoun, K.J., Cheon, M.J., Raeburn, S., Wong, W.: Is the East Really Different from the West? In: Proceeding of the 20th international conference on Information Systems, pp. 40–46.
60. Tylor, E.: B. Primitive Culture: The Science of Culture. In: Bohannan, P., Glazer, M. (eds.) High points in anthropology, McGraw-Hill, New York (1903/1988)
61. Vatrapu, R., Pérez-Quiñones, M.: Culture and Usability Evaluation: The Effects of Culture in Structured Interviews. Journal of Usability Studies 1(4), 156–170
62. Vygotsky, L.: Mind in society. Harvard University Press, Cambridge (1930/1980)

63. Walton, M., Vukovic, V.: HCI in the developing world: Cultures, literacy, and the web: dimensions of information "scent". Interactions 10(2), 64–71
64. Wertsch, J.: Vygotsky and the social formation of mind. Harvard University Press, Cambridge, MA, USA (1985)
65. White, L.A.: Energy and the evolution of culture. American Anthropologist. In: McGee, R., Warms, R. (eds.) Anthropological theory: an introductory history. Mayfield Pub. Co. (1943/1996) (reprinted)
66. Williams, R.: Keywords: A Vocabulary of Culture and Society. Oxford University Press, New York (1983)
67. Williams, R.: Marxism and literature. Oxford University Press, Oxford (1977)
68. Witkin, H., Goodenough, D.: Field dependence and interpersonal behavior. Psychol Bull 84(4), 661–689
69. Yeo, W.A.: Global-Software Development Life Cycle: An Exploratory Study. In: Proceedings of the SIG-CHI Conference on Human Factors in Computing Systems, pp. 4–11.

Adoption of Translation Support Technologies in a Multilingual Work Environment

Jahna Otterbacher

Department of Public and Business Administration
University of Cyprus
P.O. Box 20537
CY-1678 Nicosia, Cyprus
jahna@ucy.ac.cy

Abstract. We study the adoption of translation support technologies by professors at a multilingual university, using the framework of the Technology Adoption Model (TAM). TAM states that a user's perceived usefulness and ease of use for the technology ultimately determines her actual use of it. Through a survey and a set of interviews with our subjects, we find that there is evidence for TAM in the context of translation support tools. However, we also find that user adoption of these tools is a bit more complicated. Users who are able to successfully employ these tools have not only developed strategies to overcome their inaccuracies (e.g. by post-editing machine translated text), they also often compensate for the weaknesses of a given technology by combining the use of multiple tools.

Keywords: Technology Acceptance, Machine Translation, Electronic Dictionaries, Multilingual Environments.

1 Introduction

The development of technologies that support translation, such as electronic multilingual dictionaries and machine translation systems, has the potential to enhance a great number of intercultural, multilingual interactions and activities. Such technologies can be used for the purposes of information assimilation, dissemination and collaboration. For example, translation support tools allow users to access online material in languages that they do not actually know [23]. Even in cases where the current technologies cannot fully enable the user to comprehend a given text, he or she can often grasp the main ideas expressed. This allows one to judge whether or not the text is of interest, and should be passed on to a human translator, thus saving time and resources [8]. Likewise, translation tools enable users to disseminate their work in languages in which they are not fully fluent [23]. Finally, they also facilitate the interaction and collaboration of individuals who do not speak a common language [2] or who do not share the same native language and thus require support in order to communicate effectively in a common second language [12,18]. This in turn might help to mitigate some of

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 276–290, 2007.

the social and cognitive challenges associated with collaborating in international and multicultural settings [20].

In general, translation technologies offer several potential benefits to the user who needs to overcome a language barrier in his or her work. According to Theologitis [21], some of these include enhancing a user's productivity, by improving the speed and accuracy with which he or she can translate desired information into a target language. In addition, translation tools are more consistent than humans, since they employ linguistic rules or statistical models that ensure they will always return the same translation for a given input word or phrase in a particular context. Another benefit that is more difficult to characterize or measure, is the possibility that such tools will enable a user to communicate in another language such that she will be exposed to new information, ideas and people, thus enriching her work.

Despite these potential advantages, the adoption of translation technologies, and in particular, of machine translation systems, is challenged by a number of factors. For example, humans often reject the raw output of a system, such that the translated text requires manual editing in order to be considered acceptable [5]. In a simultaneous collaboration environment (e.g. chat) this means that users need to monitor the quality of the output, in order to ensure their messages will be understandable to the other parties [12]. Another potential issue is that because of the intense processing required to translate an input text, some users find that automatic translators are slow, and not easy to integrate with the rest of one's computing environment and work routine [23].

Nonetheless, adopters of translation technologies can learn to use them successfully and productively. Users adapt to the tools, and often develop strategies for getting them to perform better. For instance, upon receiving a translated sentence that is not entirely satisfactory, users often alter the original input sentence and rerun the translation, using repairs that are somewhat predictable given the language pairs involved [19]. In fact, vendors of automatic translation systems have noted that getting such a system to work properly is a process of user adaptation and training and that users should not expect "quick fix" solutions [16]. Users need to invest time at updating their translation system's dictionary and underlying linguistic tools as well as in enriching the source texts, in order to receive higher quality output.

1.1 Adoption of Translation Support Technologies

The goal of the current paper is to better understand users' decisions to adopt technologies that support translation in a multilingual work environment. Previous research has most often focused specifically on the use of machine translation systems. Research carried out by translation system vendors has most often considered what people do with a system (e.g. what kinds of texts they are likely to translate and between which languages) and their subjective evaluation of the given system (e.g. [9,23]). To contrast, other researchers have examined the adaptive strategies that users employ, such as backtranslation and post-editing, in order to improve translation performance (e.g. [19]). However, we are not aware of any previous studies that have specifically focused on examining the factors that influence one's decision to adopt translation technologies or not, and how these technologies fit into a user's work routine. In addition, we are not aware of any research that examined how people in a multilingual work environment ultimately find solutions to their translation needs (e.g. whether one technology supports all of their needs or if they rely on multiple technologies, their colleagues, etc.).

It has previously been noted that evaluating the quality of translation systems is difficult in part because system performance varies for different applications and users [16]. In other words, translating a technical document is a very different process than translating instant messages. Therefore, in the current paper, we examine the acceptance of translation support technologies in a specific professional context, in which the subjects are likely candidates for adoption of translation technologies. In particular, we will study whether or not professors in a multilingual university environment adopt translation support technologies. In addition, we include not only the use of automatic translation systems but also electronic multilingual dictionaries and termbanks. This is because it may be the case that in their work routine, users rely on a number of translation support tools in order to express themselves effectively in a target language.

The paper will proceed as follows: in Section 2, we provide some background about our subjects and their work environment, and in addition, we introduce our theoretical framework. Next, in Section 3, we describe a survey that was administered to tenure track faculty members at the University of Cyprus (UCY). Finally, Section 4 discusses the insights gained from interviews with five UCY faculty members, in which they were asked about how translation support technologies might fit into their work routines.

2 Research Context and Approach

Our goal is to examine the factors determining the adoption of translation support technologies. Thus, we studied subjects who are likely candidates for adoption of such tools - tenure track faculty members at a multilingual institution, the University of Cyprus (UCY). While the Republic of Cyprus itself has three official languages (Greek, Turkish and English), the language of instruction at the university is officially Greek for undergraduate programs, with many postgraduate programs offered in English. However, English textbooks and teaching materials are widely used, at both the undergraduate and graduate levels. In addition, administrative documents (e.g. internal reports, requests for research and travel expenses) are generally to be written in Greek. Finally, the faculty is international and the majority of tenure-track professors have doctoral degrees and post-doctoral experience from non-Greek speaking universities.

Since their expertise has been gained abroad while studying and working in a language other than Greek, and due to the widespread use of English language textbooks, we hypothesize that UCY professors are candidates for the adoption of translation support technologies. As noted by Birbili [3], having to explain knowledge one gained while working in another language is challenging, because decisions made in translating information have a direct impact on its validity. It is expected that our subjects have to make such decisions both in asynchronous settings (e.g. in preparing course materials, in writing administrative reports and funding requests for one's research activities) as well as in synchronous situations (e.g. in dialogs with students). In addition, it is likely that establishing common ground with their students and colleagues [6] is made more challenging since they need to communicate concepts (e.g. from an English textbook) and experiences across languages. Therefore, in addition to studying whether or not they have adopted translation support tools in their work, we will also consider how they use such tools and in which activities.

2.1 User Acceptance of Translation Support Technologies

Previous work has attempted to determine which factors influence users' adoption or acceptance of information technologies. In a survey of experts who work in the area of technology development, Kolo and Friedewald [13] determined that the most important factors for user acceptance are the speed and timeliness of a technology, and its robustness and reliability. Similarly, in the context of machine translation systems, Flanagan [9] found that the main system requirements cited by users were speed, robustness and language coverage. In addition, McClure and Flanagan [16] have noted that the lack of reliable means to evaluate translation quality is an obstacle to widespread adoption of these technologies. They noted that automatic evaluation procedures (e.g. counting the number of words or sentences translated in an acceptable manner) are problematic and that manual evaluation by humans is both time consuming and subjective, since users often disagree on the extent to which a given translation is acceptable. They also point out that an increase in intrinsic accuracy of a translation system (e.g. the proportion of sentences translated acceptably) does not necessarily entail a more satisfying experience overall for users. This disconnect between intrinsic accuracy and the quality of the user's experience has also been noted in other fields of nature language technology (e.g. automatic character recognition [10] and information retrieval systems [22]). Lastly, users' expectations of a given technology also affect their experiences and opinions of its performance. Particularly for translation systems, Fenn explains that users need to be trained in what to expect from the technology, as to not become disappointed by the fact that automatic systems cannot mimic human translation performance [8].

2.2 Davis' Technology Adoption Model

On a more abstract level, Davis [7] put forward the Technology Adoption Model (TAM). According to TAM, the factors that determine whether or not someone uses a particular technology are his or her perceived usefulness and ease of use of the technology. These perceptions impact the potential user's attitudes towards the technology, which is turn leads to behavioral intentions (i.e. to adopt it or not) that determine actual use. Following Lederer and colleagues [14], we adopt a more brief version of the TAM as a framework for the current study. Rather than trying to study user intentions, we will concentrate on examining the direct effect of subjects' perceived ease of use and usefulness on their actual usage of translation support technologies. In other words, we will not attempt to characterize how a user's attitude towards the technology influences his or her intention to use it. Figure 1 depicts the model and an example of possible antecedents of usefulness and use that might effect users' perceptions, and in turn, their use of translation technology.

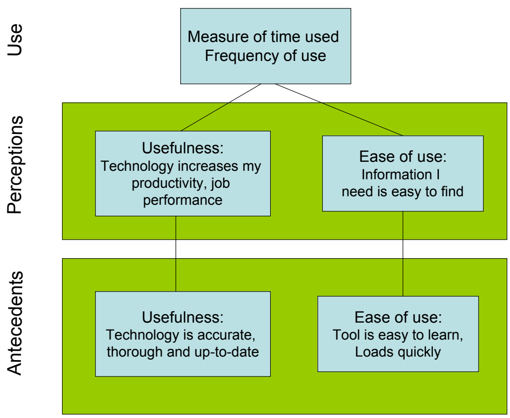


Fig. 1. Davis' [7] Technology Acceptance Model

3 Perceptions and Use Behaviors: Faculty Survey

Using the TAM framework, we developed a Web-based survey in order to examine faculty members' perceptions of translation support technologies and their actual usage. The survey questions are shown in Tables 1 (introductory questions) and 2 (items measuring respondents' perceived ease of use, perceived usefulness, and actual usage of translation support technologies). The subjects responded to the items in Table 2 using a scale from 1-7 with one indicating that they "strongly disagree" with the statement and 7 meaning that they "strongly agree."

Following Lederer and colleagues [14], who used TAM in order to study the usage of Web sites, we included the ease of use items and usefulness items from Davis' original study [7]. In addition, to measure perceived ease of use, we included items tailored for this study from Boling's work on usability testing [4]. Finally, the items testing perceived usefulness include dimensions of the usefulness of information [11] as well as the requirements necessary for successful technology adoption as identified by Kolo and Friedewald [13].

A definition of translation support technologies was prepared and six categories of tools were described: automatic translators on a local computer, automatic translators on a Web site, multilingual dictionaries on a local computer,

Table 1. Survey questions: introductory items

Background
In which department do you work? What is your most dominant language? For how many years have you been in your current faculty position? Have you ever used a Translation Support Tool in the context of any work-related task?
Usage Pattern Items (for respondents who answered that they have used translation technologies)
How often have you used the following technologies in your work? (Response choices: Never, Once or twice, A few times, Regularly, Daily) 1. Automatic translation on my computer 2. Automatic translation on a Web site 3. Multilingual dictionary on my computer 4. Multilingual dictionary on a Web site 5. Specialized terminology dictionary on my computer 6. Specialized terminology dictionary on a Web site
For which work-related activities have you used the following technologies? (Response choices: Teaching, Research, Administration, Communications, Other) 1. Automatic translation on my computer 2. Automatic translation on a Web site 3. Multilingual dictionary on my computer 4. Multilingual dictionary on a Web site 5. Specialized terminology dictionary on my computer 6. Specialized terminology dictionary on a Web site
With which of the above 6 technologies do you have the most experience?

Table 2. Survey questions: perceived ease of use, usefulness and actual use

TAM items measuring ease of use [7]
Getting the information I need is easy with this technology. Learning to use this technology is easy. Becoming skillful at using this technology is easy.
Items measuring antecedents to ease of use - Usability testing criteria [4]
I find that translation support tools on Web sites load too slowly. When using translation support tools on Web sites, the information I need is not easy to find. I find that translation support tools on Web sites are difficult to navigate. I find that graphics on translation support tools are difficult to understand. I find that the interfaces used in translation support tools are difficult to read and understand. The information I need is not easy to find on the interfaces of translation support tools.
TAM items measuring usefulness [7]
Using this technology enhances my effectiveness on the job. Using this technology in my work increases my productivity. Using this technology improves my job performance.
Items measuring antecedents to usefulness and characteristics of useful information [11,13]
I use this technology for ACCURATE information in my work. I use this technology for THOROUGH information in my work. I use this technology for TIMELY information in my work. I use this technology for RELEVANT information in my work.
Reported actual use
During an average academic year, approximately how many days per week do you use translation support technology? During the last 30 days, on how many days did you use a translation support technology?

multilingual dictionaries on a Web site, specialized terminology dictionaries on a local computer, and specialized terminology dictionaries on a Web site. (Since several previous studies reported that a technology's speed and timeliness is a factor that affects adoption (e.g. [8,13]), it was thought best to make a distinction between Web-based tools and those installed on a local machine). The survey's introduction also provided specific examples of translation support tools that fall into each of the six categories. Finally, the survey instrument was tested on two pilot subjects in order to ensure clarity of the directions and questions.

3.1 Demographics of Subjects and Respondents

An email was sent to all tenure-track faculty members in four schools (Engineering, Humanities, Pure and Applied Sciences, and Economics and Management) inviting them to participate in the survey. As reported on the UCY Web site, of the 116 faculty members in these four schools, 84% have doctoral degrees from English-speaking universities, 8% from Greek universities, 6% from German-speaking institutions, and 2% from French universities.

Of the 116 tenure track professors to whom the email was sent, 46 completed the survey, for a completion rate of 40%. While achieving high response rates in online surveys is a known challenge, some research has claimed that response quality is better in Web-based surveys as compared to traditional methods [15,17]. In this particular case, it is likely that computer-savvy individuals were more willing to complete the survey and that these same people are more likely in turn to be users of translation support technologies. Thus, it is important to note that response bias in this study may be an issue for the population of non-users of translation technologies.

In terms of academic discipline, the responses were rather balanced, with 11 respondents in engineering, 11 in languages, 9 in pure in applied sciences (including chemistry, physics and computer science) and the remaining 15 in one of the social sciences disciplines. Two-thirds (31) were users of translation support tools, while one-third stated that they did not use these technologies. Interestingly, all 11 engineers were users of translation tools, while in the other three schools there were both users and non-users.

3.2 Which Technologies are Used?

As previously mentioned, respondents were asked about three translation support technologies: automatic translation systems, electronic dictionaries and electronic dictionaries for specialized terminology. However, we distinguished between the use of these technologies on the user's local computer versus use on a Web site. As shown in Table 3, Web-based technologies appear to be more popular with the subjects. For example, while 60% of the subjects reported using an online machine translation system, only 20% had used one installed on their personal computer. This trend holds for all three tools. Another finding shown in the table is that Web-based multilingual dictionaries is the translation support technology that is used by the most subjects (87%). In fact, Web-based dictionaries was also the tool with which subjects reported having the most experience (39.3 %). This was followed by the use of automatic translation on a Web site (21.4%).

Table 3. Use of translation support on local machine versus web site: % of adopters who use technology

Technology	Local machine	Web site
Automatic translation system	20%	60%
Multilingual dictionary	60%	87%
Terminology dictionary	30%	73%

3.3 Hypotheses: The Use and Usefulness Connection

We now turn to examining some hypotheses that concern which subjects are the likely users of translation support technologies, and the relationship between perceptions of usefulness and ease of use and the actual reported use of these technologies.

H1: Users of translation technologies have not been in their academic positions as long as non-users.

We hypothesize that newer faculty members are more likely to adopt translation support technologies in their work. Particularly, if they have just arrived in Cyprus and are coming from a monolingual, non-Greek speaking working environment abroad, it seems reasonable that they may rely on translation tools more as they are building up their knowledge of technical terminology. To answer this question, we compared the average number of years that translation support users reported being in their current faculty positions versus the average among the non-users.

The distribution of years on the job is skewed for both the users and non-users of translation technology. The median number of years at UCY for the users is 2.25, while the median for the non-users is 7 years. We conducted a Mann-Whitney-Wilcoxon test in order to see if there is a statistically significant difference in the average number of years on the job between the two groups [1]. The test confirmed that this difference is in fact significant, with a p-value of 0.044.

H2: Users of translation technologies perceive them as being more useful as compared to non-users.

To address this question, we considered both the usefulness items on the survey as well as those measuring the antecedents of usefulness. For both, we created a composite score for each subject, which was the mean of the three usefulness items and of the four antecedents items, respectfully. As the composite scores follow an approximately normal distribution for both the users and non-users, we compared the two groups using the small sample t-test of means. Table 4 shows the mean composite scores (on the 7-point scale) for users and non-users, as well as the corresponding p-values for the differences in means. As can be seen, the differences are statistically significant for both perceived usefulness and its antecedents, with the users of translation technologies perceiving them as more useful than the non-users.

Table 4. Comparison of mean perceived usefulness among users and non-users

	Users	Non-users	p-value for difference
Perceived Usefulness	4.622	3.667	0.039
Antecedents to Usefulness	4.725	3.804	0.023

H3: Users of translation technologies perceive them as being more easy to use as compared to non-users.

Similarly, in order to compare the users and non-users with respect to their perceived ease of use, we examined both their perceptions and the antecedents to these perceptions. As can be seen in Table 5, we obtained some unexpected results. The users of translation technologies perceive them as being significantly more easy to use than do the non-users. However, with respect to the antecedents to ease of use, which is made up of items designed to measure respondents' ease of use for information technologies in general, we see that the non-users actually appear to perceive technologies in general to be more easy to use than do the users of translation support tools. Therefore, next, we will examine the correlation between ease of use and its antecedents. We will also compare the correlation between perceived usefulness and its antecedents.

Table 5. Comparison of mean perceived ease of use among users and non-users

	Users	Non-users	p-value for difference
Perceived Ease of Use	6.03	4.67	0.003
Antecedents to Ease of Use	3.56	5.01	0.0002

H4: There is a correlation between the antecedents of ease of use and perceived ease of use.

H5: There is a correlation between the antecedents of usefulness and perceived usefulness.

Table 6 shows the correlation between our variables of interests, and their antecedents. As can be seen, there appears to be no significant linear correlation between the antecedents of ease of use and the subjects' perceived ease of use. As mentioned previously, the antecedent ease of use items attempt to gauge how comfortable subjects are with information technologies in general. For example, the subjects are asked how difficult it is for them to find information on Web pages, and how difficult it is for them to understand graphical software interfaces. To contrast, the perceived ease of use items ask specifically about the usability of translation support tools. It may be the case that the antecedent items are not important determinants of our subjects' ease of use. In addition, it may be possible that in general, ease of use is not an important factor because our subjects (both users and non-users of translation technologies) have a similar level of comfort and ability with information technologies in general. These questions will be examined again later in the paper.

To contrast, in Table 6, we see that there is a significant correlation between the subjects' perceived usefulness and the antecedents of usefulness. As previously mentioned, the antecedents of usefulness survey items included questions related to the quality of information provided by translation technologies, as well as its speed and reliability.

Table 6. Correlation between antecedents and perceived ease of use and usefulness

	r	p-value
Perceived Ease of Use Antecedents of Ease of Use	0.24	0.115
Perceived Usefulness Antecedents to Usefulness	0.54	0.00016

H6: There is a significant relationship between perceived usefulness and actual use.

To evaluate this hypothesis, we regressed each of the measures of the subjects' actual use of translation technologies on the two composite variables for usefulness and the antecedents of usefulness. As can be seen in Table 7, the first measure of actual technology use, the number of days the subjects use the technology in a typical work week, is not significantly related to subjects' perceived usefulness. To contrast, usefulness is significantly correlated to the second measure of use, which was reported as the number of days that translation support tools were used during the last month (i.e. the month before the survey was administered). One possible explanation for this might be that subjects are more accurate at reporting their usage in a specific time period than they are at giving an estimate of average usage during a "typical week."

Table 7. Relationship between perceived usefulness and measures of actual use

	F	p-value
Ave. Days per Week	0.81	0.452
Freq. Last 30 Days	6.53	0.0034

H7: There is a significant relationship between perceived ease of use and actual use.

Similarly, to address H7, we regressed the measures of use onto the perceived ease of use and the antecedents of ease of use. The trend here was the same as in the case of perceived usefulness. Table 8 shows that the relationship between the average number of days in an average week that the translation tools are used and the perceived ease of use is not significant. However, the relationship of ease of use to the frequency of usage during the last month is statistically significant, with a p-value of 0.033.

Table 8. Relationship between perceived ease of use and measures of actual use

	F	p-value
Ave. Days per Week	2.17	0.127
Freq. Last 30 Days	3.72	0.033

4 Dimensions of Perceived Ease of Use and Usefulness: Interviews with Faculty Members

We conducted semi-structured interviews with five faculty members, using the protocol shown in Table 9. Subjects A and B, who were both professors in the School of Economics and Management, self-identified as non-adopters of translation technologies. Subjects, C, D and E, who were engineers, were users of the technologies. Here, we briefly describe the insights gained from the interviews. All of the subjects were native Greek speakers, except for subject E, who was a native speaker of German.

Table 9. Interview protocol

Respondent background
What is your dominant language? In which language(s) did you complete each of your higher degrees? In which language(s) have you previously worked (state number of years)? In which language do you feel most comfortable teaching? In which language do you feel most comfortable writing technical articles? Please describe the activities performed during a typical work week, and the languages in which they are conducted.
Teaching
Please describe the courses you have taught at UCY (e.g. level, required/elective). What are the primary instructional materials used? Do you use any translation support tools in preparing teaching materials? If yes, how do you use them? If no, how do you manage?
Research and collaboration
Do you require translation support in any collaborative projects? Do you require translation support in any of your research activities? Do you require translation support for communications?
Opinions about translation support tools
Overall, how well do you feel translation support tools perform (strengths/weaknesses)? How do you use these tools (e.g. to translate complete documents, individual terms, etc.)? Have you developed any strategies for how to make these technologies perform better?
Other
Do you have any other insights or comments about translation technologies that you would like to share?

4.1 Subject A: Self-identified Technophobe

Subject A had been in her position for four years. She completed all of her higher education in the U.S., and now teaches both undergraduates and graduate students at UCY, in Greek. However, she reported conducting all of her research and administrative activities in English. When asked about her use of technology support technologies, she quickly confessed to being a "technophobe, who doesn't even like to operate the DVD player!" She was not aware of translation support technologies before being asked about them in the context of this research, and said that she probably would have never found out about them. She said that they most likely would be useful in the context of her work. When she needs to translate in the context of her teaching responsibilities, her strategy is to first check her traditional Greek-English dictionary. If that fails, she ask colleagues for help.

4.2 Subject B: Perceived Troubles Exceed Expected Benefits

Subject B had tried out a Greek-English dictionary installed on her computer and was “very impressed at how accurate it was for translating individual words.” She had also purchased a translation system for a family member, but decided not to try it out herself because it came with 3 CDs to install, and appeared to be “too complicated and too much trouble.” Hypothetically, she said that translation support tools would be very useful for her work, however, she already has “an established routine” after being in her position for 10 years. She doubted that people like herself were likely candidates for adoption, since “people don’t upset their work routines to use something new unless they really have a need for it.”

4.3 Subjects C and D: Avid Translation Technology Adopters

Two of the engineers had very similar responses. They were both rather new to the job (being at UCY for 0.5 and 2.5 years, respectively) and reported relying on translation support technologies on a daily basis. Both conducted their research exclusively in English, although they both occasionally write funding proposals in Greek. In addition, they both taught undergraduate courses in Greek.

Subject C reported a specific strategy that he uses to translate a document such as a grant proposal. He uses an online translation system, and works paragraph-by-paragraph. He enters a paragraph into the system, and then manually post-edits it. “Technical terms are always, always wrong.” Therefore, he uses an online terminology dictionary to find the correct Greek term. Although he felt that his strategies were working well, he said that he would be willing to pay a lot of money for more accurate translation systems. He also cited a need for more resources for translating technical terms. In general, he feels that translation with currently available technologies is very time-consuming.

Subject D also noted that translation systems are “terrible with the terminology.” His strategy for translating a document is to work sentence-by-sentence. He uses an online system to translate a given sentence, and then has to “cross-check various sources in order to come up with the correct terms.” He reported using a translation system, a dictionary and a thesaurus, all online, on a daily basis. After being on the job for 2.5 years, he says that he now feels that he has improved his Greek and that he will use the tools less and less often as time goes on. He also felt that he now prefers teaching in Greek, whereas when he first started at UCY, he would have much preferred to teach in English.

4.4 Subject E: Balanced in Two Languages

Subject E carries out his work in English (research and teaching) and German (research and communications). He uses Greek for some administrative purposes and will start teaching in Greek soon. E reported that he typically uses an online German-English dictionary for finding individual terms. He is very satisfied with this technology. He used it much more frequently in the past when he first moved to an English working environment from Germany, and then had to use it when he went back to a German-speaking environment after several years. However, he now feels that his language abilities are very balanced and that his knowledge of the terminology is good. However, he anticipates using more translation support tools once he begins to teach in Greek as he expects them to be useful as well as easy to use.

5 Conclusion

Here, we summarize the main findings of the current study.

1. The adopters of translation support technologies have typically been in their positions for a shorter period of time than have the non-users. In addition, their translation support needs change/lessen as they are in their positions longer and gain more knowledge of the target language.
2. The main TAM effects of the subjects' perceived ease of use and usefulness on their actual usage behaviors were confirmed. However, the relationship was only significant when usage was measured as the "number of days that translation technologies were used in the last 30 days."
3. One of the indirect relationships according to TAM was not confirmed. The antecedents of the ease of use were not significantly correlated to the ease of use.
4. Successful adopters of translation support technologies use more than one tool to get their work done. In particular, adopters of machine translation systems can supplement their weaknesses by post-editing and by use of a terminology dictionary.

There are some limitations of the current study that should be considered when interpreting the findings. First, the survey response rate was 40%, and should be improved in future work. In particular, we need to elicit participation from non-adopters of translation tools. Another factor is that the amount of usage was self-reported, and thus, subject to error. Finally, it should be noted that there are many other statistical techniques that we could have used to examine the relationships between ease of use, usefulness and actual use. However, we leave these for future work, when we plan to carry out a larger study.

In conclusion, we have attempted to characterize the adoption of translation support technologies, using the TAM as our theoretical framework. Overall, we found that how subjects view the "usefulness" of these technologies is dynamic, and not only has to do with how well they perceive a technology to work, but also how the tool fits into their work routine. Another finding was that the way successful adopters of this technology employ them is rather complex. They often have developed both strategies (e.g. post-editing) as well as a means for compensating the weaknesses of a given tool (e.g. cross-checking with another tool). Future work in the area of the evaluation of such technologies should bear this in mind, as studies focusing on the performance or acceptance of a single tool may not shed light on the performance of these technologies as they are actually being used.

Acknowledgments. The author would like to thank the faculty members at the University of Cyprus for their participation in the study.

References

1. Anderson, D., Sweeney, D., Williams, T.: Statistics: Concepts and Applications. West Publishing Company. St. Paul, MN (1994)
2. Berner, B.: Lost in Translation: Cross-lingual Communication and Virtual Academic Communities. In: Proceedings of the 5th Annual Conference on the World Wide Web Applications. Durban, South Africa (2003)
3. Birbili, M.: Translating from One Language to Another. Social Research Update, Issue 31. Department of Sociology. University of Surrey, Guildford, England (2000)
4. Boling, E.: Usability testing for Web sites. In: Learning for the Global Community: Seventh Annual Hypermedia Conference. Bloomington, Indiana (1995)
5. Bowker, L., Ehgoetz, M.: A Pilot Study Exploring User Acceptance of Machine Translation. In: Proceedings of the 3rd Computational Linguistics in the North-East Workshop. University of Quebec (2005)
6. Clark, H.H., Brennan, S.E.: Grounding in Communication. In: Resnick, L.B., Levine, J.M., Teasley, S.D. (eds.) Perspectives of Socially Shared Cognition. Washington, DC (1991)
7. Davis, F.D.: Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. In: Management Information Systems Quarterly, vol. 13(3), pp. 319–339. Management Information Systems Research Center. University of Minnesota (1989)
8. Fenn, J.: Corporate Deployment of Machine Translation: Case Study. Gartner Research Case Studies, CS-17-9077 (2002)
9. Flanagan, M.: Two Years Online: Experiences, Challenges and Trends. In: Expanding Machine Translation Horizons: Proceedings of the Second Conference of the Association for Machine Translation in the Americas. Montreal, Quebec, Canada, pp. 206–211 (1996)
10. Frankish, C., Hull, R., Morgan, P.: Recognition Accuracy and User Acceptance of Pen Interfaces. In: Proceedings of the Conference on Human Factors in Computing Systems (ACM CHI '95). Denver, Colorado (1995)
11. Griffin, R.W.: Management. Houghton Mifflin, Boston, MA (1990)
12. Ishida, T.: Language Grid: An Infrastructure for Intercultural Collaboration. In: IEEE/IPSJ Symposium on Applications and the Internet (SAINT-06). Phoenix, Arizona (2006)
13. Kolo, C., Friedewald, M.: What Users Expect from Future Terminal Devices: Empirical Results from an Expert Survey. In: Kobsa, A., Stephanidis, C. (eds.) User Interfaces for All: Proceedings of the 5th ERCIM Workshop. Sankt Augustin (1999)
14. Lederer, A.L., Maupin, D.J., Sena, M.P., Zhuang, Y.: The Technology Acceptance Model and the World Wide Web. In: Decision Support Systems, vol. 29, pp. 269–282. Elsevier Science, Amsterdam (2000)

15. MacElroy, B., Mikucki, J., McDowell, P.: A Comparison of Quality in Open-end Responses and Response Rates Between Web-based and Paper and Pencil Survey Modes. Journal of Online Research. Interactive Marketing Research Organization (2002)
16. McClure, S., Flanagan, M.: Overcoming the Machine Translation Quality Impasse. IDC Research White Paper 1, 29917 (2003)
17. Nicholls II, W.L., Baker, R.P., Martin, J.: The Effect of New Data Collection Technologies on Survey Data Quality. In: Lybert, L., Biemer, M., Collins, M., DeLeeuw, E., Dippo, C., Schwarz, N., Trewin, D. (eds.) Survey Measurement and Process Quality, Wiley, New York (1997)
18. Nomura, S., Ishida, T., Yasuoka, M., Yamashita, N., Funakoshi, K.: Open Source Software Development with Your Mother Language: Intercultural Collaboration Experiment 2002. In: Proceedings of the 10th International Conference on Human - Computer Interaction (HCII2003). Crete, Greece (2003)
19. Ogura, K., Hayashi, Y., Nomura, S., Ishida, I.: User Adaptation in MT-mediated Communication. In: Su, K.-Y., Tsujii, J., Lee, J.-H., Kwong, O.Y. (eds.) IJCNLP 2004. LNCS (LNAI), vol. 3248, Springer, Heidelberg (2005)
20. Setlock, L.D., Fussell, S.R., Neuwirth, C.: Taking It Out of Context: Collaborating within and across Cultures in Face-to-Face Setting and via Instant Messaging. In: Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW '04). Chicago, Illinois (2004)
21. Theologitis, D.: ...and the Profession? The Impact of New Technology on the Translator. Terminologie and Traduction, vol. 98(1). Office for Official Publications of the European Communities, Luxembourg (1998)
22. Turpin, A., Scholer, F.: User Performance versus Precision Measures for Simple Search Tasks. In: Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, University of Washington, Seattle, Washington (2006)
23. Yang, J., Lange, E.D.: SYSTRAN on Alta Vista: A User Study on Real-Time Machine Translation on the Internet. In: Farwell, D., Gerber, L., Hovy, E. (eds.) AMTA 1998. LNCS (LNAI), vol. 1529, pp. 275–285. Springer, Heidelberg (1998)

Bayesian Belief Network Approach for Analysis of Intercultural Collaboration in Virtual Communities Using Social Capital Theory

Ben K. Daniel[1, 2], Gordon I. McCalla[1], and Richard A. Schwier[2]

[1] ARIES Research Group
Department of Computer Science
University of Saskatchewan
Saskatoon, Saskatchewan, Canada
[2] Virtual Learning Community Research Laboratory
Educational Communications and Technology
University of Saskatchewan
Saskatoon, Saskatchewan, Canada
{ben.daniel,gordon.mccalla,richard.schwier}@usask.ca

Abstract. We present a Bayesian belief network approach using social capital theory as a means for addressing issues that are critical to intercultural collaboration learning in virtual communities. Our work has two contributions; first, we present a computational approach that can be used for understanding social capital and intercultural factors critical to the design of virtual communities. Second, using evidence-based scenarios, we show how the Bayesian model can be tuned over time as knowledge about the system grows.

Keywords: virtual communities, social capital, Bayesian belief network, intercultural collaboration, awareness, trust, social protocols, knowledge sharing.

1 Introduction

There is a plethora of research suggesting that collaborative learning is one of the most fundamental aspects of virtual communities [41, 42]. This research builds on the grounds that much of the knowledge sharing activities, in all forms of virtual communities' take place through collaboration [18, 37, and 5]. However, collaborative learning that ultimately led to knowledge sharing under distributed circumstances is often difficult to achieve due to intercultural factors [4]. Intercultural factors especially, in the context of communication in virtual communities are concerned with the exchange, co-creation, of information, meaning and interpretation by individuals or groups. [14]. Typical intercultural factors critical to collaboration include; diversity in members' demographic cultures, organizational cultures and professional cultures and the like.

Drawing from research into social capital and virtual communities over the years [6, 5, 3], we present a Bayesian computational model of social capital as a framework for addressing issues critical to intercultural collaboration learning in virtual communities. Using evidence-based scenarios, we show how changes in different variables in a Bayesian model can affect an overall level of social capital and trust.

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 291–305, 2007.

2 Related Research

The term "culture" has multiple meanings in different contexts. In this study culture refers to the commonly shared system of general beliefs, values, and underlying assumptions held by a group of people. Culture is always believed to be a collective phenomenon. It is a collective programming of the mind, which distinguishes the members of one group or category of people from another. Culture can be learned not inherited [16]. Hall [15] described "culture as a total communication framework" comprised of words, actions, nonverbal behaviors, the handling of time, space, and materials, world view, beliefs, and attitudes passed over time from generation to generation. Culture also serves as a perceptual filter with which people determine what is and is not important to them [14]. How does culture affect knowledge sharing and learning in virtual environments?

In communities or groups, people learn as they navigate to solve problems together [44] or design representations of their understanding [13]. Quite often people who are willing to share their knowledge seek to attain a balance between donating and collecting knowledge. This implies that in order to effectively share knowledge there has to exist certain norms of reciprocity—people share their own knowledge because they expect others to contribute as well [21, 38].

But to fully understand what can be shared and the challenges associated with sharing knowledge, we explore the notion of knowledge, differentiating it from information and data. As Andriessen [25] pointed that information is basically a collection of facts and figures, while knowledge consists of insights and interpretations. Though the distinction between interpersonal and community knowledge can help us to understand what can easily be shared and what cannot, there is no agreed upon standard definition of knowledge. However, Polanyi [28] distinction between tacit and explicit knowledge has gained widespread acceptance.

Tacit and explicit knowledge are common to all kinds of virtual communities but the protocol for sharing each one of them differs from one community to another. The distinction between knowledge, information and data is context dependent [see Fig. 1]. For instance, when people exchange data, the data is processed into information. In turn, information can be situated in a particular context and turned into knowledge for a particular individual. Both information and knowledge then are grounded on data. The two can be differentiated if we consider interpretation and meaning. Information by definition is informative and, therefore, informs us about something. It can also be treated as data from which we can derive meaning. Knowledge is directly related to understanding and is gained through the interpretation of information. Knowledge enables us to interpret information i.e. derive meaning from data.

Knowledge can be derived from both information and data since the context of data needs to be known before it can be interpreted as information [5]. How specific knowledge is generated from data and information depends on how the data are stored, and how information is presented, organized, communicated and received by particular individuals within a particular community. In many cases, the process of knowledge sharing in virtual communities is mostly achieved through tacit to tacit communication, though clearly knowledge sharing can also be achieved through the tacit to explicit to tacit conversion loop and it is highly contextual [3].

Beyond contextual issues of knowledge sharing, there are also interpersonal issues relating to individuals backgrounds, interests, organizational affiliations and competence. These issues have greater impact on intercultural collaboration and can influence the process of knowledge sharing. How do we understand intercultural issues critical to collaboration in virtual communities? We propose a model of social capital to simulate the relationship of different variables and use the model's predictions to speculate on issues critical to intercultural collaboration.

2.1 Social Capital

What is social capital (SC)? Our search for understanding SC in general and in the context of technology-mediated communities in particular begins with exploration and selection of a number of definitions that appear in current literature. Table 2 presents a selected summary of some of the definitions.

From table 2, it seems there is no single definition of SC, but it is possible to categorize current definitions into two major categories. Among these there is a structural dimension [26, 21, 30, 34, 45] and content dimension in which SC resides [11, 38, 40]. The structural dimension refers to the fundamental elements of a social network such as types of ties and connections and the social organization of a community and the content dimension includes the types of norms, trusts, shared understanding and those variables that hold people together. We simplify the two dimensions, types and examples in Fig. 1. While definitions of SC can vary based on the discipline and the researcher, the main aspects of SC are trust and shared values, community involvement, volunteering, social networks and civic participation. A summary of key definitions of SC is presented in table 1.

Table 1. Definitions of social capital and key variables

Researcher (s)	***Definition***	***Key variables***
Putnam [2000]	The connections among individuals – social networks and the norms of reciprocity and trustworthiness that arise from them. In that sense SC is closely related to what some have called "civic virtue.	Connections, networks, norms/social protocols, reciprocity, trust
World Bank [1999]	The institutions, relationships, and norms that shape the quality and quantity of a society's social interactions.	Relationships, norms/social protocols, social interactions
Cohen & Prusak [2001]	The stock of active connections among people: the trust, mutual understanding, and shared values and behaviors that bind the members of human networks and communities and make cooperative action possible	Connections, trust, mutual understanding/shared understanding, shared value/goals, networks

Table 1. (*continued*)

Bourdieu [1996]	The aggregate of the actual or potential resources which are linked to possession of a durable network of more or less institutionalized relationships of mutual acquaintance and recognition	Relationships, resources, networks
OECD [2001]	The network, together with shared norms, values and understandings that facilitates cooperation within and among groups.	Network, shared norms, shared understanding, cooperation
Nahapiet & Ghosal [1998]	The sum of actual and potential and resources embedded within, available through and derived from a network of relationships possessed by an individual or social unit.	Resources, network, relationships
Loury [1977]	Natural occurring social relationships among persons which promote or assist the acquisition of skills and traits valued in the market place.	Values, social relationships, skills, traits
Woolcock [1998]	Information, trust and norms of reciprocity inhering in one's social networks.	Information, trust, norms/social protocols, social networks
Inglehart [1997]	The culture of trust and tolerance, in which extensive networks of voluntary associations emerge	Culture, trust, tolerance/shared understanding, networks, voluntary associations
Resnick [2004]	Productive resources that inhere in social relations	Resources, social relations
Rafaeli, Ravid & Soroka [2004]	A collection of features of the social network created as a result of virtual community activities that lead to development of common social norms and rules that assist cooperation for mutual benefit.	Social network, common social norms/social protocols, co-operation, mutual benefit

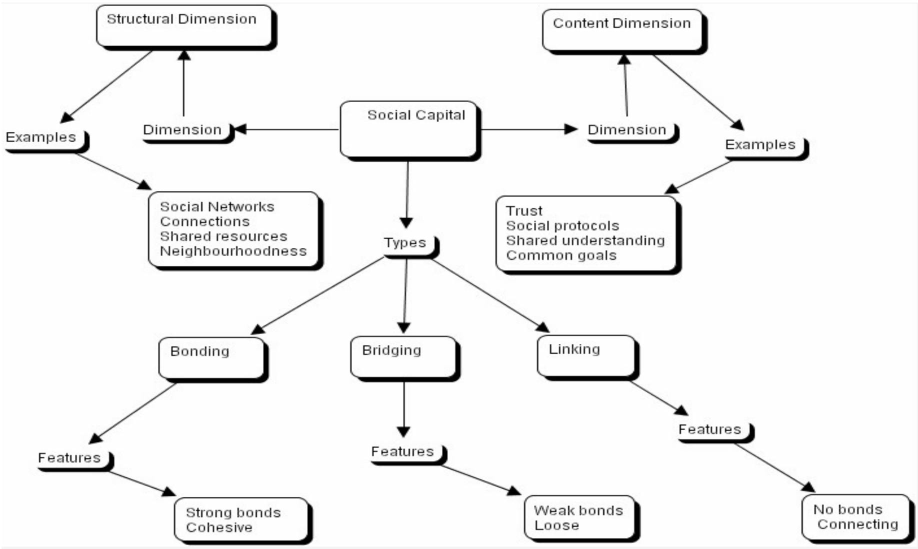


Fig 1. Dimensions and types of social capital

Further benefits of social appear in the literature. Putnam [40] suggested that SC allows people to resolve collective problems more easily; that is, people are normally better off if they collaborate on a task with each other. Mechanisms such as social sanctions are used for coping with breaches in social protocols (e.g., a case in which individuals benefit by shirking their responsibilities, hoping others will do their work for them). He also observed that SC greases the wheel that allows communities to advance smoothly. For instance, when people are trusting and trustworthy, and maintain continuous interaction, everyday business becomes easier and more enjoyable.He added that people who are well connected usually receive good news first. Further, SC can help preserve social norms in the community and reduce delinquent or selfish behaviour. For instance, people who are well connected in a community and have active trusting connections with others are likely to behave in the accepted social manner.

The community benefits of SC appear to extend to formal educational institutions. The World Bank [45] found that schools were more effective when parents and local communities were actively involved. Teachers were more committed and students had high tests scores. The mentoring, networking and mutual support associated with high levels of SC contributes to success in education [26].

Firms benefit from SC because it facilitates cooperation and coordination, which minimizes transaction costs, such as negotiation and enforcement, imperfect information and layers of unnecessary bureaucracy. Reciprocal, interdependent relationships reinforce compliance, which helps firms minimize financial risks. For instance, in a production system such as automobile company, where there is a lack of cooperative agreement, a parts manufacturer might be able to take advantage of others by strategically altering prices. In the corporate sector, SC can provide a competitive edge because efficiency gains in time and information allow more resources to be devoted to producing and marketing a better product at a higher volume.

SC can bridge cultural differences by building a common identity and shared understanding. SC requires continuous interaction and enables people to identify common interests and build trust. This raises their level of shared commitment, and encourages a sense of solidarity within a community. Furthermore, from the perspective of organizational management, Prusak and Cohen [11] pointed out that SC can promote better knowledge sharing due to established trust relationships, common frames of reference and shared goals.

Narayan and Pritchett [21] further suggested that communities with high SC have frequent interaction, which in turn cultivates norms of reciprocity through which learners become more willing to help one another, and which improve coordination and dissemination of information and knowledge sharing. SC has been used as a framework for understanding a wide range of social issues in temporal communities. It has been used for the investigation of issues such as trust, participation, and cooperation.

3 Modelling Social Capital in Virtual Communities

Daniel, Schwier and McCalla [3] refer to social capital in virtual learning communities as a common social resource that facilitates information exchange, knowledge sharing, and knowledge construction through continuous interaction, built on trust and maintained through shared understanding. A growing body of research shows that building social capital requires continuous and positive interaction [6, 11, 40, 43]. Positive interaction as a function of social capital provides value to its participants especially when it is built upon positive attitudes among individuals in a community [3]. Further, positive interaction and attitudes enable people to identify common goals, achieve shared understanding and social protocols, build trust, and commit themselves to each other's well being [43]. The value derived especially from positive interactions and attitudes can include sharing personal experiences with others, endorsing behavior, surfacing tacit knowledge, sharing information, recommending options, and providing companionship and hospitality [1].

In essence, the nature of members' attitudes toward each other and the kinds of relationships that form from their interactions can be used as starting point for analyzing social capital in virtual communities. Resnick [35] noted SC can be understood through the kinds of relationships among individuals, and that the lack of SC in a community reveals the absences of productive relationships. Further, productive relationships occur when participants have a common set of expectations, mediated by a set of shared social protocols [23]. In addition, when members of a virtual community establish a certain level of shared understanding, they can attain a certain level of social capital. The process of establishing shared understanding often draws upon a set of shared beliefs, shared goals and values, experiences and knowledge [4, 41].

Further, maintaining different forms of awareness in a virtual community can lubricate the value of interaction. For instance, in order to effectively collaborate and function as a community, people need to be aware of others, where they are located (demographic awareness), what they do (professional awareness), what others know (knowledge awareness) and what they are able to do (capability awareness) [3, 10]. Different forms of awareness in a virtual community can encourage positive interaction and productive collaboration [7]. Awareness can also foster trust. It has also been suggested that trust develops as people become sufficiently aware of others in the community (who is who, who knows who, who knows what, who is located where) and learn what to expect from each other [9, 1]. Trust is a critical ingredient and a lubricant to almost many forms of social interactions [6]. Trust enables people to work together, collaborate, and smoothly exchange information and share knowledge without time worsted on negotiation and conflict [11].

3.1 Bayesian Belief Network

A Bayesian network is a particular type of graphical model, frequently used in applications of artificial intelligence for building probabilistic expert systems. Bayesian networks can be used to model probabilistic relationships among variables. In some cases, their graphical structure can be loosely interpreted as the result of direct causal dependencies between variables. In domains with many causal relations, such as in medical diagnosis (symptoms cause diseases), human experts are usually able to express their domain knowledge in the graphical structure of the network. For example, in a model for medical diagnosis, the parameters of the network are the conditional probabilities of effects given the state of their direct causes.

Bayesian Belief Network (BBN) techniques are increasingly being used for understanding and simulating computational models [31]. BBN models enable reasoning when there is uncertainty [22]. They combine the advantages of an intuitive visual representation with a sound mathematical basis in Bayesian probability. Though BBN techniques are elegant ways for capturing uncertainties, considerable effort is required to create conditional probability values for the given variables in a network. We have propose a computational framework drawing from both qualitative and quantitative descriptions of conditional probability values [3], and we illustrated in previous studies [5] how the framework has been used to build models of trust and social capital that can be used as decision tools for the design of social software systems.

Constructing a BBN generally involves several steps. The first step in constructing BBN models is to define the problem or opportunity that needs to be modelled. This is followed by the identification of possible critical variables in a particular domain, along with their possible states [5]. From our synthesis, social capital is not only a function of single variables. The various instances of social capital in table 1 were identified within two kinds of virtual communities; distributed communities of practice and virtual learning communities. In many studies, trust is used as proxy for measuring SC in communities [45, 40]. In our work we have identified variables constituting social capital drawn mainly from the literature (see table1) and added others that we believe may be relevant in the context of virtual communities [3]. The variables identified are then assigned potential states (see table 2).

Table 2. Variables of social capital and their definitions in virtual communities [8]

Name	Definition	States
Interaction	Exchanging of information between two or more individuals via text, video, or any other digital media	Present/Absent
Attitudes	Individuals' general perception about each other and others' actions	Positive/Negative
Community Type	The type of environment, tools, goals, and tasks that define the group	Virtual learning community (VLC) or Distributed community of practice (DCoP)
Shared Understanding	A mutual agreement/consensus between two or more agents about the meaning of an object	High/Low
Awareness	Knowledge of people, tasks, or environment and or all of the above	Present/Absent
Demographic Awareness	Knowledge of an individual: country of origin, language and location	Present/Absent
Professional Awareness	Knowledge of an individual's background training, affiliation etc.	Present/Absent
Competence Awareness	Knowledge about an individual's capabilities, competencies, and skills	Present/Absent
Capability Awareness	Knowledge of an individual's competences and skills in regards to performing a particular task	Present/Absent
Social Protocols	The mutually agreed upon, acceptable and unacceptable patterns of behaviour in a community	Present/Absent
Trust	A particular level of certainty or confidence with which an agent uses to assess the action of another agent	High/Low

The second step involved mapping the variables into a network structure based on logical, and coherent qualitative reasoning [3]. The resulting network shows dependencies among variables (see Fig. 2).

The graph presented above relates only to two forms of virtual communities (VLCs and DCoP), the graph topology enables different forms of model updating to be conducted. Once a BBN graph is developed, the third stage is to obtain initial probability values to populate the network. In our case the initial probabilities were generated by qualitative descriptions of the strength of the relationship among variables in a network. This approach takes into account the number of states of a

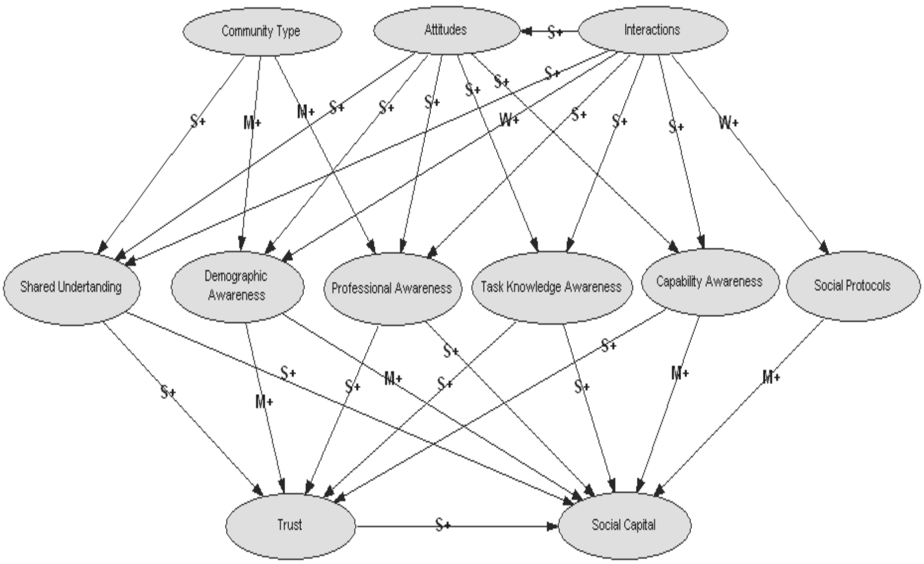


Fig 2. A BBN model of trust and social capital in virtual communities [3]

variable, the number of parents, the relative strength of a variable (e.g., strong -S, medium -M, weak -W) and the kind of relationship/influence of the variable (e.g., positive or negative influence - +/-) to produce initial prior and conditional probabilities. Once an initial model is elicited, particular scenarios are used to refine and document the network [3].

The conditional probability values were obtained by adding weights to the values of the variables depending on the number of parents and the strength of the relationship between particular parents and children. For example, say Attitudes and Interactions have positive and strong (S+) relationships with Knowledge Awareness; the evidence of positive interactions and positive attitudes will produce a conditional probability value for Knowledge Awareness to be 0.98 (threshold value for strong = 0.98). The weights were obtained by subtracting a base value (1 / number of states, 0.5 in this case) from the threshold value associated to the degree of influence and dividing the result by the number of parents (i.e. (0.98 - 0.5) / 2 = 0.48 / 2 = 0.24), this follows the fact that in the graph Knowledge awareness is a child of both interactions and attitudes. Table 3 shows the threshold values and weights used in this example. Since it is more likely that a certain degree of uncertainty can exist, value $\alpha = 0.02$ leaves some room for uncertainty when considering evidence coming from positive and strong relationships. These threshold values can be adjusted based on expert opinion.

Using this approach it is possible to generate conditional probability tables (CPTs) for each node (variable) regardless of the number of parents. Depending on how the initial knowledge is elicited and what decisions are made to process the knowledge into initial probabilities. For instance, assuming some subject matter experts are consulted to obtain initial probabilities, this knowledge is translated into the threshold

Table 3. Threshold values and weights with two parents [3]

Degree of influence	Thresholds	Weights
Strong	$1-\alpha = 1 - 0.02 = 0.98$	(0.98-0.5) / 2 = 0.48 / 2 = 0.24
Medium	0.8	(0.8-0.5) / 2 =0.3 / 2 = 0.15
Weak	0.6	(0.6-0.5) / 2 =0.1 / 2 = 0.05

weighted values as described in table 3 above depending on the degree of influence among the variables (i.e. evidence coming from one of the parent's states), decision which can also be obtained from the subject mater expert in a particular domain. However, when experts define degrees of influence for more than one of the parents' states, adding weights could result in ties, which could generate inconsistent CPT. In such cases, one could ask the expert which parent should be used, or has the most probable high degree of influence depending on the case under investigation.

3.2 Querying the Network

The fourth step in the modeling process is to observe changes in the network when new evidence affecting one or more of the nodes is added to the network. Querying a BBN refers to the process of updating conditional probability tables and making inferences based on new evidence. One way of analyzing and refining a BBN is to develop a detailed number of scenarios or cases grounded on data, intuition, and experts' opinions describing a set of phenomena in a virtual community. A scenario/case refers to a written synopsis of inferences drawn from observed phenomena, intuition or empirical data. For illustration, we describe a scenario in a distributed community of practice, whose members have a high level of shared understanding, professional awareness and positive interactions. Using evidence based scenarios; we query the model and observe changes in different kinds of variables that are likely to affect intercultural issues critical to collaborations in two communities.

3.3 Scenario One

Governance Knowledge Network (GKN) [http://www.icgd.usask.ca/gkn/] is a distributed community of practice (DCoP) that reflects diversity in membership in terms of background in training, organizational affiliation and geographical location. The Governance Knowledge Network (GKN) research was launched to address a perceived need to span geography and cross-organizational boundaries to enhance the scholarship on, and the practice of, governance and its role in advancing international development.

A large-scale distributed community of practice such as GKN usually draws their members from hundreds and sometimes thousands of people, who are geographically distributed and professionally diverse. Distance and geography can make knowing each other difficult. Even in the presence of robust technologies building awareness among individuals in such communities can be a great challenge. How do different forms of awareness affect the social capital of this community?

We use this scenario to query the network to answer this question. In the process, we assume there is a positive interaction in the group, the type of community is DCoP and since this group draws individuals from different domains, it is more likely that the cultural diversity in the group can have influence on the level of shared understanding, so we set shared understanding to "no exist" and equally assume no any other kind of awareness is known in the community. After updating the network, the posterior probability of social capital high = 0.207 and low = 0.792; likewise, the posterior probability of trust high = 0.02 high and low = 0.98. Fig.4 shows the graphical representation of the network after the query was performed.

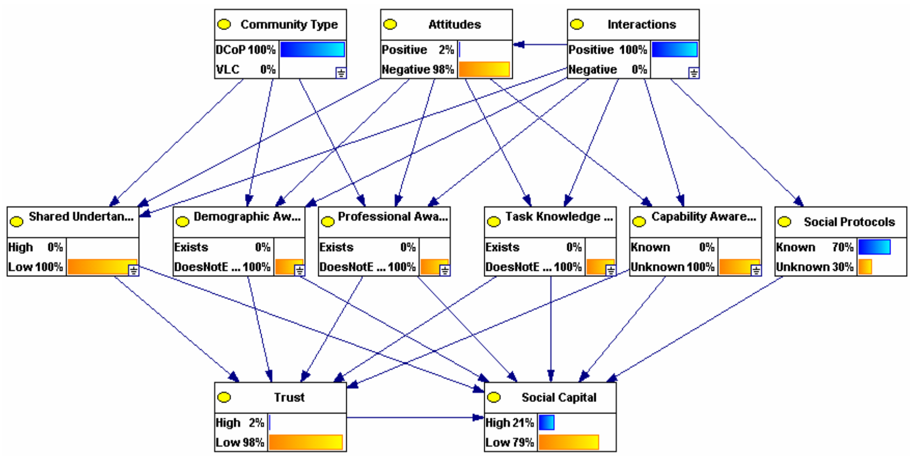


Fig. 3. Posterior probability values in scenario one

The low levels of posterior probabilities in trust and social capital can be attributed to the absences of different forms of awareness and shared understanding. At this point however, no evidence regarding which kind of awareness has what impact on the levels of social capital and trust. But one can generally infer that there is a possible correlation between trust, social capital and different forms of awareness. At this point more studies or scenarios are needed to perform further analysis.

The scenario presented above was developed with the assumptions that a distributed community of practice is typically a group of geographically dispersed professionals in different fields who share common practices and interests in a particular area of concern, and whose activities can be enriched and mediated by information and communication technologies. Such a group usually maintains high level of shared understanding and professional awareness, and so variation in the level of shared understanding and awareness of each other within such a group can affect the level of trust and social capital as demonstrated by changes in the probability distributions in the network.

3.4 Scenario Two

The second scenario involved analysis of a formal virtual learning community of graduate students learning fundamental concepts and philosophies of E-Learning. The members of this community were drawn from diverse cultural backgrounds and different professional training. In particular, participants were practising teachers teaching in different domains at secondary and primary schools levels. Some individuals in the community had extensive experiences with educational technologies, while others were novices but had some experience in classroom technology and pedagogy. Members in this community were not exposed to each other before and thus were not aware of each other's talents and experiences. However, individuals knew about each other i.e. to say they were aware of others demographic information and professional affiliations.

Since the community was a formal one, there was a formalized discourse structure and the social protocols for interactions were explained to participants in advance. The special protocols required different forms of interactions including posting messages, critiquing others, providing feedback to others postings, asking for clarifications etc. As the interactions progressed in this community, intense disagreements were observed in the community and there was a little shared understanding among the participants. Given the conditions in this scenario, we updated the network and observed low values in the posterior probabilities of social capital = high is 0.637 and low =0.362, trust high is 0.656 and low is 0.344 respectively.

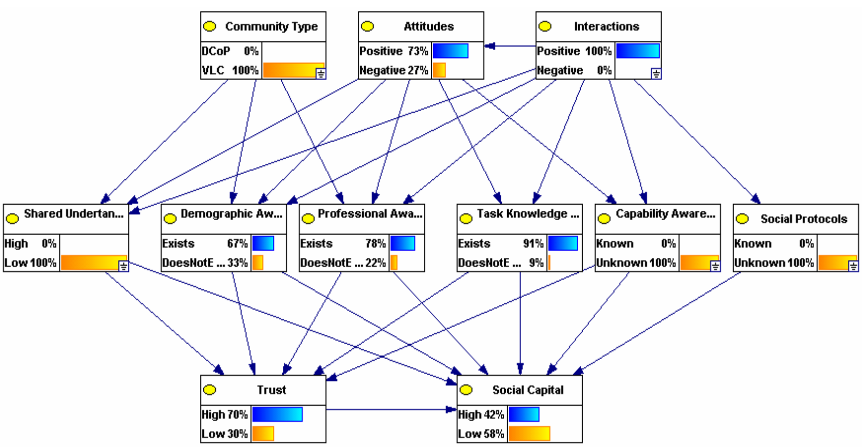


Fig. 4. Posterior probability values in scenario two

The lack of shared understanding, task knowledge awareness and capability awareness and presence of social protocols in scenario two has a moderate influence on the level of social capital and trust. Similarly it is hard to come up with final conclusions without any kind of sensitivity analysis in explaining which variables are the most influential in the model.

4 Conclusion and Model Implications

In virtual communities where face-to-face interactions is lacking, intercultural collaboration becomes a delicate process, where different individuals have a variety of means and standards to express themselves and interact with each other in diverse settings and features obvious in face-to-face interaction—such as age, gender, or physical appearance—are only inferred from the interactions in the community. Nonetheless, in both virtual communities and terrestrial communities, there are common features, for instances in both communities, people tend to establish meaningful relationships; they share a common purpose, shared understanding, trust each other, operate within certain social protocols [23][24]. Understanding these features helps us to develop robust tools and processes that effectively encourage collaboration and knowledge sharing. As Raybourn [14] suggested that designers of collaborative virtual communities have both the responsibility and the opportunity to consider the impact of underlying dynamics of culture and intercultural interactions such as identity, negotiation, conflict, power, equity, and trust on virtual spaces and collaborative communities.

We suggest that using social capital as a framework can enhance our understanding of a culture sensitive design for intercultural communication in virtual communities. Within the social capital concept, variables such as awareness and trust can be used as proxies to further address intercultural issues that are critical to collaboration. Predictions of the model suggested that trust in virtual communities can be influenced by variables such as awareness, social protocols and shared understanding. Further, trust is context can depend on the nature of the community and individuals. It is more likely that evidence that suggest otherwise can be easily fed into the model and the model update to observe the results. Follow up work will explain how to construct the network structure with concrete evidence and compare the simulation results and observation in the real world in order to further validate the BBN model.

References

1. Lee, A., Danis, C., Miller, T., Jung, Y.: Fostering Social Interaction in Online Spaces. In: Hirose, M. (ed.) Human-Computer Interaction (INTERACT'01) — Eighth IFIP TC.13 Conference on Human-Computer Interaction, pp. 59–66. IOS Press, Amsterdam (2001)
2. Mowshowitz, A.: Virtual Organizations: A Vision of Management in the Information Age. The Information Society, 267–294 (1995)
3. Daniel, B.K., Zapata-Rivera, J.D., McCalla, G.I.: A Bayesian computational model of social capital in virtual communities. In: Huysman, M., Wenger, E., Wulf, V. (eds.) Communities and technologies, pp. 287–305. Kluwer, London (2003)
4. Daniel, B.K., Sarkar, A., O'Brien, D.: A Participatory Design Approach for a Distributed Community of Practice on Governance and International Development. In: The proceedings of Ed-Media World Conference on Educational Multimedia. Hypermedia and Educational Telecommunications, Lugano, Switzerland (June 21st-26th, 2004)
5. Daniel, B.K., McCalla, G., Schwier, R.A.: Social Capital in Virtual Learning Communities and Distributed Communities of Practice. The Canadian Journal of Learning Technology 29(3), 113–139 (2003)

6. Daniel, B.K., McCalla, G.I., Schwier, R.A.: A Process Model for Building Social Capital in Virtual Learning Communities. In: Proceedings of the International Conference on Computers in Education (ICCE), Auckland New Zealand, pp. 574–577 (December 2-4, 2002)
7. Daniel, B.K., Schwier, R.A., Ross, H.M.: Intentional and incidental discourse variables in a virtual learning community. In: Proceedings of E-Learn 2005–World Conference on E-Learning in Corporate. Government. Healthcare and Higher Education. Vancouver, British Columbia, Canada (October 2005)
8. Daniel, B.K., Zapata-Rivera, J.D., McCalla, G.I.: A Bayesian Belief Network Approach for Modelling Complex Domains. In: Mittal, A., Kassim, A., Tan, T. (eds.) Bayesian Network Technologies: Applications and Graphical Models (in press)
9. Daniel, B.K., McCalla, G.I., Schwier, R.A.: Data mining and modeling social capital in virtual learning communities. In: Proceedings of the 12th International Conference on Artificial Intelligence in Education, Amsterdam, pp. 200–208 (July 18-22, 2005)
10. Gutwin, C., Greenberg, S.: Effects of Awareness Support on Groupware Usability. In: Proceedings of the ACM Conference on Human Factors in Computing Systems, pp. 511–518 (1998)
11. Cohen, D., Prusak, L.: In good company. How social capital makes organizations work. Harvard Business School Press, Massachusetts (2001)
12. Narayan, D., Pritchett, L.: Cents and Sociability: Household Income and Social Capital in Rural Tanzania. World Bank, Washington DC. USA (1997)
13. Suthers, D.: Representational support for collaborative inquiry. In: Proceedings of the 32nd Hawaii International Conference on the System Sciences, IEEE Computer Society Press, Los Alamitos. CA (1999)
14. Raybourn, E.M.: Designing an Emergent Culture of Negotiation in Collaborative Virtual Communities: The Case of the DomeCityMOO. In: SIGGROUP Bulletin. 129 (April 21, 2000)
15. Hall, E.T.: Beyond Culture. Anchor Press, Garden City. N.Y. (1976)
16. Hofstede, G.: Cultures and organizations: Software of the mind. McGraw-Hill, London (1991)
17. Loury, G.: A dynamic theory of racial income differences. In: Wallace, P.A., Le Mund, A. (eds.) Women, minorities, and employment discrimination, pp. 153–186. Lexington Books, Lexington. MA (1977)
18. McCalla, G.: The fragmentation of culture, learning, teaching and technology: Implications for artificial intelligence in education research. International Journal of Artificial Intelligence 11(2), 177–196 (2000)
19. Ogata, H., Yano, Y.: Combining knowledge awareness and information filtering in an open-ended collaborative learning environment. International Journal of Artificial Intelligence in Education 11, 33–46 (2000)
20. Nonaka, I., Takeuchi, H.: The Knowledge-Creating Company: How. In: Japanese Companies Create the Dynamics of Innovation, Oxford University Press, Oxford (1995)
21. Nahapiet, J., Ghoshal, S.: Social Capital, Intellectual Capital and the Organizational Advantage. Academy of Management Review 23(2), 242–266 (1998)
22. Pearl, J.: Probabilistic reasoning in intelligent systems: Networks of plausible inference. Morgan Kaufmann, San Mateo, CA (1988)
23. Preece, J.: Supporting community and building social capital. Special edition of Communications of the ACM 45, 37–39 (2002)
24. Andriessen, J.H.E.: Knowledge Communities in fives. IS - 2006-01 Delft Innovation System Papers (2006)

25. Coleman, J.S.: Social capital in the creation of human capital. American Journal of Sociology 94, 95–120 (1988)
26. Sveiby, K.E.: Measuring intangibles and intellectual capital–an emerging first standard. 1998. Retrieved January 15th from (2006), http://www.sveiby.com/articles/EmergingStandard.html
27. Polanyi, M.: The tacit dimension. Routledge and Kegan Paul, London (1967)
28. Turoff, M.: Virtuality. Communications of the ACM 40(9), 38–43 (1997)
29. Woolcock, M.: Social capital and economic development: Toward a theoretical synthesis and policy framework. Theory and Society 27(2), 151–208 (1998)
30. Druzdzel, M.J., Gaag, L.C.: Building probabilistic networks: Where do the numbers come from?Guest editor's introduction. Data Engineering 12(4), 481–486 (2000)
31. Sharratt, M., Usoro, A.: Understanding Knowledge-Sharing in Online Communities of Practice. Electronic Journal on Knowledge Management 1(2), 187–196 (2003)
32. OECD: The Well-being of Nations: The Role of Human and Social Capital, Paris. Policy-makers should not have to wait for a couple of decades of detailed research before asking whether attentiveness to social capital might be worth their while. OECD Observer. No 226/227. (Summer 2001)
33. Bourdieu, P.: On the family as a realized category. Theory, Culture and Society 13(1), 19–26 (1996)
34. Resnick, P.: Beyond bowling together: Sociotechnical capital. In: Carroll, J.M. (ed.) HCI in the new millennium, vol. 47, p. 272. Addison-Wesley, New York (2002)
35. Fahy, P.J., Crawford, G., Ally, M.: Patterns of interaction in a computer conference transcript. International Review of Research in Open and Distance Learning (2)1. Retrieved August 28. 2006 (2001), from http://www.irrodl.org/content/v2.1/fahy.html
36. Schwier, R.A.: Catalysts, emphases and elements of virtual learning communities. Implication for research. The Quarterly Review of Distance Education 2(1), 5–18 (2001)
37. Putnam, R.D.: Making Democracy Work: civic traditions in modern Italy. Princeton University Press, Princeton NJ (1993)
38. Inglehart, R.: Modernization and Postmodernization: Cultural, Economic, and Political Change in 43 Societies. Princeton University Press, Princeton (1997)
39. Putnam, R.: Bowling Alone: The Collapse and Revival of American Community. Simon Schuster, New York (2000)
40. Schwier, R.A., Daniel, B.K.: Did we become a community? Multiple methods for identifying community and its constituent elements in formal online learning environments. In: Lambropoulos, N., Zaphiris, P. (eds.) User- Evaluation and Online Communities. IDEA Group, Hershey (2006)
41. Schwier, R.A., Daniel, B.K.: Implications of Virtual Learning Communities for Designing Online Communities of Practice in Higher Education. In: Kimbel, C., Hildreth, P. (eds.) Communities of Practice: Creating Learning Environments for Educators. Information Age Publishing, Greenwich. CT (2007)
42. Rafaeli, S., Ravid, G., Soroka, V.: De-lurking in virtual communities: A social communication network approach to measuring the effects of social capital. In: Paper presented at the HICSS 37. Hawaii, Big Island (2004)
43. Koschmann, T., Kelson, A.C., Feltovich, P.J., Barrows, H.S.: Computer-Supported Problem-Based Learning: A principled approach to the use of computers in collaborative learning. In: Koschmann, T. (ed.) CSCL: Theory and practice of an emerging paradigm, pp. 83–124. Erlbaum, Mahwah. NJ (1996)
44. World Bank: Social capital research group. Retrieved from (May 29, 2003) http://ww.worldbank.org/poverty/scapital/

Social Conventions and Issues of Space for Distributed Collaboration

Hillevi Sundholm

Dept. of Computer and Systems Sciences
Stockholm University/The Royal Institute of Technology
Forum 100, S-164 40 Kista, Sweden
hillevi@dsv.su.se

Abstract. We followed the work of an international research network that holds regular meetings in technology-enhanced environments. The team is geographically distributed and to support its collaborative work it uses a set of technical artifacts, including audio- and videoconferencing systems and a media space. We have been studying some of the techniques and social conventions the team develops for its collaboration, and different aspects of what it mean to be located in a shared but distributed workspace. Our approach has been to analyze the initiatives and responses made by the team members. Over time the group created conventions; e.g. the chair introduces team members participating only by audio and members turn off their microphones when not talking. The latter convention led to the side effect of faster decision making. We also identified two characteristics, *implicit excluding* and *explicit including*, in a situation where the majority of the team members were co-located.

Keywords: Social conventions, Distributed Collaboration, Co-location, Shared Workspace, Group-to-group collaboration.

1 Introduction

Although current technologies offer many different possibilities for communicating, interacting and sharing information simultaneously at a distance, people still prefer to work at the same place using a common collaborative space [32]. Computer-mediated communication may be efficient for disseminating information but for decision-making it is rarely if ever more effective than face-to-face meetings [4]. Moreover, the frequency and quality of communication declines when the distance increases between participants' offices [22]. This finding has recently been supported in an experimental study [7]; the authors concluded that researchers in the field of Computer-Supportive Cooperative Work (CSCW) need to pay more attention to the design of technology to overcome social and geographical distance. Group-to-group collaboration, which is becoming increasingly common, also deserves more study as it places different demands on the participating individuals and on the design of the systems, compared to individual distance collaboration [28].

In this paper we present a study with a group of researchers who regularly engage in geographically distributed meetings. The team uses a set of technical artifacts to

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 306–320, 2007.

support its collaborative work, including large wall displays and a media space. An important characteristic of the setting, one that makes the situation more complex, is that both video- and audio conferences have been used as communication channels. In this area of research many studies focus on systems and users, and on the specific design of shared tools for distributed collaboration [14, 18, 25, 31, 41]. However, few long-term empirical studies have aimed at furthering our understanding of teamwork in these settings. In a previous paper [37] we looked at how the team members created mutual understanding about the current situation, how available artifacts mediated the collaboration, and the role of visual representations. We investigated those issues by looking at the turn-taking between both the team members and when using the shared media space, in addition to which communication channel they use (i.e. video and/or audio). We saw that in this multi-channel setting the participants often have to clarify who is attending, and those who are only present on audio risk being left out of conversation. Another outcome was that the communication space is limited; when many want to participate in the communicative activity, it becomes harder to make a successful initiative.

In this paper we investigate some of the techniques and social conventions the team members developed in order to collaborate and to handle situations that arose. So far social conventions have received too little attention in CSCW [27]. We also investigate what it means to have a "distributed" location, that is, to operate in a shared but distributed workspace. Initiative-Response Analysis [26] helped us study the interaction between the team members.

2 Related Work

In all collaborative work, an important issue is how to maintain, at least to some degree, both a shared view and a shared understanding so that the team members can perform the work and reach common goals. In long-term collaboration the team members must establish and maintain a shared awareness of their actions, plans, goals and activities [29]. Face-to-face interaction provides people with many contextual cues such as facial expressions, body postures and gestures that guide them as they interpret others' communication and interact with them [16]. Some or all of these cues disappear in distributed meetings, depending on which media are used to enable communication. Video is shown to be especially important in distributed collaboration if the participants do not have the same native language: the picture supports them in showing their understanding through facial expressions and gestures [40]. However, it is often more important to share the view of the work than see each other [12]: by demonstrating their activity to the other team members they can efficiently establish a shared understanding within the group [15].

2.1 Social Conventions

All actions are situated and all communication is embedded in the present and in the past. This means that we cannot separate an action from the context where it takes place; we also make use of earlier experiences to handle a current situation. We have to learn which verbal and nonverbal behaviors are appropriate in which situation [11], and we must understand what the current context means; this shared knowledge is essential if we are to be socially competent.

At least two techniques help people to coordinate their interaction: making explicit agreements and using conventions [10]. The conventions are normative and guide people towards the correct and acceptable behavior, as well as predicting group behavior [27]. The explicit conventions (as well as prescribed procedures) mediate the articulation of cooperative activities [34].

The community members share the conventions, which help them to coordinate action and avoid problems [24], because they make the social system more stabile, efficient, and coherent [5]. An example of a convention in our culture is that we greet one another with our right hands when we first meet. Conventions are arbitrary in the sense that they usually result from historical coincidences [10, 24]; for example, we greet with the right hand and not with the left [10].

Whereas Clark [10] makes the distinction between explicit agreements and conventions, Becker and Mark [5] differentiate between explicit and implicit social conventions. The explicit conventions are agreed upon, while the implicit ones are embedded in the social practice. When newcomers join a group or community they must become aware of the implicit social conventions [5].

Similar to our notion of social conventions Schön [33] talks about norms as a contract of shared rules between actors that govern their behavior. Those norms consist of both formal and informal understandings of what to expect from each other. These reciprocal expectations are often important in practice, e.g. when designing systems: the user and designer must share their expectations about the kinds of communications the system needs to handle in various kinds of situations [36].

Becker and Mark [5] compared three different virtual environments, focusing on the role social conventions play in communication and how the on-line systems supported those conventions. Their findings indicate that people use the same social conventions in those kinds of settings as in face-to-face interactions; however, they supported and/or expressed those conventions differently, depending on the media. For example, when team members had to leave a conversation, they signed off by saying goodbye, and included an explanation if they had to leave early. Social conventions are important for creating common understanding of behaviors, and therefore, they are important for maintaining the consistency of the space as a social system [5].

In a longitudinal study lasting about 4 years, Mark [27] followed a group of workers in a German ministry who used an electronic system to share and coordinate their work between different units and cities. The distributed groups failed to develop any normative conventions, partly because it takes time and a certain amount of communication to recognize each others actions and behaviors. As the co-workers were geographically distributed they could not create conventions implicitly; instead the local subgroups created their own conventions, which never became common to the larger group. They could have avoided this situation if they had committed to creating explicit conventions, or if the system had given them more feedback. Feedback functions in two ways: it can provide information about appropriate behavior, and/or it can reinforce certain behaviors [26].

Gay and Hembrooke [13] revealed the limitations on communications in a study of a tool that enabled students to view and discuss art online. The major problems they encountered were knowing when and where to make contributions because normal social conventions and cues were restricted. The lack of instant feedback may disturb the flow of communication, and the lack of visual cues, like facial or physical expressions, makes it difficult to interpret inactivity or silence.

2.2 Distributed Collaboration: Being Co-located, Being Distributed

Many researchers have studied the role of audio only in a distributed setting (e.g. [17]) and the role of video for remote collaboration (e.g. [19]). However, few have studied multi-channel or mixed-media settings, a complex situation where team members often are unequally distributed [9]. In an experimental study Bos et al. [8, 9] looked at teams that were partly co-located and partly distributed; a total of 130 subjects participated, divided into thirteen groups. Each group engaged in five sessions lasting 15-20 minutes each, in which half the group was co-located and half was distributed. Their task was to play a game that required them to buy and sell "skills", and each player had a special skill to offer. To succeed well in the game the player had to sell to and buy from both co-located and distributed participants. Their results show that the co-located players failed to pay enough attention to the distributed ones, a phenomenon called 'collocation blindness' [8]; in turn the distributed members created an 'in-group' with one another [9]. However, the distributed players were still able to do very well on the game if they had skills that were scarce in the co-located group [8].

3 Method

3.1 General Description of Corpus

Between April and December of 2004 we followed nine meetings of an international research network that consists of ten laboratories spread out across Europe and North America. The teams are not working on a common project but they share interests and use these occasions to share information and expertise, through discussions and presentations.

Every month they have a geographically distributed meeting. They use a multiplex videoconferencing system to transmit video and audio; an audio conference system is available in case the video link fails. They also have access to a shared media space and a wiki[1] website to share information; both are accessible on the Internet. They use the wiki website mainly to store internal information about the team members and the labs, along with meeting dates and agendas; this information is primarily used between the meetings. General information is also available to individuals who visit the page. The media space, on the other hand, is used as an information resource, a place where they can upload and download documents such as their presentations and working documents. The media space is used during the meetings and functions as a shared virtual workspace where everyone who is logged in can work simultaneously. One restriction is that only one person at a time can manipulate a document. To

[1] Wiki is a type of server software that allows invited users to create, add and remove web page content while using any browser (http://wiki.org/wiki.cgi?WhatIsWiki).

handle meta-communication and provide silent support during the meetings they use an instant messenger (IM), to which they log on before the meetings begin.

The laboratories use two screens to display the information, and usually at least one of the screens is large: one screen shows the team members who are present on video and the other displays the shared media space. Figure 1 shows the meeting situation in one of the labs; the video connections are shown on the left screen and the shared media space on the right one.

Fig. 1. Meeting situation

The number of participating teams may differ from one meeting to the next. During the meetings we followed, 4 to 7 teams were present at a given videoconference, and sometimes one or more teams were present on audio. Anywhere from 11 to 19 individual team members have participated in various meetings. The meeting time is a compromise between the labs since they are located in different time zones. Depending on which labs are connected, the local time may differ by as much as 11 hours.

The meetings have both a formal and an informal character. They are formal in the sense of having a clear meeting time, a chair, an agenda, and a procedure for getting connected. But they are also informal: the team members, specially the lab managers, know each other very well and the meetings function as a way to keep in touch. The meetings are divided into two parts. During the first part, for which 45 minutes is reserved, all the labs are to be connected and technical issues are discussed. The second part, which lasts about an hour, is the research seminar, where network activities and research are presented and discussed.

3.2 Data Collection

All meetings were recorded at Laboratory of Design for Cognition, EDF R&D, one of the network members. We used two to four fixed cameras in order to cover different angles in the space: one camera for the shared media space, one for the screen that shows the videoconference picture, and finally one or two for the local space (as illustrated in Figure 1). We also used a 360° angle camera and ceiling cameras to position the local participants, and at two of the meetings one of the participants used a wearable camera [23]; however, we do not include data from those cameras in this analysis. The data we analyze here consists of about 18 hours of video recordings. We also handed out or e-mailed questionnaires before and after the meetings. The questionnaires helped us understand the personal objectives of the people attending and asked for suggestions on how to improve the meetings.

3.3 Data Analysis

Our work is based on Initiative-Response Analysis [26], which analyses dialogue or multiparty communication (a neutral term is 'communicative activity'), in terms of initiatives and responses. The unit of analysis is the turn, and it is a useful model for understanding the global aspects of communicative activities where the turns are relatively short. We have focused especially on how people take the initiative to introduce a new episode [20]. The initiative is an attempt to request, claim or dominate and it refers forwards; the response refers backwards, and can be more or less immediate. In contrast to many other theorists, such as [35], Linell and Gustavsson [26] do not talk about 'follow-up moves'. Instead they mean that every utterance can be classified as either an initiative or a response. They developed a 6-level system for evaluating how strong or weak the initiatives and responses are, but for our purposes we have restricted the analysis to the following four: strong initiative (introducing a new topic and explicitly requesting a response), weak initiative (introducing new content by claiming something that possibly requires a response), extended response (response that adds new content to the preceding turn, or implicitly asks for a response) and minimal response (response without any initiative).

Interaction Analysis [21] has also influenced our analysis, particularly the following four foci: 'Beginning and Endings', 'Turn-taking', 'Trouble and Repair', and 'The Spatial Organization of Activity'.

We transcribed five of the meetings using regularly indicated time stamps. During the transcription and analysis, we noted the most interesting episodes. The notes covered a wide range, from what was monitored and manipulated on the displays to social interaction between the team members. We re-transcribed the parts we thought to be most interesting, adding detailed information, including the exact time stamps for beginnings and endings. We divided the excerpts following the work of [30], but modified the format slightly. Our 'Transcript of Interaction' does not indicate the times of pauses in the talk, and our 'Characteristic of Action' clarifies the action more abstractly: Is the turn is an initiative (I) or a response (R), and which line (L) does it refers to? In addition we categorized each turn in line with the work of Baker et al. [2, 3]. The categories we use in this report are Social relation, Interaction management, and Task management. By social relation we mean interaction concerned with managing the social relations and verbalizing the situation (the latter is not included in [2, 3]). Interaction management concerns the interaction itself: coordination, establishing contact, understanding, topic shifting etc. Finally, task management is about planning the task and making progress on it.

In the excerpts presented in the results all names were changed and personal information replaced by ellipses '[…]'.

4 Results and Analysis

For this analysis we have chosen to illustrate two different situations. In the first one, which represents the most common meeting situation, the labs are geographically distributed from each other; in these situations the difference is whether the teams are present on both video and audio, or only on audio. In the other situation, the "April meeting", most of the team members were present at a conference venue in Vienna (i.e. outside the laboratories), and only two teams[2] with a total of three participants were present on video from their labs.[3]

4.1 Being Present: Introduced, Forgotten and Re-introduced

Each team member can only completely view his or her own local situation, that is, who is attending (locally), what they have access to (e.g. the media space and the IM) and the quality of sound and picture. Normally the organizing lab, which also helps the other labs to connect, is the one that can have the best overview of the situation. For this reason the convention is that the chair explicitly shares his knowledge about the presence of the participants with the other team members. For this reason the convention is that the chair explicitly shares his knowledge about the presence of the participants with the other team members. For those participating by audio, it is especially important to be recognized as present and also to know who are present on video. In Excerpt 1 we see how the chair introduces a team member (present on audio) who joined the meeting late.

Excerpt 1. June meeting: Introduction to the other team members

Time 0:19:45	**Person, team, mode**	**Transcript of Interaction**	**Characteristics of Action**
1	Olivier, #1, video	"I, I think hmm… Andy? Did you join us, Andy?"	Strong I: Interaction management
2	Andy, #9, audio	"Yes, yes I am here."	Minimal R to L1: Interaction management
3	Olivier,#1, video	"Yes, you are here, okay… because the others were not aware that, that you had come in, so… I was the only one to know so I share the news."	Extended R to L2: Social relation
4	Andy, #9, audio	"Well, thank you."	Minimal R to L3: Social relation

End of Excerpt: 0:19:59

By using a strong initiative (line 1) we see that Olivier, the chair, introduced Andy to the others at the meeting. But the convention of introducing team members early on in the meeting does not guarantee that each participant will be remembered throughout. Excerpt 2 illustrates how a team member who is participating only on audio has been forgotten.

---

[2] A third lab was connected to handle the technical support.

[3] This situation is what Bos et al. [8] call 'partially-distributed groups'.

Excerpt 2. June meeting: Forgotten in the audio

Time 0:46:42	Person, team, mode	Transcript of Interaction	Characteristics of Action
1	Olivier, #1, video	"By the way, Thomas, are you still there?"	Strong I: Interaction management
2	Thomas, #5, audio	"I am still here, yeah."	Minimal R to L2: Interaction management
3	Olivier, #1, video	"Okay, because I realized we have forgotten you for a while."	Weak I: Social relation

End of Excerpt: 0:46:49

When Olivier realizes that Thomas had been forgotten, he assures the others of his presence by using a strong initiative (line 1). Our interpretation of these two situations is that someone participating only on audio may find it very difficult to make an initiative unless they have a direct request from someone who is already part of the communicative activity[4] (compare this to the analysis in section 4.3). Here, the convention of having the chair introduce audio-only participants and continue to check on their presence helps to overcome the problem to some extent.

4.2 Making Decisions Across Spaces

A meeting situation that includes up to seven teams at the same time[5] can be demanding for all the participants. It requires that they be focused and strict in the way they interact and that they all minimize disturbances during the meetings. Over time they have developed several conventions for handling the meeting situation. One convention is using an instant messenger to handle meta- communication, for example to say hello when connecting late or to inform others about connection problems. This allows them to communicate without interrupting the ongoing discussion. They also have created an explicit convention of turning off the microphone when not speaking, in order to minimize the risk of transmitting unwanted sounds to the others or causing an echo. Excerpt 3 illustrates that turning off the microphone can change the way people contribute to the conversation and make decisions. The team members are discussing whether they should exchange their IM system for another one and if so, to what other one.

Excerpt 3. November meeting: Decision-making

Time 0:54:00	Person, site, mode	Transcript of Interaction	Characteristic of Action
1	Olivier, #1, video	"All right, what, what I suggest is that some, somebody takes the lead on this, on this issue and takes the decision technically and, and with whoever they want, but that someone takes a decision. I think you on your side, Peter, you are probably the ones who have the most experience of this instant messaging system, if I... You seem to have tried a lot of things, so maybe somebody could just decide, and we'll, we'll just take your solution."	Strong I: Task management

[4] Another relevant aspect that we have argued for earlier is that the communication space is limited [37].

[5] More is possible, but this we have not seen during these nine meetings.

Excerpt 3. (*continued*)

2	Peter, #2, video	"Okay, we, we just talked about this here locally with the microphone off, ha-ha, and what we can do is, we will, Mattias will go to the wiki website, the […] website, and put up a little section on, on IM clients and with links to recommended systems for both Macintosh and, and Windows users. Do we have any people who are using Unix, Linux or anything?"	(Less than 1 s. between L1 and L2) Extended R to L1: Task management
3	John, #1, video	"Yes."	Minimal R to L2: Task management
4	Peter, #2, video	"Okay, we'll, we'll include a link to, to people, for solutions for those people as well."	Extended R to L3: Task management

End of sequence: 0:55:00

Olivier's strong initiative in line 1 gets a quick response from team #2 in line 2 since Peter is able to discuss it in parallel with his local team members. Turning off the microphone enabled them to accept the request quickly, and meanwhile manage their local interaction (who will do what). Line 2 ends with a request from Peter so that his team can fulfill the undertaking. The whole sequence takes one minute.

At the April meeting, in Vienna, the group made two formal decisions that needed approval from all the others. Excerpt 4 shows how they made the first decision.

Excerpt 4. April meeting: Decision-making

Time 0:52:17	**Person, team, mode**	**Transcript of Interaction**	**Characteristic of Action**
1	Olivier, Vienna	"Let's, let's go on. So we'll have to give an answer to Anna more formal than... In the way I understand it you more or less agree on this, or? More or less?"	Strong I: Task management
2	Several people, Vienna	[*Inaudible*]	Simultaneous talk (2 s.)
3	Aaron, Vienna	"[*Inaudible*] let's raise your hands."	Weak I: Task management
4	Andy, Vienna	"That's a great format!"	Extended R to L3: Social relation
5	Olivier, Vienna	"Yes." [*Everybody raising hands*]	Minimal R to L1: Task management
6	Peter, Vienna	"Yeah, we're okay."	Minimal R to L1: Task management
7	Several people, Vienna	[*Inaudible*]	Simultaneous talk (5 s.)
8	Olivier, Vienna	"What, what about you who are remote? […]"	Strong I, related to L1: Task management
9	Michael, #6, video	"Yeah."	Minimal R to L1 through L8: Task management
10	Olivier, Vienna	"Yeah?"	Minimal I: Task management
11	Wolfgang, #4, video	"Yes."	Minimal R to L1/L8: Task management
12	Unrecognizabl e voice, Vienna	"Raise your hands!"	Weak I: Social relation

Excerpt 4. (*continued*)

13	Nils, #6, video	"Yes, yes."	Minimal R to L1/L8: Task management
14	Olivier, Vienna	"Yes, raise your hands, that's correct."	Extended R to L1/L8: Social relation

End of Excerpt: 0:52:48

In this excerpt we see how the decision was first made among those co-located in Vienna, then by the distributed team members (from line 8 on). Since the majority of the team members were located in the same physical space they were naturally more able to react when the chair made the strong initiative of suggesting they come to a decision, compared to the distributed ones. The bonds between the team members in Vienna were stronger than between them and the two distributed teams. The distance between the team members was probably also influenced by the fact that the co-located team members started to talk more intensively with each other at two moments (lines 2 and 7), and this further excluded the two distributed teams. In line 8 Olivier used a strong initiative, turning to the distributed team members to include them in the decision-making process.

Their second decision proceeded more smoothly, but the co-located team members were still in focus first, followed by a direct question to the distributed ones.

4.3 Physical Space Matters: What Does It Mean to Be Distributed?

In this section we will look more closely into what it means to be participating in a geographically distributed meeting in relation to the mode. In our data we have three conditions: video and audio, audio only, and "co-located" (as in the case in the April meeting). Excerpt 4 raises the question of what it means to attend a meeting from a distance when most team members are in the same physical space. Compared to their normal meeting setting—a group-to-group meeting in which all the teams are "distributed" from one another—the setting in Vienna led to stronger bonds between the co-located participants by putting the few distributed team members in the situation of being "on a screen far away". Figures 2a and 2b illustrate the setting in Vienna.

Fig. 2a. Team members sitting in the same location

Fig. 2b. View of the other teams and of the shared media space

In the Vienna meeting we identified two phenomena that arose from the participants' situation at the meeting. It is important to note, and we will return to this, that these phenomena are not exclusive to the situation in Vienna. The first phenomenon is *implicit excluding*: without anyone intending to exclude the distributed participants from the co-located team members it can happen anyway (cf. [8] who talk about 'collocation blindness'). Sometimes it happens because the co-located participants share something that is hard or impossible to transfer over the video link; for example, the co-located participants in Vienna talked about the dinner they would attend that evening. Another example, as we saw in Excerpt 4, is that co-located participants started to have more intensive local discussions (lines 2 and 7); though they were short, 2 and 5 seconds respectively, they excluded the distributed team members. In that sense we might say that the situation is fairer when all the teams are distributed: the risk of being implicitly excluded is reduced, or equal, for everybody. On the other hand, we have seen in our data and illustrated in Excerpt 2, there is also the risk of being excluded in this more "equal" situation. Being invisible is certainly a disadvantage in this complex meeting situation, and the implication is that all teams need to be visible to the others.

One way to overcome some of these limitations in the system is the chair's active effort to include the others. This leads us to the second characteristic: *explicit including*. By explicit including we mean that it is easier for someone to make a successful turn (i.e. a response or a weak initiative) if someone else has already made a strong initiative to include them in the discussion (as we saw in Excerpts 1 and 2). Again this is especially evident when most of the group is gathered together and only a few team members are "distributed". In the data from the April meeting we saw only a few initiatives by the distant team members that succeeded without a direct request – thus, an explicit including. Also, team members who tried to contribute to the conversation rarely succeeded after the first try, a phenomenon we do not have space to illustrate here. One team member, who was connected in the videoconference, also stressed this problem a couple of times during the meeting. This shows how the distributed team members experienced what [28] describe as the "space between".

5 Discussion

Participants' ways of communicating and interacting in geographically distributed meetings are constrained by the available media as well as the composition of the media. We have seen in this study that the mix of video and audio conferencing added considerable complexity to the meeting situation (see also [37]). In this paper we have described some techniques and conventions that the team members developed to handle the situation and overcome some difficulties. To better understand the aspects of being distributed versus co-located we have paid special attention in this analysis to a meeting—the one in Vienna—where most team members were co-located. Identifying the *implicit excluding* and *explicit including* helps us to understand the team members' situation during the meetings.

At the Vienna meeting we saw that it was difficult for the few persons participating on video to enter into the discussion. This led us to what we call implicit excluding, that is being excluded without anyone intending that to happen. The local chats in Excerpt 4, of 2 and 5 seconds respectively, also show how the distributed team members were excluded. Similarly [1] showed that turning off the video and/or audio may make distributed participants feeling excluded. To understand why this happens we must remember that the participants in the face-to-face situation have the advantage of being part of the 'principal setting' [10] and that those participating from a distance are part of what can be seen as a 'derivative' of that setting (cf. [6], p. 43), which therefore requires particular management techniques and practices [10]. Correspondingly, in their habitual setting with group-to-group collaboration, we have seen how those participating by audio only have the disadvantage of not being seen; in this setting the audio only mode could be seen as the derivative of the principal setting of the video and audio mode. This can explain why the persons on audio find it hard to join into the conversation, and have the need of being introduced and re-introduced if they are not to be totally forgotten; that is to say, they need to be explicitly included. This is also the reason why the group has developed the convention of having the chair introduce all participating teams, no matter the mode, at the very beginning of the meeting. Clearly, an ideal system would communicate the presence of all the team members.

Another social convention they have developed over time is that those who are not speaking turn off their microphones. This convention was developed explicitly to minimize disturbances during the meeting, but it led to a new behavior: now people can talk to their local colleagues at the same time that they are part of the network activities. In Excerpt 3 we identified one outcome of this new behavior: a parallel local discussion led to a quicker decision for the team as a whole. This shows that it is possible to make faster decisions when using computer-mediated communication, an opposite result to what [4] have found. It also illustrates how the meeting space supports the group decision process as well as the division of labor in a way that would not be possible in a co-located meeting. The set-up makes it possible for the team members to communicate and interact with each other in several ways in parallel: with the team members in the common discussion, with the co-located team members, and with individual distributed team members through the IM. In this way different networks may co-exist [28].

The next step we plan to take is to look at how this kind of setting, with large displays and shared tools, impacts the roles the team members take on. We will look at both this corpus with distributed collaboration and another corpus with co-located work (described earlier in [38, 39]). In the distributed setting we have seen that the roles become accentuated [37], and that certain constraints make it more or less plausible to make successful initiatives like those reported on here. In contrast, in the co-located setting we saw indications of how the workspace helps to equalize the roles between the team members [38]. We will also continue to investigate *implicit excluding* and *explicit including* since we think that these phenomena will help us improve the design of systems for distributed collaboration.

Acknowledgement. This work was partly sponsored by EDF R&D and the European Commission through a Marie Curie Fellowship in Social Representation and Communication. I want to thank Jakob Tholander for valuable comments on an earlier draft and Michael Baker for fruitful discussions regarding the analysis. Finally I thank the researchers in the RUFAE network who have allowed me to observe and take part in their meetings.

References

1. Abel, M.: Experiences in an Explanatory Distributed Organization. In: Galegher, J., Kraut, E.K., Egido, C. (eds.) Intellectual Teamwork: Social and Technological Foundations of Cooperative Work, pp. 489–510. Lawrence Erlbaum Associates, Hillsdale, NJ (1990)
2. Baker, M., Andriessen, J., Lund, K., van Amelsvoort, M., Quignard, M.: Rainbow: A Framework for Analysing Computer-Mediated Pedagogical Debates. Submitted to the International Journal of Computer-Supported Collaborative Learning (submitted)
3. Baker, M.J., Quignard, M., Lund, K., Séjourné, A.: Computer-supported collaborative learning in the space of debate. In: Wasson, B., Ludvigsen, S., Hoppe, U. (eds.) Designing for Change in Networked Learning Environments: Proceedings of the International Conference on Computer Support for Collaborative Learning 2003, pp. 11–20. Kluwer Academic Publishers, Dordrecht (2003)
4. Baltes, B.B., Dickson, M.W., Sherman, M.P., Bauer, C.C., LaGanke, J.S.: Computer-Mediated Communication and Group Decision Making: A Meta-Analysis. Organizational Behavior and Human Decision Processes 87, 156–179 (2002)
5. Becker, B., Mark, G.: Social Conventions in Collaborative Virtual Environments. In: Churchill, E., Snowdon, D. (eds.) Proceedings of CVE'98: Collaborative Virtual Environments, Manchester (1998)
6. Berger, P., Luckmann, T.: The Social Construction of Reality: A Treatise in the Sociology of Knowledge. Doubleday, New York (1966)
7. Bradner, E., Mark, G.: Why Distance Matters: Effects on Cooperation, Persuasion and Deception. In: Proceedings of the Conference on Computer Supported Cooperative Work, pp. 226–235. ACM Press, New York (2002)
8. Bos, N., Olson, J., Nan, N., Shami, N.S., Hoch, S., Johnston, E.: 'Collocation blindness' in Partially Distributed Groups: Is There a Downside to Being Collocated? In: Proceedings of the SIGCHI conference on Human Factors in Computing Systems, pp. 1313–1321. ACM Press, New York (2006)
9. Bos, N., Shami, N.S., Olson, J.S., Cheshin, A., Nan, N.: In-group/Out-group Effects in Distributed Teams: An Experimental Simulation. In: Proceedings of the Conference on Computer Supported Cooperative Work, pp. 429–436. ACM Press, New York (2004)
10. Clark, H.H.: Using Language. Cambridge University Press, Cambridge, UK (1996)
11. Erickson, F., Schultz, J.: When is a Context? Some Issues and Methods in the Analysis of Social Competence. In: Cole, M., Engeström, Y., Vasquez, O. (eds.) Mind, Culture and Activity, pp. 22–31. Cambridge University Press, Cambridge, UK (1997)
12. Gaver, W.W., Sellen, A., Heath, C., Luff, P.: One is Not Enough: Multiple Views in a Media Space. In: Proceedings of the SIGCHI conference on Human Factors in Computing Systems, pp. 335–341. ACM Press, New York (1993)
13. Gay, G., Hembrooke, H.: Activity-Centered Design: An ecological approach to designing smart tools and usable systems. MIT Press, Cambridge, MA (2004)
14. Geyer, W., Richter, H., Fuchs, L., Frauenhofer, T., Daijavad, S., Poltrock, S.: A team collaboration space supporting capture and access of virtual meetings. In: Proceedings of the 2001 International ACM SIGGROUP Conference on Supporting Group Work, pp. 188–196. ACM Press, New York (2001)

15. Gergle, D., Kraut, R.E., Fussell, S.R.: Action as Language in a Shared Visual Space. In: Proceedings of the Conference on Computer Supported Cooperative Work, pp. 487–496. ACM Press, New York (2002)
16. Goffman, E.: The Presentation of Self in Everyday Life. Doubleday, New York (1959)
17. Hindus, D., Ackerman, M., Mainwaring, S.D., Starr, B.: Thunderwire: A Field Study of an Audio-Only Media Space. In: Proceedings of the Conference on Computer Supported Cooperative Work, pp. 238–247. ACM Press, New York (1996)
18. Isaacs, E.A., Morris, T., Rodriguez, T.R.: A forum for supporting interactive presentations to distributed audiences. In: Proceedings of the Conference on Computer Supported Cooperative Work, pp. 405–416. ACM Press, New York (1994)
19. Isaacs, E., Tang, J.C.: What Video Can and Cannot Do For Collaboration: A Case Study. Multimedia Systems 2(2), 63–73 (1994)
20. Korolija, N., Linell, P.: Episodes: Coding and Analyzing Coherence in Multiparty Conversation. Linguistics 34, 799–831 (1996)
21. Jordan, B., Henderson, A.: Interaction Analysis: Foundations and Practice. The Journal of the Learning Sciences 4(1), 39–103 (1995)
22. Kraut, E.K., Egido, C., Galegher, J.: Patterns of Contact and Communication in Scientific Research Collaborations. In: Galegher, J., Kraut, E.K., Egido, C. (eds.) Intellectual Teamwork: Social and Technological Foundations of Cooperative Work, pp. 149–171. Lawrence Erlbaum Associates, Hillsdale, NJ (1990)
23. Lahlou, S.: Observing Cognitive Work in Offices. In: Streitz, N.A., Konomi, S., Burkhardt, H.-J. (eds.) CoBuild 1998. LNCS, vol. 1370, pp. 150–163. Springer, Heidelberg (1998)
24. Lewis, D.K.: Convention: A Philosophical Study. Harvard University Press, Cambridge (1969)
25. Liao, C., Liu, Q., Kimber, D., Chiu, P., Foote, J., Wilcox, L.: Shared Interactive Video for Teleconferencing. In: Proceedings of ACM Multimedia 2003, pp. 546–554. ACM Press, New York (2003)
26. Linell, P., Gustavsson, L., Juvonen, P.: Interactional Dominance in Dyadic Communication: A Presentation of Initiative-Response Analysis. Linguistics 26, 415–442 (1988)
27. Mark, G.: Conventions and commitments in distributed groups. Computer Supported Cooperative Work: The Journal of Collaborative Computing 11(3-4), 349–387 (2002)
28. Mark, G., Abrams, S., Nassif, N.: Group-to-Group Distance Collaboration: Examining the ”Space Between”. In: Proceedings of the Conference on Computer Supported Cooperative Work, pp. 99–118. Kluwer Academic Publ., Dordrecht (2003)
29. Neale, D.C., Carroll, J.M., Rosson, M.B.: Evaluating Computer-Supported Cooperative Work: Models and Framworks. In: Proceedings of the Conference on Computer Supported Cooperative Work, pp. 112–121. ACM Press, New York (2004)
30. Pomerantz, A., Fehr, B.J.: Conversation Analysis: An Approach to the Study of Social Action as Sense Making Practices. In: van Dijk, T.A. (ed.) Discourse as Social Interaction, pp. 64–91. Sage Publications, London (1997)
31. Rogers, Y., Brignull, H., Scaife, M.: Designing Dynamic Interactive Visualisations to Support Collaboration and Cognition. In: Proceedings of 1st International Symposium on Collaborative Information Visualization Environments, pp. 39–50. IEEE, London (2002)
32. Rosenberg, D., Foley, S., Kammas, S., Lievonen, M.: Interaction space theory: A framework for tool development. In: Proceedings of the 1st International Symposium on Information and Communication Technologies, pp. 427–432 (2003)

33. Schön, D.A.: The Reflective Practitioner: How Professionals Think in Action. Basic Books, New York (1991)
34. Simone, C., Divitini, M., Schmidt, K.: A notation for malleable and interoperable coordination mechanisms for CSCW systems. In: The Proceedings of the Conference on Organizational Computing Systems, pp. 44–54. ACM Press, New York (1995)
35. Sinclair, J.McH., Coulthard, M.: Towards an Analysis of Discourse: The English Used by Teachers and Pupils. Oxford University Press, London (1975)
36. Suchman, L.: Plans and Situated Actions: The Problem of Human-Machine Communication. Cambridge University Press, Cambridge, USA (1987)
37. Sundholm, H.: To Share or Not to Share: Distributed Collaboration in Interactive Workspaces. In: Hassanaly, P., Herrmann, T., Kunau, G., Zacklad, M. (eds.) Cooperative Systems Design: Seamless Integration of Artifacts and Conversations - Enhanced Concepts of Infrastructure for Communication, pp. 270–285. IOS Press, Amsterdam (2006)
38. Sundholm, H., Artman, H., Ramberg, R.: Backdoor Creativity: Collaborative Creativity in Technology Supported Teams. In: Darses, F., Dieng, R., Simone, C., Zacklad, M. (eds.) Cooperative Systems Design: Scenario-based Design of Collaborative Systems, pp. 99–114. IOS Press, Amsterdam (2004)
39. Sundholm, H., Ramberg, R., Artman, H.: Learning Conceptual Design: Activities with Electronic Whiteboards. In: Agger Eriksen, M., Malmborg, L., Nilsen, J. (eds.) CADE2004 Web Proceedings of Computers in Art and Design Education Conference, Copenhagen Business School, Denmark and Malmö University, Sweden (2004), asp.cbs.dk/cade2004/proceedings/
40. Veinott, E.S., Olson, J., Olson, G.M., Fu, X.: Video Helps Remote Work: Speakers Who Need to Negotiate Common Ground Benefit from Seeing Each Other. In: Proceedings of the SIGCHI conference on Human Factors in Computing Systems, pp. 302–309. ACM Press, New York (1999)
41. Yankelovich, N., Walker, W., Roberts, P., Wessler, M., Kaplan, J., Provino, J.: Meeting Central: Making Distributed Meetings more Effective. In: Proceedings of the Conference on Computer Supported Cooperative Work, pp. 419–428. ACM Press, New York (2004)

Board-Based Collaboration in Cross-Cultural Pairs

Gregorio Convertino[1], Yang Zhang[2], Brian Asti[3], Mary Beth Rosson[1], and Susan Mohammed[2]

[1] School of Information Sciences and Technology
{gconvertino,mrosson}@ist.psu.edu
[2] College of Liberal Arts, Industrial and Organizational Psychology Dept.
{yzz107,sxm40}@psu.edu
[3] College of Engineering, Dept. of Industrial Engineering
The Pennsylvania State University, University Park, PA, 16803
bpasti@gmail.com

Abstract. This paper reports on an ongoing study of cross-cultural collaboration mediated by board-based collaborative systems. Twenty-one pairs (American-Chinese and American-American pairs) performed collaborative design tasks either face-to-face or remotely. Survey data, video recording, and design products were collected to examine the impact of Culture (American-American vs. American-Chinese), Medium (Face-to-Face vs. Computer-Supported), and Board-based System (Mimio™ vs. SMART Board™) on the process and outcomes of collaboration. Results from the survey showed significant effects of these variables on measures of common ground, cognitive consensus building, perceived performance, and satisfaction. The effects on perceived performance were robust. American-Chinese pairs reported a significantly lower level of consensus when using a system that supports uni-directional (Mimio™) rather than bi-directional (SMART Board™) interaction on the board.

Keywords: Culture, Board-Based, Computer-Supported Cooperative Work.

1 Introduction

Globalization of markets, internationalization of companies, diffusion of new technologies into the workplace, and the growing emphasis on teamwork in organizations are transforming the nature of work. Along with these changes, designers and researchers are facing the new challenge of designing technologies that adequately support cross-cultural and remote collaboration. With the hope of shedding light on this emerging topic, this study focuses on two main aspects of the collaboration process: common ground and cognitive consensus. The aim is to gain insights into how cross-cultural teams build common ground and cognitive consensus while performing design tasks with the aid of board-based collaborative systems that support differing levels of interaction. In prior work we reviewed research on culture and examined implications for designing and implementing Computer-Supported Cooperative Work (CSCW) systems in cross-cultural contexts [52]. In this paper we extend this review and present preliminary results from a laboratory study of cross-cultural collaboration mediated by board-based collaborative systems.

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 321–334, 2007.

1.1 Defining Culture

The first step in understanding the role of culture in collaboration is to define culture. The existing literature provides us with at least two different definitions. One definition describes culture as a constant entity that includes shared beliefs, attitudes, values, and assumptions [16, 36]. Hofstede also describes culture as "the collective programming of the mind which distinguishes the members of one group or category of people from another" [16, p. 25]. Another definition considers culture as a variable entity dependent upon and evolving together with the context. In this view, the cultural values and attitudes are seen as fragmented, variable, contentious, and "in-the-making" (e.g., [5, 33]). Collective patterns of behavior influence and are influenced by cultural values and attitudes that characterize the culture of groups, organizations, or nations [11]. A common implication of both definitions is that *culture has an important role in shaping people's collaborative behavior* [52].

A key dimension that may affect collaborative attitudes is individualism-collectivism. Hofstede [16] observed that collectivistic cultures value the ideologies of harmony, consensus, and equality more than individualistic cultures. In contrast, individualistic cultures value the ideologies of self-actualization, self-realization, self-government, and freedom more than collectivistic cultures. Individual interests and concern for privacy may prevail over the interests of the group [20, 9]. Studies show that East Asians, compared to their European or North American counterparts, are more attuned to interpersonal relationships, more concerned with group harmony, more sensitive to the presence of a common ground of knowledge, and prefer accommodation over direct confrontation [15, 19, 6, 12, 29]. Consequently, managers more often consult their superiors before making decisions (e.g.[40, 43]). However, when interacting with members of individualistic cultures, members of collectivist cultures might be more competitive [32].

These between-culture differences can translate into difficulties and misunderstandings in cross-cultural communications. In negotiation scenarios, for example, people from East Asian cultures may value relationship building and indirect communication styles, whereas North Americans of European descent may focus more on the negotiation itself and take a more direct approach (e.g. [35, 14, 39]. It was also observed that Hong Kong Chinese students responded more to cooperation and yielded to an in-group negotiator more than to an out-group negotiator (e.g. [32, 42]).

Vatrapu and Suthers [54] reviewed prior findings on culture and pointed to implications for the design and evaluation of user interfaces and collaborative applications. For example, psychology studies suggest that East-Asian learners might differ from Western learners in their attention style: the former attend more individual objects whereas the latter attend more the whole field. This can change how they carry out references (deictic communication) within a shared workspace. As a result, this may translate in different preferences for tools and visualizations for the two cultural groups that system designers and evaluators need to account for. This is particularly relevant to the design of board-based collaborative systems.

1.2 Communication, Team Cognition, Team Performance and Cultural Differences

Our description of communication draws on Herbert Clark's concept of common ground. A general function of communication is to test, refine and increment the common ground. Clark considered common ground between two people as "the sum of their mutual, common, or joint knowledge, beliefs, and suppositions" [3]. It is also defined as "the things that we know about what is known by the person we are talking to." [28, p. 270]. In this study, we primarily focus on the communal common ground described by Clark, which includes the cultural background of people such as historical facts, beliefs, norms, values, and skills. A similar but more specific concept is cognitive consensus. Mohammed [5] defined cognitive consensus as similarity among group members in interpreting key issues of the matters at hand. Awareness in collaboration is the understanding that collaborators have of others' activities, providing a context for their own activity [53]. They understand and seamlessly coordinate various aspects of collaboration: who and where are the collaborators, what tasks they are working on, when, how, etc. We view the awareness process as closely related to their efforts to ground communication and achieve consensus in decision-making – the focus of this study.

Greater amount of common ground leads to more efficient communication, coordination, and collaboration [28, see conceptual mode in 58]. Similarly, cognitive consensus has important implications for collective decision-making in organizations. Mohammed and Ringseis [6] found that groups reaching higher levels of cognitive consensus were more likely to report higher satisfaction with their collaboration and anticipate less difficulty in implementing decisions in the future. In this study, we investigate the influence of cross-cultural communication on building common ground and cognitive consensus. We expect that better common ground and consensus building will lead to better team performance, and will discuss part of our results in this paper.

1.3 Board-Based Systems

Nearly everyone is familiar with a dry-erase whiteboard on which brainstorming sessions are held, to-do lists are written, and concepts elaborated. Board-based systems are a method of expanding upon or augmenting the functions of the basic dry-erase whiteboard. The core functionality of a board-based system can include, gestures, scribbling, editing, saving, retrieving, printing, and importing.

Board-based systems have been developed to support a shared focus of attention for users [31]. Board-based systems are being increasingly used to hold virtual meetings. In these scenarios, they facilitate communication among meeting participants by allowing drawing and writing on a shared visual space. The following are few examples of collaborative board-based systems. Flatland is an augmented, pen-based computing board system designed for informal office work [28]. ClearBoard [17] is a shared drawing medium that was developed to give a seamless shared drawing space and also to give eye contact between the two users that are not at the same location (gaze awareness). This technology gives space for drawing or writing as well as space for the users' live facial pictures. Other similar systems include LiveBoard [13] and Tivoli (Pedersen et al., 1993), both of which are large interactive displays supporting virtual group meetings. They can be used for presentations and meetings as well as any kind of remote collaboration by the visually shared space and cordless pens. All the systems presented above are board-based systems used in a work scenario similar to the one considered in this study.

2 Hypotheses

Two pilot studies preceded the current study, with the first examining the effect of two collaborative systems, Mimio™ and SMART Board™, using 4 pairs of professionals performing a design task. This pilot suggested that the two above systems could lead to significantly different performances. The other pilot used 8 pairs to explore the effect of cultural and system factors on performance.

Our dependent variables in the current study include: common ground, cognitive consensus building processes, cognitive consensus, perceived quality of communication, and quality of work.

H1. Culture Effect. Compared to same-culture pairs (American-American), cross-cultural pairs (American-Chinese) will be associated with lower values of the dependent variables.

H2. Medium Effect. Compared to face-to-face collaboration, computer-supported collaboration will be associated with lower values of the dependent variables.

H3. System Effect. Compared to the collaborative system that supports uni-directional communication, the system that supports bi-directional communication will be associated with higher values of the dependent variables.

H4. Interactions. The negative impact of (1) computer mediation and (2) uni-directional collaborative system, respectively, on the dependent variables will be greater for cross-cultural pairs than same-cultural pairs.

3 Method

3.1 Study Design

The experiment has a 2 (Culture, between) x 2 (Medium, between) x 2 (System, within and nested under Medium) factorial design.

Table 1. Study design: 2 x 2 x 2 nested factorial

2 (Culture) x 2 (Medium) x 2 (System) design	**Medium (Within)**		
	Face-to-Face	*Computer-Supported*	
Culture (Btw)		**System (Btw)**	
		Mimio	Smart Board
Same-culture	11	6	5
Cross-culture	10	4	6

3.2 Participants

Twenty-one same gender pairs (13 female and 8 male) of Penn State students participated in this experiment. The participants had an average age of 25.4 (SD = 5.14). Eleven were American-American (AA) pairs, and 10 were American-Chinese (AC) pairs. The American students were born and had been living in the United States. The Chinese participants, on average, had been living in the United States for 2.5 years (three for less than 1 year and two for about 6 years). They were all enrolled students in the United States, ensuring basic proficiency with the English language. Recent research suggested that language proficiency affects meaning making, cooperation, and social dynamics in international work [56].

3.3 Procedure and Tasks

Each pair performed two 25-min floor-plan design tasks (Prison and Welcome Center), in which members assumed distinct roles with opposing priorities but were required to collaborate. Task instructions and role descriptions were provided to the participants in their original language: US Americans received instructions in US English, Chinese in Mandarin Chinese (all translations were verified with several native Mandarin speakers). Chinese students were also given the US English translation since this was the language used in the collaborative task. The experimental situation reflects the nature of cross-cultural work: one of the parties works in a secondary language or environment. This asymmetry addresses concerns of external validity. International teams across America and Asia that use English as the lingua franca are a common case. Each experimental task was performed either face-to-face with a traditional dry-erase board, or with the aid of one of the two systems: Mimio™ or SMART Board™ (SB). The order in which the media were used and the tasks performed was balanced across all pairs. A set of measures was collected at the end of each task. Procedure and materials are viewable at [57].

3.4 Tools and Lab Setting

Mimio™, by Virtual Ink [1], is used in conjunction with a physical whiteboard. Mimio™, captures what is being written or drawn on the whiteboard and makes a digital copy of it visible on a computer screen. Remote collaborators can view what is being drawn in real time through screen sharing software but cannot modify the drawing. The second system, SB, is an interactive electronic board that uses a touch-sensitive screen. The SB system by Smart was paired with a standard computer. Both users could view and edit the collaborative design. In the remote setting, the team members also used telephones with headsets to communicate verbally. Remote collaborators worked across two separate lab rooms. The lab setting remained the same across the two remote conditions. The only changing aspect was the board-based system used (Mimio or SB), with one of the two collaborators working on the board and the other on a desktop workstation (Figure 1).

The lab setting remained the same across the two remote conditions. The only changing aspect was the board-based system used (Mimio or SB), with one of the two collaborators working on the board and the other on a desktop workstation (Figure 1). In the face-to-face condition, the task was performed in the room where the boards

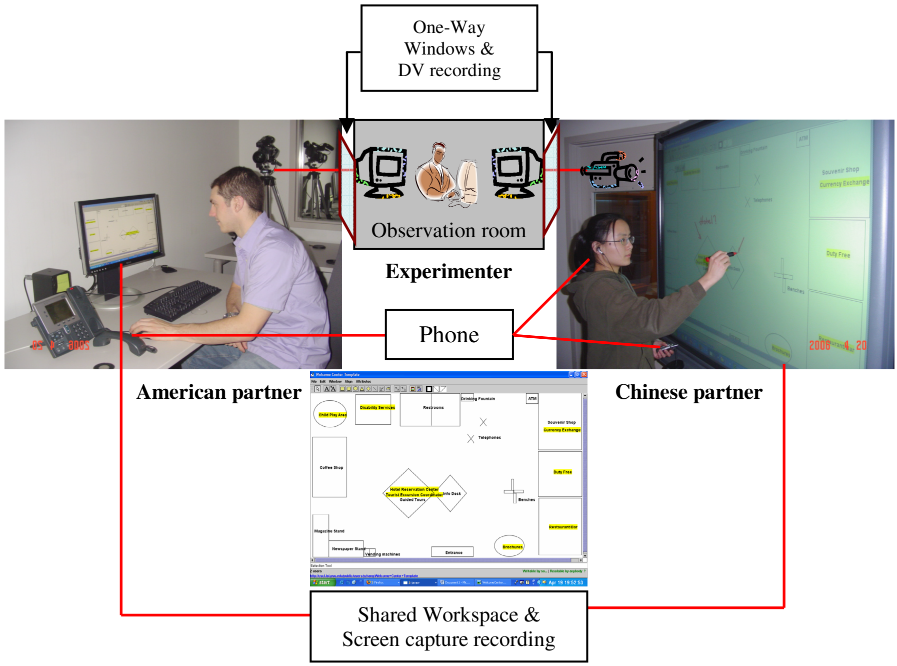


Fig. 1. Lab setting for the remote condition (Computer-Supported) and using the Smart Board system

were located, and the pairs used the same physical board that was used for the Mimio remote condition.

3.5 Measures

Prior to each run, participants received a task sheet containing the scenario and instructions. Maintaining their roles, participants chose items from a list to include in the design. They also recorded novel ideas.

Questionnaire items assessing common ground, awareness, cognitive consensus building process, cognitive consensus, perceived satisfaction, and performance were administered. The questionnaire consisted of 74-items (7-point Likert scales). For common ground, the following facets were considered (in parentheses: number of items, alpha from 1st run, alpha from the 2nd run): quality of communication (6, .80, .83), ease of communication (7, .83, .90), global communication (13, .87, .92), ease of understanding (8, .90, .89), ease of expression (8, .83, .86), and gains of shared knowledge over time (3, .74, .67). Note that global communication is composed of the items in both the quality of and ease of communication measures. Cognitive consensus measures include a cognitive consensus building processes measure (10, .75, .71) and actual cognitive consensus reached (4, .51, .54). Other measures are: awareness (8, .85, .93), perceived satisfaction (4, .92, .92) and performance (4, .90, .85). The items within each cluster were generally highly correlated among themselves, indicating good reliability of the measures. Questionnaire items, clusters and sources are viewable at [57].

After each run, participants were interviewed separately about each item in the design. In order of placement, participants recalled the pair's reasons for including that item and rated their level of agreement. They also recalled the pair's reasons for each item's final position and rated their level of agreement.

4 Results

4.1 Questionnaire Data

Effects of Culture and Medium. Considering culture and medium as main predictors and the scores from the questionnaire measures (84 cases) clustered as listed in Table 1, the effects observed are the following: culture had a significant effect on quality of communication (p = 0.017), ease of communication (p = 0.015), global measure of communication (p = 0.006), ease of understanding (p = 0.001), ease of expression (p = 0.001), perceived satisfaction (p = 0.002), and perceived performance (p = 0.007). The effect of culture on collaborators' level of awareness appeared to be close to significance (p = 0.071). Means and standard deviations are shown in Table 2.

Table 2. Means and standard deviations for effects of Culture (1 = AA, and 2 = AC) (n=44-40) and Medium (1 = FtF, 2 = Remote) (n=42). In the first column, the labels “C.01” & “C.05” indicate “p <=.01” & “p <=.05” for Culture, and “M.01” & “M.05” indicate “p <=.01” & “p <=.05” for Medium

FACETS		Culture Mean	Culture SD		Medium Mean	Medium SD
Quality of Communication (C.05)(M.05)	1	5.8	0.69		5.9	0.71
	2	5.3	0.92		5.2	0.84
Ease of Communication (C.05)(M.01)	1	5.7	0.91		5.8	0.79
	2	5.2	1.09		5.0	1.07
Global Communication (C.01)(M.01)	1	5.8	0.69		5.9	0.71
	2	5.3	0.92		5.2	0.84
Shared Knowledge Gains (M.05)	1				5.7	0.72
	2				5.3	0.88
Awareness (M.01)	1	5.6	0.92		5.8	0.75
	2	5.3	1.04		5.1	1.05
Satisfaction (C.01)	1	6.1	0.66			
	2	5.6	0.88			
Performance (C.01)(M.01)	1	6.0	0.55		6.1	0.56
	2	5.6	0.79		5.6	0.75
Ease of Understanding (C.01)	1	5.6	0.78			
	2	5.0	0.95			
Cognitive Consensus	1	5.8	0.69		5.9	0.71
	2	5.3	0.92		5.2	0.84
Ease of Expression (C.01)	1	5.6	0.7			
	2	5.0	0.9			

Medium also had a significant effect on quality of communication ($p = 0.029$), ease of communication ($p = 0.000$), global communication ($p = 0.000$), gains of shared understanding ($p = 0.015$), awareness ($p = 0.000$), and perceived performance ($p = 0.001$). Means and standard deviations are shown in Table 2.

Both culture and medium had an evident effect on collaborators' evaluation of communication. As an exemplar the item, "It was very hard to communicate effectively", was strongly influenced by both culture ($p = .010$) and medium ($p = .001$). Means and standard deviations are shown in Table 3.

Table 3. Means and standard deviations for Culture (1=AA, 2=AC) and Medium (1=FtF, 2=CM) on Question 8.2

Culture			
		Mean	Sd
p=.01	1	5.8	0.69
n=84	2	5.3	0.92

Medium			
		Mean	Sd
p=.001	1	5.9	0.71
n=84	2	5.2	0.84

Effects of Culture, System, and their Interaction. The following result pertains to the 42 cases where the two board-based systems were used. Note that the factor system (Mimio™ vs. SMART Board™) is nested under the factor Medium (Face-to-Face vs. Computer-Supported). Therefore, only one of the two levels of medium (Computer-Supported) was used (i.e., half of the 84 cases).

Culture continued to have a significant effect on global communication ($p = 0.027$), ease of understanding ($p = 0.001$), perceived satisfaction ($p = 0.004$), and perceived performance ($p = 0.012$). The effect of culture on quality of communication ($p = 0.055$), and ease of communication ($p = 0.051$) approached significance. Means and standard deviations are shown in Table 4.

System had a significant a significant effect on ease of understanding ($p = 0.045$), perceived satisfaction ($p = 0.004$), and perceived performance ($p = 0.012$). The effect of culture on collaborators' quality of communication ($p = 0.055$) and ease of communication ($p = 0.051$) appeared to approach significance. Means and standard deviations are shown in Table 4. The pairs that used SB, the system supporting bi-directional interaction on the board, rated their ease of understanding and expression significantly higher than those using Mimio™, the system supporting uni-directional communication. This can be justified by additional constraints placed on the users' ability to interact by the uni-directional system (see third hypothesis).

Interestingly, when considering cognitive consensus as the response, we observed an *interaction between system and culture*. AA pairs' ability to build consensus was unaffected by the type of system. Due to their reduced communal common ground [3] *AC pairs report a significantly lower level of consensus building when using a system that supports uni-directional (Mimio™) rather than bi-directional (SB) interaction on the board*. Cross-cultural pairs need to cover a greater gap as they ground their communication. Therefore, the reduction imposed by the unidirectional system takes a greater toll on their ability to build consensus. Figure 2 illustrates the interaction.

Building on the assumption that the quality of collaborative work is affected by grounding in communication and cognitive consensus, we tested a regression model

Table 4. Means and standard deviations for the effects of Culture (1=AA, 2=AC) (n=22-20) and System (1=MimioTM, 2=SB) (n=20-22). In the first column, the labels "C.01" & "C.05" indicate "p <=.01" & "p <=.05" for Culture, and "S.01" & "S.05" indicate "p <=.01" & "p <=.05" for System

FACETS		Culture Mean	Culture SD	System Mean	System SD
Quality of Communication (C.05)	1	5.7	0.67		
	2	5.2	0.94		
Ease of Communication (C.05)	1	5.3	1.01		
	2	4.7	1.05		
Global Communication (C.05)	1	5.5	0.68		
	2	4.9	0.92		
Ease of Understanding (C.01)(S.05)	1	5.6	0.67	5.0	1.05
	2	4.8	1.02	5.4	0.83
Satisfaction (C.01)	1	6.1	0.64	5.6	0.91
	2	5.4	0.96	5.9	0.82
Performance (C.01)	1	5.8	0.51	5.5	0.70
	2	5.3	0.88	5.7	0.78
Ease of Expression (C.01)(S.05)	1	5.6	0.67	5.0	1.05
	2	4.8	1.02	5.4	0.83

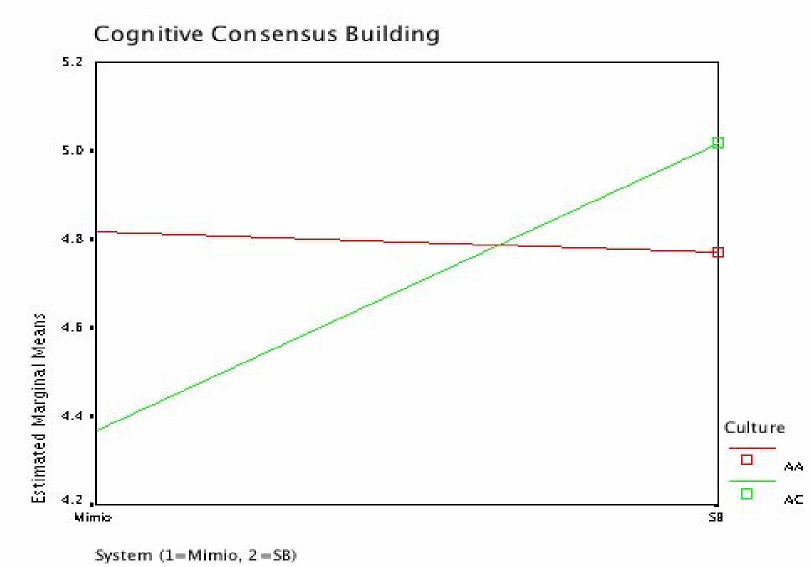


Fig. 2. Interaction between System (1 = Mimio, 2 = SB) and Culture (Red line = AA, Green line = AC). Lower cognitive consensus in AC pairs (red) using Mimio. No difference in AA pairs (green).

with performance as the response variable. The predictors were measures of global communication (13), understanding (8), expression (8), and cognitive consensus process building (10). The model is significant ($p < .001$) and the $R^2 = .47$.

4.2 Outcomes Measures

To complement the questionnaire data, we are analyzing the products of each pairs' runs, including final graphical designs and data from the post-task interviews.

Final Graphical Designs. Each pair's design was analyzed for adherence to the task rules and scored for violations. The scores were then analyzed for differences between cultural pairs. AC pairs were significantly more likely ($p = 0.043$) to manipulate objects in the template. Also, users of SB were significantly more likely ($p = 0.046$) to have final designs deviating from the specified number of items.

5 Discussion

CSCW researchers have recently pointed out a lack of research on cultural differences in computer mediated remote collaboration [8]. A major aim inspiring this research was to orient the design of effective technology for cross-cultural and remote collaboration using board-based systems. The ability to provide clear design specifications depends on the identification of relevant constructs and reliable measures that relate to culture. In this study, drawing on existing literature, we constructed a questionnaire for measuring common ground, cognitive consensus, and a few related constructs (e.g., awareness).

High Cronbach alpha values (alpha ≥ .7) from the analysis of two administrations of the questionnaire to the 21 pairs indicated that the clusters of items measuring the constructs of common ground, awareness, cognitive consensus building process, cognitive consensus, perceived satisfaction, and performance were reliable measures (see http://cscl.ist.psu.edu/public/projects/crossculturecollab/quest.html).

The questionnaire data shows an interaction effect and a main effects of culture, medium, and system.

Interaction Effect. Lower cognitive consensus in AC pairs (red) using Uni-directional system (Mimio). No difference in AA pairs (green) (see H4). Cross cultural pairs need to cover a greater gap as they ground their communication. Therefore, the reduction imposed by the unidirectional system takes a greater toll on their ability to build consensus.

American-Chinese pairs rated multiple aspects of their communication and work lower than American-American pairs. This confirms that cross-cultural pairs need to make a greater effort in order to communicate and collaborate (Culture, H1).

Pairs working face-to-face assessed their communication (quality, ease, understanding, and expression), level of awareness, perceived performance as being significantly higher than pairs working remotely (Medium, H2).

Finally, pairs using the Bi-directional system rated their Understanding and Expression higher than those using the Uni-directional system. This is consistent with Clark's prediction that common ground is more difficult to establish when constraints such as copresence and visibility are not available [5]). (System, H3).

6 Conclusions

Groups of collaborators need to agree on goals and means in order to complete their tasks. In our study, pairs of collaborators with conflicting roles performed a design task: complete a floor plan by adding a limited number of items. The task required participants to establish a shared language, exchange information about the task, understand each other's role-based priorities and negotiate. Our results show that these processes are significantly influenced by the team members' cultural background, the mode of communication, and the tool that they use.

The authors are currently extending the data analysis. Process data from video analysis and outcome data from artifact analysis are being integrated with the results from the questionnaire. The analysis is extracting and comparing regularities in the communication and the interaction with the board-based system across the different conditions (Culture: AA vs. AC, Medium: FtF vs. CM, and System: SmartBoard vs. Mimio). The data collected will be compared with data corpora from related studies of collocated and remote collaboration in cross-cultural teams [e.g., 55]

A future study with both same-culture and cross-culture teams, who work using the lab setup and method presented is being planned to refine the current findings and extend the investigation to include learning and development issues. This – we hope - will help to better understand the needs of cross-cultural teams and provide insight on how those needs can be addressed in the design of board-based collaborative systems.

Acknowledgments. We thank Lu Xiao, John M. Carroll, the members of the CSCL lab at the College of IST of Penn State University, and Susan Fussell for their support and encouragement to continue this line of research.

References

1. Apperley, M., Masoodian, M.: Supporting collaboration and engagement using a whiteboard-like display. In: Shared Environments to Support Face to Face Collaboration: A CSCW Workshop. Philadelphia, Pennsylvania, pp. 22–26 (2000)
2. Beattie, G.: Talk: An analysis of speech and non-verbal behaviour in conversation. Open University Press, Stony Stratford, England (1986)
3. Beattie, G., Cutler, A., Pearson, M.: Why is Mrs. Thatcher interrupted so often? Nature 300, 744–747 (1982)
4. Boyle, E., Anderson, A.H., Newlands, A.: The effect of visibility on dialogue and performance in a cooperative problem-solving task. Language and Speech 37(1), 1–20 (1994)
5. Brightman, R.: Forget culture: replacement, transcendence and relexification. Cultural Anthropology 10(4), 509–546 (1985)
6. Briley, D.A., Morris, M.W., Simonson, I.: Reasons as carriers of culture: dynamic versus dispositional models of cultural influence on decision-making. J. Consum. Res. 27(157), 78 (2000)
7. Clark, H.H.: Using language. Cambridge University Press, Cambridge (1996)
8. Clark, H.H., Brennan, S.E.: Grounding in communication. In: Resnick, L.B., Levine, J., Teasley, S.D. (eds.) Perspectives on socially shared cognition, pp. 127–149 (1991)

9. Cohen, D., Gunz, A.: As seen by the others: perspectives on the self in the memories and emotional perceptions of Easterners and Westerners. Psychol. Sci. 13(55), 59 (2002)
10. Daly-Jones, O., Monk, A., Watts, L.: Some advantages of video conferencing over high quality audio conferencing: Fluency and awareness of attentional focus. Int. Journal of Human Computer Studies 45, 21–58 (1998)
11. Davison, R., Jordan, E.: Cultural factors in the adoption and use of GSS. In: Information Federation for Information Processing Working Group 8.4 2nd International Office of the Future Conference. Tucson, Arizona (1996)
12. Derlega, V.J., Cukur, C.S., Kuang, J.C.Y., Forsyth, D.R.: 'Interdependent construal of self and the endorsement of conflict resolution strategies in interpersonal, intergroup, and international disputes. J. Cross-Cult. Psychol. 33(610), 25 (2002)
13. Elro, S., Bruce, R., Gold, R., Goldberg, D., Halasz, F., Janssen, W., Lee, D., McCall, K., Pedersen, E., Pier, K., Tang, J., Welch, B.: Liveboard: a large interactive display supporting group meetings, presentations, and remote collaboration. In: Proceedings of the Conference on Human Factors in Computing Systems (CHI'92), pp. 599–607. ACM Press, New York (1992)
14. Faure, G.O.: The Cultural Dimension of Negotiation: The Chinese Case. Group Decision and Negotiation 8(3), 187–215 (1999)
15. Haberstroh, S., Oyserman, D., Schwarz, N., Kuhen, U., Ji, L-j.: Is the interdependent self more sensitive to question context than the independent self? Self-construal and the observation of conversational norms. J. Exp. Soc. Pyschol. 38, 323–329 (2002)
16. Hofstede, G.: Culture's consequences: international differences in work-related values. Sage Publications, Beverley Hills, CA (1980)
17. Ishii, H., Kobayashi, M.: Clearboard: a seamless medium for shared drawing and conversation with eye contact. In: Proceedings of the Conference on Human Factors in Computing Systems (CHI'92), pp. 525–532. ACM Press, New York (1992)
18. Kroeber, A.L., Kluckhohn, C.: Culture: a critical review of concepts and definitions. Vintage Books, NY (1963)
19. Lehman, D.R., Chiu, C.Y., Schaller, M.: Psychology and Culture. Annual Review of Psychology 55, 689–714
20. Markus: International and Intercultural User-Interface Design. In: Stephandis, C. (ed.) User Interfaces for All. Lawrence Erlbaum, New York (2000)
21. McCarthy, J., Monk, A.F.: Measuring the quality of computer-mediated communication. Behaviour and Information Technology 13, 311–319 (1994)
22. McCarthy, J.C., Wright, P.C., Monk, A.F.: Coherence in text-based electronic conferencing: Coupling text and content. Journal of Language and Social Psychology 11, 267–277 (1992)
23. McCarthy, J.C., Miles, V.C., Monk, A.F.: An experimental study of common ground in textbased communication. In: Robertson, S., Olsen, G., Olsen, J. (eds.) In: Proceedings of CHI91 Conference on Computer Human Interaction, pp. 209–215. ACM Press, New York (1991)
24. McCarthy, J.C., Miles, V.C., Monk, A.F., Harrison, M.D., Dix, A.J., Wright, P.C.: Four generic communication tasks which must be supported in electronic conferencing. ACM/SIGCHI Bulletin 23, 41–43 (1991)
25. McKinlay, R., Proctor, O., Masting, R., Woodburn, J.: Studies of turn-taking in computer-mediated communication. Interacting with computers 6, 151–171 (1994)
26. Mohammed, S.: Toward an Understanding of Cognitive Consensus in a Group Decision-Making Context. Journal of Applied Behavioral Science 37(4), 408

27. Mohammed, S., Ringseis, E.: Cognitive Diversity and Consensus in Group Decision Making: The Role of Inputs, Processes, and Outcomes. Organizational Behavior and Human Decision Process 85(2), 310–335 (2001)
28. Monk, A.: Common Ground in electronically mediated communication: Clark's theory of language use. In: Carroll, J.M. (ed.) Towards a multidisciplinary science of HCI. MIT Press, Cambridge (2003)
29. Monk, A., McCarthy, J., Watts, L., Daly-Jones, O.: Measures of process. In: MacLeod, M., Murray, D. (eds.) Evaluation for CSCW, pp. 125–139. Springer, Heidelberg (1996)
30. Monk, A.F., Watts, L.A.: Peripheral participation in video-mediated communication. International Journal of Human-computer Studies 52, 775–960 (2000)
31. Mynatt, E.D., Igarashi, T., Edwards, W.K., LaMaarca, A.: Flatland: New dimensions in office whiteboards. In: Proceeding CHI '99, pp. 346–353. ACM Press, New York (1999)
32. Ohbuschi, K.-I., Fukushima, O., Tedeschi, J.T.: Cultural values in conflict management: goal orientation, goal attainment, and tactical decision. J. Cross-Cult. Psychol. 30(51), 71 (1999)
33. Olson, G.M., Olson, J.S.: Making sense of the findings: common vocabulary leads to the synthesis necessary for theory building. In: Finn, K.E., Sellen, A.J., Wilbur, S.B. (eds.) Video-Mediated Communication. Lawrence Erlbaum, New Jersey (1997)
34. Pedersen, E.R., McCall, K., Moran, T., Halasz, F.: Tivoli: An Electronic Whiteboard for Informal Workgroup Meetings. In: CHI '93. ACM Press, New York (1993)
35. Probst, T.M., Carnevale, P.J., Triandis, H.C.: Cultural Values in Intergroup and Single-group Social Dilemmas. Organizational Behavior and Human Decision Processes 77, 171–191 (1999)
36. Prus, R.C.: Subcultural Mosaics and Intersubjective Realities and Ethnographic Research Agenda for Pragmatizing the Social Sciences. State University of New York Press, Albany, NY (1997)
37. Saavedra, R., Earley, P.E., Van Dyne, L.: Complex Interdependence in Task-performing Groups. Journal of Applied Psychology 78(1), 61–72 (1993)
38. Salacuse, J.W.: Intercultural Negotiation in International Business. Group Decision and Negotiation 8(3), 217–236 (1999)
39. Schein, E.: Organizational Culture and Leadership, 2nd edn. Jossey-Bass, San Francisco, CA (1992)
40. Sellen, A.J.: Speech Patterns in Video-Mediated Conversations. In: CHI '92: Human Factors in Computing Systems, pp. 49–59 (1992)
41. Sellen, A.J.: Remote Conversation: The Effects of Mediating Talk with Technology. Human Computer Interaction 10, 401–444 (1995)
42. Setlock, L., Fussell, S.R., Newirth, C.: Taking it out of Context: Collaborating within and across Cultures in Face-to-Face Settings and via Instant Messaging. Computer Supported Collaborative Work 6(3), 604–613
43. Shakun, M.F.: Consciousness, Spirituality and Right Decision/Negotiation in Purposeful Complex Adaptive Systems. Group Decision and Negotiation 8(1), 237–249 (1999)
44. Smith, P.B., Peterson, M.F., Misumi, J.: Event Management and Work Team Effectiveness in Japan, Britain and the USA. Journal of Occupational and Organizational Psychology 67, 33–44 (1994)
45. Thompson, J.D.: Organizations in Action. McGraw-Hill, New York (1967)
46. Trubisky, P., Ting-Toomey, S., Lin, S.L.: The Influence of Individualism-Collectivism and Self-monitoring on Conflict Styles. International Journal of Intercultural Relations 15, 65–84 (1991)

47. Tse, D.K., Lee, K.H., Vertinsky, I., Wehrung, D.A.: Does Culture Matter? A Cross-cultural Study of Executives' Choice, Decisiveness and Risk Adjustment in International Marketing. Journal of Marketing 52, 81–95 (1988)
48. Van de Ven, A.H., Delbecq, A.L., Koenig, R.: Determinants of Coordination Modes within Organizations. American Sociological Review 41, 322–338 (1976)
49. Watts, L., Monk, A., Daly-Jones, O.: Inter-personal awareness and synchronization: assessing the value of communication technologies. International Journal of Human-Computer Studies 44, 849–873 (1996)
50. Wilson, P.: Computer Supported Cooperative Work: An Introduction. Intellect Books, Oxford (1991)
51. Convertino, G., Asti, B., Zhang, Y., Rosson, M.B., Mohammed, S.: Board-based collaboration in cross-cultural pairs. In: CHI '06 extended abstracts on Human factors in computing systems CHI '06, pp. 658–663 (2006)
52. Xiao, L., Convertino, G., Trauth, E., Rosson, M.B., Carroll, J.M.: An Overview of Computer Supported Cooperative Work in Cultural Context. In: Gaoui, C. (ed.) Encyclopedia of Human Computer Interaction (2005)
53. Dourish, P., Bellotti, V.: Awareness and coordination in shared workspaces. In: Proceedings of the ACM CSCW '92, pp. 107–113. ACM Press, NY (1992)
54. Vatrapu, R., Suthers, D.: Culture and Computers: A review of the concept of culture and implications for intercultural collaborative online learning. In: Ishida, T., Fussell, S.R., Vossen, P.T.J.M. (eds.) IWIC 2007. LNCS, vol. 4568, Springer, Heidelberg (2007)
55. Sundholm, H.: Social Conventions and Issues of Space for Distributed Collaboration. In: Ishida, T., Fussell, S.R., Vossen, P.T.J.M. (eds.) IWIC 2007. LNCS, vol. 4568, Springer, Heidelberg (2007)
56. Beyene, T., Hinds, P.J., Cramton, C.D.: The Impact of Uneven Language Fluency on Meaning Making and Relationships in Internationally Distributed Work. In: Poster presented at The First International Workshop on Intercultural Collaboration (IWIC2007) (2007), http://langrid.nict.go.jp/iwic2007/index.html
57. Project web page: http://cscl.ist.psu.edu/public/projects/crossculturecollab/index.html
58. Carroll, J.M., Convertino, G., Ganoe, C., Rosson, M.B.: Toward a Conceptual Model of Common Ground in Teamwork (in Press) In: Letsky, M., Warner, N., Fiore, S., Smith, C. (eds.) Macrocognition in Teams: Understanding the Mental Processes that Underlie Collaborative Team Activity, ch. 6, Elsevier, New York

Atoms of Bonding: Communication Components Bridging Children Worldwide
(Invited Talk)

Yumiko Mori

NPO Pangaea
1-20-15, Minamiaoyama, Tokyo, 107-0062, Japan
yumi@pangaean.org

Abstract. Connecting children around the world using ICT with the mind of respecting various cultures and language, NPO Pangaea is challenging to create "universal playground" where children can feel a bond regardless of their physical locations, languages spoken, or economic circumstances. We develop Package consisting contents, or activities, facilitator training program, and net environment utilizing pictogram designed by adults and children. Two years of conducting activities, over 100 occasions, to create bonds among children, four major communication components became apparent. 1) Shared tasks, 2) Shared personal information, 3) Enjoyable face to face meeting, 4) Attractive communication method. Lacking any one of these four components, it is difficult for participants, aged 9-17, to actively seek the opportunities to bond. Pangaea activities now taking places in Tokyo, Kyoto, Seoul, Vienna, and Kenya, what works for intercultural communication activities will be presented focusing on four components.

Keywords: Pictogram, intercultural communication, atoms of bonding, CMC, simultaneous activity.

1 Atoms of Bonding

9.11.2001 is the memorable date for every one of us. I am sure that all of us can recall where you were or what you were doing when the tragedy happened. For me and the co-founder, the incident happened very near from us. We were both in US, and were going to take that UA93 on 9.11. It happened that my appointment at New Jersey on 9.10 was postponed to 9.12 only three days before as I received the phone call from the company I planned to meet (Toys R Us) saying the marketing report won't be ready on 10th, that they wanted our meeting needs to be on 12th. I complained, as I had my meeting in San Francisco with Exploratorium staff on 12th, and the flight to Japan was planned on 13th. My colleague and I started the project named Pangaea.[1]

What happened on 9.11 is terrible. Yet many of us started to think why this sort of things happens, and some new projects started. ICE [4], Intercultural experiment was

[1] http://www.pangaean.org/

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 335–343, 2007.

inspired by the incident as we did and started. For some, it is to break language barrier, and for us, it is to create "Universal Playground" where children can feel the bond regardless of culture, language, or physical distance which can be considered as barriers for the formation of bond.

Internet connects the world. So we say and hear for quite some time. But has it really achieved that in last twenty years? It provided many good things, windows to the world, information access easily available for everyone. No more expensive international phone charge. We use Skype, or emails without paying postage. OK, then do we feel the connection or bond under these circumstances? For children, I must say that it limits the population. To feel a bond to another individual or group, both parties need to be themselves, in another word, they need to feel comfortable to each other before bonding takes place. In today's ICT environment, we are still not successful to create such situation where two parties can be on the same platform and express themselves freely and positively [5]. Infrastructure of ICT is rapidly developing. Internet Café can be found in Refugee Camp in Ghana, or very rural area in Cambodia. In some areas, still electric power lines are not available, nor telephone lines, yet mobile phones work there. Lines are connected, but as the world we see today, people are hardly connected and still searching the way to create ways for intercultural, inter-lingual communication passages. Computer Mediated Communications (CMC) as a new dimension of intercultural communication between East Asian and North American College students by Ringo Ma studied interactions, and presented five propositions among those who participated such interactions and those who did not [2]. Walther pointed out that in spite of its limitation on the transmission of non-verbal cues, CMC allows relational development [3]. Many intercultural communication studies seem to take an underlying assumption that it is English or another language, even though it may not be the native language. Researches in machine translation and sociological, or psychological researches seem very far from each other. Please excuse my ignorance in the world of researches. As NPO, the aim is to reach the mission ASAP, finding the most effective road to find a path to connect children, these distance between machine translation group and sociology, psychology, and education group needs to be shortened. Only having one is not usable for activists like us and many others.

This created situation that still NO ICT activities providing inter-lingual communications with people not college-educated can be found, closest is in the field of sports. Then there was a possibility to create environment, or Universal Playground. Playground can be found near everyone's house. Some are fancy with many playthings, some are just green fields, or ground. Children are genius to find something to play with. One ball can turn the ground into sports field. When different language children come to the playground, they may not interact at the beginning. However, what happens is if they come everyday, there are more chances that they will start interact. In current ICT, the lack of such common playground can be observed. Also being told not disclosing one's identity on internet, due to the unfortunate existences of net criminals, very artificial environments are now formed at regular net society.

Current ICT accommodates the same language group communication or when it is international, usually English is the choice language. At feasibility study, we faced questions many times, “Pictogram messages? Why not use English? On Internet, it is the global language.” In Japan, or in other Asian countries, many children aged between 8 and 15, do not speak or use English fluently. It is not mandatory in Japanese elementary schools, so it will stop the majority of elementary school children’s participation if we ask them to use English. Selecting any foreign language as the communication language, mission of Pangaea, creating individual bonds across the world becomes far reached, as children need to concentrate how to use foreign language. Many international exchange programs ask participants to use foreign language when they interact. This is natural as two groups don’t have a common language, the non native language is selected.

Many intercultural communication projects currently available force children to use non-native language when two groups have different native languages. A person may have rich experiences or things to say, when he/she can not speak a foreign language. It imposes problems when a person can say very good opinions in their native language, but does not understand what is going on around the roundtable because foreign language is used.

For Pangaea Project, we wanted to children, not those children from rich or elite families who can afford traveling overseas with money or qualify competitive exchange programs, but those who may never travel or thought of traveling overseas. Saying please use particular foreign language, there is a hidden implication. “It is the superior language.” UNESCO is now eagerly seeking ways how people, including children, understand to live in the era of multicultural coexistence. Project upholds the spirit of the Cultural Diversity Treaty approved by UNESCO on October 20th 2005, by creating the space where participants can learn to respect cultural diversity and recognize individuals as they are regardless of their social backgrounds. To respect others do not mean that one must like others. Understanding there are different ways of thinking and living in various parts of world, appreciating the differences, and finding similarities interesting, they are crucial for such coexistence. Being able to speak and express your own language is good start to tell story about oneself and community. Therefore, our project challenges to create contents, system, and facilitation to send the message to the world, “We have our own uniqueness, and let’s play together.”

Project uses Internet technologies effectively to transcend the distance, the dividing factor of children across the world, and to create more opportunities for them to feel the bonds. At the same time, it focuses on building personal bonds among them in the face-to-face environment. This is why it holds activities at such locations as schools and community centers, where children participate in group, talk to each other and make new friends.

Bond is all about people. Definition of Bond differs from one culture to another. At Pangaea, it is a connective feeling to want to understand children in distance using different language living in each unique culture. Atoms of Bonding appear as they keep certain lengths of repeated communications which genuinely pleasurable experiences of some sorts are produced. Only by identifying atoms of bonding, we can start seeing the effective way to utilize ICT in the most suitable manner.

2 Formula for Universal Playground

2.1 Tools for Universal Playground

Three phases of developing a feeling of bond are found. 1) Meet, 2) Communicate and 3) Bond. Through these phases, children start seeing how others understand or do not understand, different style of expressions and communications, and eventually learn how to effectively present themselves positively. We use Pictogram to break the language barrier and to prompt heart-to-heart communication. Single thing, event or the idea are expressed with different images depending on the cultural background of each child. Project differentiates itself from normal chat. Pictogram is one way to communicate in project, but there are many other ways to communicate such as drawing, creating animation, taking photo, or writing comics.

PangaeaNet is the intranet system, where children's creative works are stored and shared without being exposed to unknown public which may pose some dangers, and Pictogram messages are exchanged. It has four levels of communities with house, village, country and the earth, allowing all participants to get to the earth with four clicks from their houses. PangaeaNet requires the authentication using RFID card to log-on, to protect the security of participants.

Shortly after project started in four locations in Tokyo at its testing phase, it was discovered that children needed to learn how to present themselves effectively to others. For the effective presentation, skills are necessary. First, explain about themselves to others, and learn also how to listen to others. This led into creating activity such as "I love map". Children just put the small photo in the center, and they can write all the things they like in any ways they wish. At first they all do this in their native language. This way, they express freely without hesitation. "I love Map" surprised adults as how different each children present, yet there were many things common in which they like. Knowing what you like and telling people tells others about oneself more than we know. Children who love playing soccer can be found in Tokyo, Kyoto, Seoul, and Vienna. Common interests bring them together. Program focuses on preparing sets of activities which are fun to do, and at the same time, it gives opportunities for them to get interests to each other.

2.2 Communication Components Learned from Preliminary Activities

In last two and half years, Project has run over 120 activities, and participants in total are now over 1000. Tokyo, Kyoto, Mie, Seoul, Vienna, and Nairobi now run activity at least once a month or more. With over 70 facilitators in action, project evolves with help from engineers and researchers. Being NPO means having very tight budget. Under this circumstance, project is not able to waste any cent or yen, but plan and move quickly and efficiently if project visions to provide Bonding experiences to children around the world. Therefore, planning, observations and reviews are all very critical to correct mistakes, and to find better ways. The realization of four components came apparent from this environment. Children mostly participate regularly, once a month or twice a month. Some in last 2 1/2 years never missed the activity. Program is successful not only due to its contents, but real interactions which take places among children, differs from interactions they sometimes have in school classroom. Group of children work together, share their creations, and sometimes go outside to take photo to introduce their community in activities like Pangaea Playing Cards.

Shared Tasks. Children at Project all start from Tako Introduction. It serves as the icebreaking activity for children who gather to participate for the first time. In adult world, it is self-introduction. But for those of us, who are shy, usually get caught up in awkward situation when someone asks to tell about yourself in front of strangers. Therefore, at project, they introduce their partners who are randomly paired. This brings out the feeling of empathy, children feel sorry for not giving good introduction for their partner, and they are much more talkative to speak about their partners than about themselves. In addition, they are asked to introduce their partners in a way which others will be interested in this partner and even being liked. After this, children start drawing their houses, and rooms. All these steps are followed exactly in any locations. As they work on these common tasks, and see others' houses and rooms, participants know exactly what other children are drawing. And their creations are very unique in their own way. Through this experience, they become much more interested in not only about their creations, but differences and similarities found in other countries creations.

Shared tasks need to be fun and enjoyable to all 9 to 16 year old children. We have found that at different age, still they seem to enjoy to draw their home which to be visited by others. Some children remodel their house time to time, as they see other

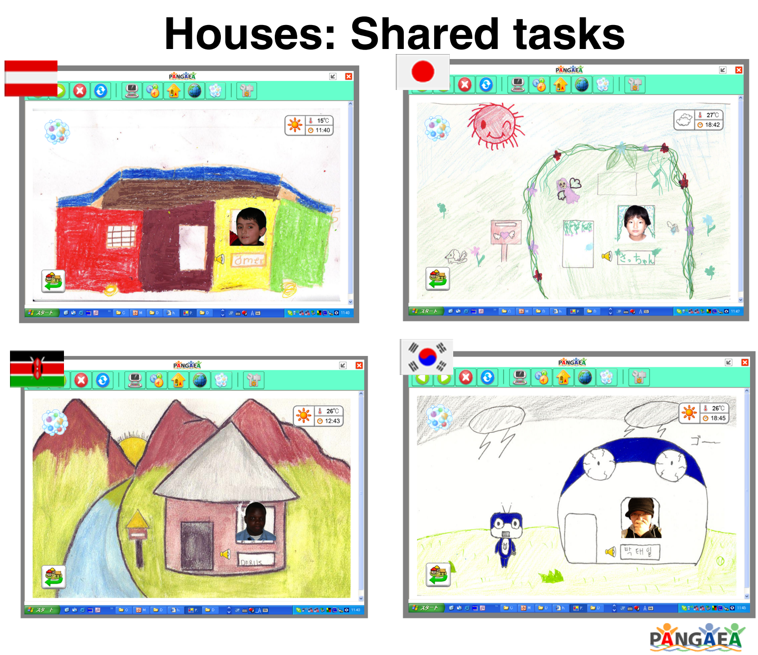


Fig. 1. Shared task

work, they want to make it more attractive. Drawing materials are crayons, markers, colored pencils, and pencils. Some will make their house looks like Pokemon Monster, and some draws the music instrument which they play. After drawing “I love Map”, the awareness of their interests becomes clear, and without any instructions, they take them in.

Tasks need to reveal personal quality of each individual. Mission is to create bonds among children. Project is not about art lesson. In order to facilitate communication in multi-lingual, multi-cultural groups, the program contents must be well planned. It must present individuals and community information which triggers many questions and comments. Program uses many analogue activities because this seems to reveal more personal quality and information than digital creation. This is partly due to the age of participants. They are still not skillful using PowerPoint, or Illustrator. Therefore, their creations appear not so different.

Common experiences make instruction unnecessary or minimum. Creating Intercultural platform, anyone will try not to have too many text information, as various language groups join, customizing to each language group will cost in terms of time and money. Project found that having common tasks, children know how things work, and what are presented where at PangaeaNet.

Shared Personal Information. ICT created artificial intelligence, but also artificial communication ground. Revealing their real names, addresses, and faces, such personal information exposure to public is considered dangerous in many countries. Korea seems to be one of the exceptions. This is the way to protect oneself at ICT. News about early teenage girl in US, who thought she was becoming a friend with the same age boy, went to see him at shopping mall one day. She then found was it was someone who was older than 40, and he kidnapped her and somehow managed to take her to Europe. In last two years, project pointed out the crucial value about providing personal information. Being NPO, working with children, project went thorough amid discussion how to consider this fact. The conclusion reached was that we find a way where non participants are not able to see these personal information. Choosing this path, amazing things were discovered.

When children have their faces from their unique windows at their house in PangaeaNet, and faces the face reflects child’s feeling on the last accessed day, this provides more familiarity. Creation, Face, Voice, and gestures, more information about an individual, they bring more closeness. It led us to conclude no personal information, no personal bonding.

Enjoyable Face to Face Meeting. FTF is now studied and Gudykunst and Kim mention about the concept of “strangers, represents the closeness of physical distances [1]. Icebreaking activity – Fun first encounter to bring two groups closer is the critical part to foster a feeling of “bond”.

At the initial encounter at FTF, both parties are excited, and also nervous. The program is planned as Ice-breaking, melting the nervousness between two groups by playing a game which does not require spoken language. Koetsuna works as it lets children shouts to compete which group has larger voice. At every Koetsuna activity, it has proven that whether they loose or win, both parties enjoy their encounter. Voicing out works to remove their tensions, and therefore, able to move into the real

Profile: Height, birthday, things I like, languages I speak
Shared Personal Information

Fig. 2. Shared personal information

small group FTF meeting. Otobiko serves to small group (4 children/group) ice-break. Now they can see face to face, more closer, and gestures can be seen from web camera. FTF is planned from the large group-small group-individual. This gives opportunity for children to get used to be on Web Cam, and much more comfortable to each other than asking them to start interacting each other without such order.

Attractive Communication Methods. The most difficult part in inter-lingual communication without selecting non-native language is how to communicate. The project developed "Communicator" system which uses pictogram for the communication. From our study, rather than putting stress, children seem to enjoy sending pictogram messages than writing in texts. Idea about able to communicate any language groups without being able to write in English excites them as our interview revealed. There have been quite interesting research seeds on this subject. Pictograms are proven to be quite effective to communicate emotions. Rather than reading texts, which many children hate especially if it is not native language, in our interviews, children expressed that it is more fun to use communicators than writing letters. However, we do not mean for children to abandon their native language of course. In fact, having exchanges of Pictogram messages, interactions in PangaeaNet, and meeting children on Web Cam activities, some got so interested in their friend's language, they started to write with them.

3 Summary

Identifying four components which break barriers between groups, and implementing all four into the program, what we observed is the imagination of how other people of group who live in distance and have different language and culture perceive your creation or message. Children are genuine to wish other people like you, not dislike you, as they reveal their own face, voice, and creations.

Shared tasks, shared personal information, fun communication method, and enjoyable FTF meeting pave the road for bonding. Bonding, as we define as emotional connection, is very difficult to measure and evaluate. One of the ways the program evaluates is to observe how children present themselves through messages or creations. At the beginning, children are devoted to Doing Something (sending messages, or drawing 4 clip comic strip). Then the awareness of their works is viewed by other children. When four components work together, at this phase, children ask themselves if their creations can be understood by others who do not understand their language and culture. In other words, they start imagining what others think in their shoes. At this phase, they will also ask people around them if their works can be understood. Following this process, they will start to understand what other people intend to say even some messages may not be so clear. Having this extra step forward to each other shortens distances between children. One time event may have an impact to children's memory, yet to foster the imagination, some prolonged time period is necessary to internalize how to be in others' shoes.

Observations made in last two and half years, over 150 activities with 1000 children in four countries, ICT can be a powerful vehicle to promote inter-language, intercultural experiences for not elite populations. Also from high tech gritted society like Shibuya, Japan to economically striving community at Nairobi, Kenya, children seem to have common interests and something to share with. Still there are long road ahead which not so many people have traveled. Researchers at this great workshop are the kind of people who are interested to pave the rough road for intercultural communication, not only by themselves, but willing to collaborate with others who may even be in different fields. Being an invited speaker at this workshop is the great honor, and I am grateful to be able to represent NPO in intercultural collaboration. There are hundreds and thousands of NPOs and NGOs around the world, but not many are so fortunate to be near information about intercultural collaboration. We hope to utilize what have been found in researches to create effective programs, and I hope our observations will be able to guide researchers to new seeds to greater researches.

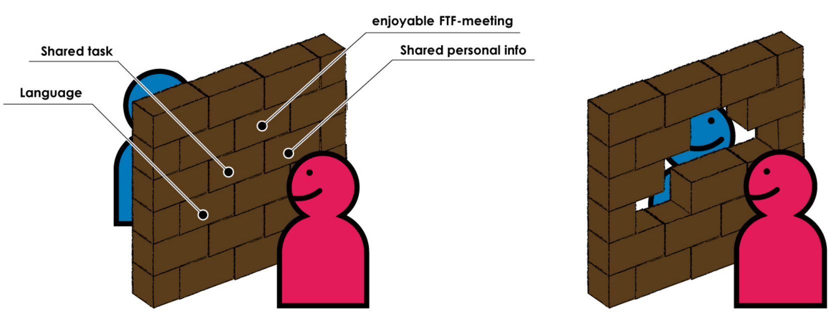


Fig. 3. Model of bonding

Acknowledgments. Special thanks to Toru Ishida who had introduced us to wonderful community of intercultural communication researchers. To Yohei Murakami who helped me to write this paper. This research is partially supported by National Institute of Information and Communications Technology (NICT).

References

1. Gudykunst, W.B., Kim, Y.Y.: Communicating Strangers: An Approach to Intercultural Communication. Random House, New York (1984)
2. Ishida, T.: Communicating Culture. IEEE Intelligent Systems, Special Issue on the Future of AI 21(3), 62–63 (2006)
3. Ma, R.: Computer-Mediated Conversations as a New Dimension of Intercultural Communication between East Asian and North American College Students. Computer-Mediated Communication, pp. 173–185 (1996)
4. Nomura, S., Ishida, T., Yamashita, N., Yasuoka, M., Funakoshi, K.: Open Source Software Development with Your Mother Language: Intercultural Collaboration Experiment. In: Proceedings of the International Conference on Human-Computer Interaction, pp. 1163–1167 (2003)
5. Walther, J.B.: Interpersonal effects in computer mediated interaction: A relational perspective. Human Communication Research 19, 52–90 (1992)

Teaching HCI: A Challenging Intercultural, Interdisciplinary, Cross-Field Experience

Cristian Rusu and Virginia Rusu

Pontificia Universidad Católica de Valparaíso, Av. Brasíl No. 2241,
Valparaíso, Chile
cristian.rusu@ucv.cl

Abstract. The extraordinary development of the Human-Computer Interaction (HCI) is still poorly reflected in South America, with the notable exception of Brazil. Our interest in HCI came from practice, as we have an extensive experience in software development. We are teaching HCI in a Chilean university (Pontificia Universidad Católica de Valparaíso) since 2003. Teaching HCI was a highly challenging intercultural, interdisciplinary, cross-field, but very rewarding experience. It was an intercultural experience, as we were born, raised, educated, with work experience both as professionals and professors in an East-European ex-communist country (Romania), but we taught HCI in a Latin-American country (Chile). Moreover, we did it in English, for Spanish speaker students. It was an interdisciplinary experience as HCI is a highly interdisciplinary science itself. It was a cross-field experience, as it allowed us to build a bridge between theory and practice.

Keywords: Human-Computer Interaction, Computer Science Curricula, Intercultural, Interdisciplinary.

1 Introduction

The *Special Interest Group on Computer-Human Interaction* (SIGCHI) of the *Association for Computing Machinery* (ACM) defines *Human-Computer Interaction* (HCI) as a discipline concerned with the design, evaluation, implementation of interactive computing systems for human use, and with the study of major phenomena surrounding them [24]. Basically, HCI tries to make people's experience with computers more productive, more time-efficient, and more pleasant.

The importance of HCI education for software professionals should be evident and well understood, when designing *Computer Science* (CS) programs, at all levels. It sounds good in theory, but is difficult in practice. Unfortunately the professors involved in CS programs usually consider HCI to be a secondary matter. This is probably a reflection of the well known conflict between HCI specialists and software engineers [8]. It has serious consequences for the training of the software professionals, as they will focus on the inner part of the software systems, ignoring the importance of the outer part (the interface). What will we get then? Top models dressed in poor clothes, and not in the "haut couture" creations they deserve!

The extraordinary development of the HCI field, both as theory and practice, is poorly reflected in South America, with the notable exception of Brazil [1], [2]. This

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 344–354, 2007.

unfortunate "rule" is followed by Chile. There is an obvious lack of HCI courses in Chilean CS programs. These courses are rather exceptions, due to enthusiasts, and are usually optional, not compulsory subjects.

We are teaching HCI as optional subject in undergraduate CS curricula in *Pontificia Universidad Católica de Valparaíso* (PUCV), Chile, since 2003. A major step forward was the introduction of HCI as compulsory subject into the curriculum of a Master Degree (MD) program in CS (in PUCV), in 2006. It was a major "battle" that we won.

2 Teaching HCI: An Intercultural Experience

Heaving a previous teaching experience in Europe (Romania), teaching in a different environment, a Latin-American country (Chile), in a language that we did not study formally (Spanish), was quite a challenge. Comparing the two systems was particularly interesting [15], [16], [18].

The problems of higher education in CS are, generally speaking, similar in any country. The education system has to develop in harmony with the systems used abroad, taking advantage of the experience of others. There are no "best" or "worst" systems. The experience of other countries should be selectively used, adapting it to the realities of the own country, and to the policy and strategy of a particular university. Exchanges are always useful and lead to benefits to all levels: personal (for professors and students), universities, and higher education systems.

2.1 Computer Science Education Systems in Chile and Romania: A Brief Comparison

The Chilean and Romanian Higher Education systems are different in many aspects. However, a changing process affected the education system both in Chile and Romania approximately in the same period, the '90s. Changes in Romania were due to the transition from a dogmatic to a liberal regime. Romania will be soon a member of the European Union (January 2007), and all changes to the education system were made in the spirit of the Bologna Process, the main purpose being the integration into the *European Higher Education Area* [21]. On the other hand changes in Chile were part of the general process of transition from dictatorship to democracy.

The education in CS in Romania has a tradition of more than 30 years. Nowadays the education system in CS has components at all levels, from elementary school to graduate programs. There are no special high schools of *Informatics* (CS) in Chile. There certainly are classes of Informatics, at all levels, from primary schools to high schools, but not special, systematic programs in Informatics, like in Romania. There are technical-oriented high schools, but not exclusively Informatics-oriented. The technical-oriented high schools are not elite schools, like Informatics high schools in Romania.

There are three traditional branches at undergraduate and graduate levels in Romania: *Informatics* (more theoretical-oriented), *Computer Science* (more technical-oriented), and *Business Informatics* (Informatics applied in Economy). The higher education in CS in Chile is quite heterogeneous. Various academic units are offering programs in Informatics, as a response to the market demand. The response is rather empiric. The traditional three branches that exist in Romania are not present in Chile. Undergraduate programs in Computer Science are generally called *Informatics Engineering* (*Ingeniería Informática*). There are two traditional types of programs:

- *Ingeniería Civil Informática* (12 semesters),
- *Ingeniería de Ejecución en Informática* (8 semesters).

A third type of program is gaining popularity lately, a 10 semesters program (*Ingeniería Informática*). The differences between *Ingeniería Civil Informática* (*Civil Informatics Engineering*), and *Ingeniería de Ejecución en Informática (Execution Informatics Engineering)* are quite obvious, but it is more difficult to establish the specificity of *Ingeniería Informática*.

Usually all Chilean undergraduate programs last an average of 2 semesters more than the similar programs in Romania. The explanation consists mainly in the higher complexity of the fundamental science programs (Mathematics, Physics, Chemistry, Biology) in the Romanian high schools, load that usually passes to undergraduate level in Chile. For instance, teaching Calculus in the Romanian high schools is a rule, but it is usually omitted in the Chilean high schools.

The Chilean curricula are usually more flexible than the Romanian ones. Students have open possibilities to plan their study program, according to their abilities and specific interests. Romanian curricula only recently became more flexible, due to the requirements of the Bologna process and the adopting of the *European Credit Transfer System* (ECTS).

The Romanian high education system includes both *theoretical* (lectures) and *practical* activities (seminars and/or laboratories), for all courses. The only teaching activities in Chile are lectures. There are additional practical activities ("*ayudantías*"), usually in the case of the compulsory courses. These are taught by students, not by professors, so their quality might be questionable.

All programs that run in Chile are taught in Spanish. There are no English programs in Informatics, neither undergraduate nor graduate. Some programs do not include the study of English as foreign language. Teaching CS courses in English is even less common. Students have a quite low level of (applied) English knowledge, and this is a major handicap when graduates want to work abroad.

There are many undergraduate programs in CS in Chile, but not a proportional number of graduate programs. There is a lack of MD programs, and especially PhD programs. The rather small number of MD programs that are currently running usually last from 3 to 4 semesters. The few PhD programs in CS last 8 semesters. There are two kinds of MD programs:

- "*científico*" – research oriented,
- "*profesional*" – technical oriented.

The PhD programs in Chile are not exclusively research-oriented, as in Romania. They include formal courses, for at least 4 semesters, then a graduation exam, that leads to the preparation of PhD thesis (PhD candidate status).

Our teaching experience in Chile takes place in *Pontificia Universidad Católica de Valparaíso* (PUCV), a prestigious and long tradition university, which is considered one of the best universities not only in the Central area (Viña del Mar – Valparaíso), but all over Chile. The university was founded in 1925, being the first university ever opened in the 5^{th} Region of Chile (Valparaíso Region). It continuously developed itself, as an academic and research establishment, having nowadays more than 12,000 students, enrolled in over 65 undergraduate and over 50 graduate programs (*diplomado*, *magister*, and *doctorado*).

The *Center of Computer and Information Sciences* (*Centro de Ciencias de Computación e Información*) of (Pontificia) Universidad Católica de Valparaíso was created in 1972. Later, in 1981, it was incorporated into the Faculty of Engineering, and became *Informatics Engineering School* (*Escuela de Ingeniería Informática*) in 1982. An 8-semesters undergraduate program of Informatics Engineering started the same year, being the first program of the new academic unit. Fifteen years later, in 1997, a 12-semesters undergraduate program was initiated. The two programs are presently running, *Ingeniería de Ejecución en Informática* (8 semesters), and *Ingeniería Civil Informática* (12 semesters), with over 650 undergraduate students. An MD program in *Informatics Engineering* (*Magíster en Ingeniería Informática*) runs in Escuela de Ingeniería Informática of PUCV since 2006. It started with 19 students, a rather large number, which certainly exceeded the expectations.

2.2 The HCI Course in a Computer Science Undergraduate Program

Once we discovered HCI as a well established science, we were tempted to share our experience with our students. As there were no HCI subjects in CS curricula in Pontificia Universidad Catolica de Valparaiso, Chile (PUCV), the only way to do that was to propose an HCI course as optional subject. We did that for the very first time back in 2003. It was a double challenge, as it was our first experience in teaching HCI, and we did it in English, not in Spanish, which is the official language in Chile (and in PUCV). The experience was highly successful, so we repeated it every year. We never intended to teach an English course; English was just a tool, but not a goal. The course was designed from the very beginning as a CS course. However, we intended to also offer an opportunity to practice applied technical English [19].

Table 1. Number of students of HCI undergraduate course

Year	No. of students	Language of teaching
2003	15	English
2004	33	Spanish
2005	26	English
2006	27	English

When we tried to teach the same course in Spanish, the interest among students was even higher, as only a relatively small number of students have enough knowledge of English to participate in HCI classes taught in English (see Table 1). This second version of our HCI course was, however, not a simple translation of the course previously taught in English. We had to redesign the course, especially due to the rather poor Spanish references. So we came to realize the disparity of HCI bibliography in English and Spanish. We do not have yet the necessary resources to offer two simultaneous versions of the HCI course, one in English, and another one in Spanish. We decided to teach only the English version, since it offers a unique opportunity to our students, being the only subject taught in English.

The structure that we are using for the undergraduate HCI course is as follows:

- *The field of Human-Computer Interaction,*
- *The nature of Human-Computer Interaction,*
- *Computer system and interface architecture,*
- *Usability,*
- *Interaction design,*
- *Web design.*

The most important area of HCI, from a software developer's point of view, is the user interface design and development. This mainly involves user-centered design and task analysis, dimensions of interface variability, dialogue tools and techniques, interface design, interface implementation, interface quality and evaluation methods.

Our course mainly focuses on Usability, which is practically its core [9], [11], [12]. We think usability is the most important attribute of the software products, and the course deals with topics like:

- *Usability Paradigms and Principles,*
- *Usability Engineering,*
- *Usability Evaluation,*
- *Usability in Practice.*

The ISO 9241 standard defines usability as the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use. Designing for maximum usability should be the main goal of the software design. There cannot be software without software users! That is why we consider usability as the most important software attribute. Two main problems occur: (1) how can an interactive system be developed to ensure its usability? and (2) how can the usability of an interactive system be demonstrated or measured?

We strongly believe in a user-centered approach for interface design and software development [3], [4], [10]. A user-centered development methodology differs from traditional software engineering methodologies in three key areas: (1) it is user centric, not data centric, (2) it is highly interdisciplinary and draws knowledge from many areas, and (3) it is highly iterative and involves as much testing and revision as possible.

The main changes that we have done since 2003 were into the theory-to-practice balance. Our belief is that we have to prepare HCI practitioners in CS undergraduate programs. Students have to prepare a final HCI project (preferable a group project), consisting in the development of user interface prototypes.

Students' feedback was excellent; they really enjoyed the HCI course. This is proved by the increasing number of the students that are choosing the course (from a wide range of available courses), the increasing number of students that are choosing HCI related subjects for their final (graduation) thesis, and the contact that they are maintaining with us over the time.

We supervised the first undergraduate thesis developed in PUCV in the usability field, back in 2003. Since than, more and more students are choosing a large range of HCI or HCI-related subjects for their graduation thesis (see Table 2).

Table 2. Undergraduate HCI thesis supervised in PUCV

Year	No. of thesis	Subjects
2003	1	• Methodology for Usability Evaluation of Software Systems
2004	1	• Usability Evaluation by Formal Methods
2005	1	• Tool Prototype for Usability Evaluations
2006	5	• Usability on E-learning Platforms • Interaction Modeling Methodologies in Software Development • Accessibility in Software Systems for Users with Mental Disabilities • 3D Interface for Operating Systems • Usability and Communicability of Software Systems

2.3 The HCI Course in a Computer Science Graduate Program

A major step forward was the introduction of HCI as compulsory course into the curriculum of the MD program in CS, a program that PUCV offers. We consider that the introduction of the HCI as a compulsory course into the new curricula was a major “battle” we won. Some of the arguments that helped us win this battle were:

- the successful experience of teaching HCI as optional course in undergraduate program,
- the interest of the students,
- the feedback of the ex-undergraduate students (now software professionals) that have graduated the HCI course,
- the demand of the potential MD students.

The last two arguments were proved by an inquiry that was made during the preparation of the MD program curriculum. Surprisingly, the software professionals seem to be more convinced of the importance of HCI than our own colleagues (professors of CS programs). The belief comes from practice, from their professional experience. We find this extremely important, as it proves the awareness of the importance of HCI practices. This is stronger and more important than any bureaucratic argument that may occur when designing CS curricula. It is particularly encouraging for the Latin American context.

The new MD HCI course represents a new challenge. Potential students are coming either from professional environment (being experienced software professionals) or from school (recently graduating a CS program, without practical experience). The audience is more heterogeneous than in the case of the undergraduate HCI course: junior professors, software professionals, and (still) undergraduate students.

Obviously, the new course has to be different. It has to include some HCI basics (for students without prior experience in HCI), without entirely repeating the content of the undergraduate HCI course. In addition, it has to bring the concepts to a higher level. We have to focus on preparing both HCI practitioners and researchers in CS graduate program. We are building the new course over the structure of the old undergraduate one, but with emphasis on different topics. The teaching process is more personalized, tanking advantage of the reduce number of students in a graduate program (19, comparing to approximately 30, in undergraduate programs).

We intend to keep the main focus on usability, and usability evaluation, but we introduce new topics, as elements of semiotic engineering that can help to bridge the gap between HCI practitioners and software engineers [6]. Besides classical evaluation methods we will also introduce a communicability evaluation method [5], [13]. Regarding the infrastructure, a usability laboratory becomes necessary, and it is already under construction.

2.4 Adapting to Changes: A Subjective Experience

Adapting ourselves to a new education system was not as difficult as we thought, probably because we experienced huge changes during the transition process in Romania. This made us more adaptable to new environments and new challenges.

One of the challenges that we had to deal with was a language that we were not familiar to (Spanish). We never formally trained in Spanish; we learned it step by step, in everyday situations. In fact, we started teaching classes in Spanish after only a few weeks of submerging ourselves in a Chilean environment. Both Romanian and Spanish are Latin languages, similar in many aspects, but this has advantages and disadvantages. The advantage was that speaking Romanian and having knowledge of French was quite easy to understand Spanish, Italian or Portuguese. The disadvantage was that having to adapt ourselves to Spanish in a very short period, we had to give emphasis to quantity instead of quality. This caused us serious problems later, as we rapidly acceded to a relatively acceptable level, but it was tremendously difficult to improve it. The lack of formal training and the (chronic) lack of time for properly studying a (still) foreign language (even if a familiar one) had long term consequences. It is particular bothering for perfectionist people as professors usually are!

We had the wrong idea that Spanish is the same, all over the world. Obviously we were wrong! Moreover, Chilean Spanish is quite different from whatever else. The difference is more difficult to deal with in human contacts, as the professor - students relationship is.

Teaching a HCI course in English, for Spanish speakers, was an even bigger challenge, as English is not our native language. The major difficulty was (and it still is) the (very) heterogeneous level of English that our students posses. As we like to improvise, to motivate the students to interact, to be active listeners, it was twice as difficult as we thought, both for our students and for ourselves.

We had to deal with a different environment, a different country, and different idiosyncrasies. Life gave us many lessons, and not all of them were pleasant experiences. Maybe the most valuable lesson was that a certain feeling of European superiority (“Chile is the end of the world!”) is completely wrong. Fortunately, we never had that feeling. Moreover, we were sometimes embarrassed by the feeling of inferiority that we perceived in the Chilean society (“Oh, Europe…!”).

As professors, maybe the biggest challenge was to adapt ourselves to a higher education system that does not explicitly include practical academic activities (seminars, laboratories), like the Romanian system does. We had to include all these activities during the regular classes’ schedule, which was quite a challenge. We didn’t count anymore on the support of an assistant professor. We didn’t even have a tutor student (*ayudante*), as optional courses do not include such activities. As we believe that a HCI course at undergraduate level should put a special emphasis on practice, it was difficult to include as many practical activities as we would.

Apart from the negative side that the Chilean education system may have, it also has a positive side: it forces students to push themselves, to solve problems without (much) external help, it builds stronger personalities. Students are more “aggressive”, more interrogative, they “don’t buy” everything that the teacher says, which is far more interesting for a professor.

Graduate students were a new challenge, as they represent a larger spectrum of academic, professional, and even HCI backgrounds. Some of them are our former HCI undergraduate students, but most of them are having the very first (formal) contact with HCI. We had to personalize the graduate HCI course more, to offer both HCI basics and advanced topics.

3 Teaching HCI: An Interdisciplinary Experience

HCI is a highly interdisciplinary area, involving mainly Computer Science (application of design and engineering of user interfaces), Psychology (application of theories of cognitive processes and the empirical analysis of user behavior), Sociology/Anthropology (interactions between technology, work, and organization), and Industrial Design (designing of interactive products). HCI focuses on Computer Science, and belongs mainly to Computer Science area. Other disciplines serve as supports. HCI uses supporting knowledge on both machine and human side. Actually HCI may be considered as a combination of Science, Engineering and Design.

We may think interdisciplinary at least at two levels: (1) students level, and (2) inter-academic units’ level.

The HCI course is very appealing for the students as it offer them a whole new world. They have a solid background in CS courses, but the curricula are definitely neglecting the most important counterpart of the interactive software systems, the users! In order to support interdisciplinary and to make connections with other subjects of CS curricula, students are encouraged to choose HCI project subjects that can help them in improving some other software development projects.

Interdisciplinary at the inter-academic units’ level is still an intention. It is necessary, it is important, it can help giving different points of view, it can support the interdisciplinary research, but it is quite a challenge in practice. There may be many reasons, but at least in our case (and the case of our university, PUCV), the major challenge is the lack of time, the overload agenda of the professors.

4 Teaching HCI: A Cross-Field Experience

Our interest in getting better user interfaces came from practice, as we have a rather extensive previous experience in software development [20], [22], [23]. Later on, we discovered HCI as a well established field of CS. So, we made our way from practice to theory, and we had the natural impulse to get back to practice. We tried to do this in our HCI classes [17].

Teaching HCI every year offered us the possibility to improve the course, and to adapt it to the necessities of local software companies, mainly based on the feedback received from ex-students. We kept the general structure of the course more or less the same, but we changed the weights of different topics.

We gradually increased the weight of practical activities, and we came to focus more and more on teaching the students how to put the HCI theory into practice. That is why we are trying to include as many practical exercises as possible. For example, students have to perform at least 3 – 4 heuristic evaluations during the semester. They also have to practice brainstorming sessions and user interviews, among others. They are always encouraged to interact, to expose their ideas, to make proposals, to analyze and evaluate them.

During the development of their HCI project, students have to apply the usability concepts, to cross-evaluate the prototypes, and to improve them based on the evaluations they performed. They have to highlight the changes and the improvements they have made, in public presentations. As the undergraduate HCI course is open for two different programs, we strongly encourage the formation of mixed teams, including students that belong to different programs. Even if they don't have yet (much) working experience, they bring different points of view, which definitely enrich the proposed solutions.

Ex-students are keeping contact with us after graduation. They offer a valuable feedback from their practical experience as software developers. They are trying hard to apply in practice what they learned in HCI classes. Even if it is difficult, and sometimes discouraging, they are actually developing, step by step, an awareness of the importance of HCI practices into their professional environment.

There is a whole debate over the conflict between HCI specialists and software engineers. Many authors analyzed the integration of the usability issues in CS curricula at different levels [14]. We perfectly agree about the existence and the relevance of the conflict, but fortunately it is easier for us to solve it, as we teach both Software Engineering and HCI. Of course, this is an exception, and the problem should be formally solved when (re)designing CS curricula. Having the mentioned privilege, we are trying to take full advantage of it, always highlighting the strong relationship that should exist between the Software Life Cycle and the Usability Engineering Lifecycle.

There are many arguments to support the introduction of HCI topics early in the CS curricula [7]. Introductory CS courses focus on systems, ignoring users most of the time. Once again we are fortunate enough to be able to informally influence the situation, because we are also teaching the very first introductory course in CS. So it is up to us to stress from the very beginning the importance of the user over the system.

5 Conclusions

A CS curriculum has to be dynamic, to adapt itself to the continuous IT changes. We consider HCI as a basic part of the formative process of the future software professionals. Changing curricula is usually a long and bureaucratic process. Improvements are faster and much easier to implement by offering new or redesigned optional subjects.

We proposed an HCI course as optional subject back in 2003. The experience was highly successful, so we repeated it every year, in CS undergraduate programs. As we strongly believe in a user-centered approach, we are mainly focusing our HCI course on usability. We try to improve the course each year, and to adapt it to the necessities of local software companies, mainly based on the feedback received from ex-students. We gradually increased the weight of practical activities, and come to focus more and more on teaching the students how to put HCI theory into practice.

A major step forward was the introduction of HCI as compulsory course into the curriculum of the MD program in CS, a program that PUCV offers. We focus on preparing HCI practitioners in undergraduate programs, but also on preparing both HCI practitioners and researchers in graduate programs.

Surprisingly, the software professionals are more convinced of the importance of HCI than our own colleagues (professors of CS programs). This proves an awareness of the importance of HCI practices into the software professional environment, and is particularly encouraging for the Latin American context.

Teaching HCI was a highly challenging intercultural, interdisciplinary, cross-field, but very rewarding experience. It was an intercultural experience, as we were born, raised, educated, with work experience both as professionals and professors in an East-European ex-communist country (Romania), but we taught HCI in a Latin-American country (Chile). Moreover, we did it in English, for Spanish speaker students. It was an interdisciplinary experience as HCI is a highly interdisciplinary science itself. It was a cross-field experience, as it allowed us to build a bridge between theory and practice.

Acknowledgments. Our thanks to *Pontificia Universidad Católica de Valparaíso* (Chile), which considers internationalization a priority, and has made possible the experience described here.

References

1. Baeza, R., de Souza, C., Rivera, C.: Enseñanza de Interacción Humano-Computador en Latinoamérica. In: Proceedings of I Jornadas de Trabajo sobre Enseñanza de HCI, Puertollano, Spain, pp. 21–31 (2005)
2. Collazos, C., Vivas, N., Ramírez, M.: La enseñanza de HCI en Colombia: Un trabajo multidisciplinario basado en competencias. In: Proceedings of XXXI Conferencia Latinoamericana de Informática (CLEI 2005), Cali, Colombia, pp. 51–58 (2005)
3. Constantine, L., Lockwood, A.: Software for Use. ACM Press, New York (1999)
4. Dix, A.: Human-Computer Interaction. Prentice Hall, Englewood Cliffs (1998)

5. de Souza, C.: The Semiotic Engineering of Human-Computer Interaction. MIT Press, Cambridge (2005)
6. de Souza, C.: Semiotic Engineering: bringing designers and users together at interaction time. Interacting with Computers 17(3), 317–341 (2005)
7. Faulkner, X., Culwin, F.: Integration of usability issues within initial software development education, ACM SIGCSE Bulletin. In: The proceedings of the thirtieth SIGCSE technical symposium on Computer science education SIGCSE ’99, vol. 31(1), pp. 296–300. ACM Press, New York (1999)
8. Folmer, E., van Welie, M., Bosch, J.: Bridging patterns: An approach to bridge gaps between SE and HCI. Information and Software Technology 48(2), 69–89 (2005)
9. Ivory, M., Hearst, M.: The State of the Art in Automating Usability Evaluation of User Interfaces. ACM Computing Surveys 33(4), 470–516 (2001)
10. McCracken, D., Wolfe, R.: User-Centered Web Site Development: A Human-Computer Interaction Approach. Prentice-Hall, Upper Saddle River (2004)
11. Nielsen, J.: Designing Web Usability. New Riders Publishing, Indianapolis (2000)
12. Nielsen, J.: Usability Engineering. Academic Press, Boston (1993)
13. Prates, R., de Souza, C., Barbosa, S.: A method for evaluating the communicability of user interfaces. ACM Interactions 7(1), 31–38 (2000)
14. Rozanski, E., Schaller, N.: Integrating Usability Engineering into the Computer Science Curriculum – A Proposal, ACM SIGCSE Bulletin. In: Proceedings of the 8th annual conference on Innovation and technology in computer science education ITiCSE ’03, vol. 35(3), pp. 202–206. ACM Press, New York (2003)
15. Rusu, C.: Computer Science Higher Education in Chile. In: Creative Mathematics, North University of Baia Mare, Romania, pp. 121–128 (2003)
16. Rusu, C.: Estudio comparativo entre los sistemas de enseñanza superior en informática entre Rumania y Chile, Final Report of Project No. 209.731/2003, Pontificia Universidad Católica de Valparaíso, Chile (2004)
17. Rusu, C.: Human-Computer Interaction. Lecture Notes, Pontificia Universidad Católica de Valparaíso, Chile (2006)
18. Rusu, C., Rusu, V.: El sistema de enseñanza en informatica en Rumania, Ingenerare. Revista Facultad de Ingeniería, vol. 15, pp. 54–60 (2003)
19. Rusu, C., Rusu, V.: Human-Computer Interaction from theory to practice, VIII Congreso Iberoamericano de Informática Educativa, Costa Rica (accepted paper, November 2006)
20. Rusu, C., Rusu, V.: Proiectarea unei interfețe utilizator în Microsoft Excel folosind Visual Basic for Applications. Lucrarile seminarului de creativitate matematica, North University of Baia Mare, vol. 10, pp. 149–166 (2001)
21. Rusu, V., Rusu, C.: El Espacio Europeo de la Enseñanza Superior entre Teoríay Práctica. In: Rivera, S., Nuñez, J. (eds.) Experiencias Docentes en Ingeniería, Argentina, vol. 1, pp. 157–166 (2006)
22. Rusu, V., Rusu, C.: Ergonomia Software in birotică. Buletinul Ştiinţific al Universităţii de Nord Baia Mare, Romania, Fascicola ŞTIINŢE ECONOMICE, Seria A, vol. 21, pp. 144–146 (2004)
23. Rusu, V., Rusu, C.: Intefata utilizator pentru functii de analiza bursiera. In: Economia contemporana. Prezent si perspective, pp. 303–312. AGIR Publishing House, Bucharest (2004)
24. Rusu, V., Rusu, C.: ACM SIGCHI Website: (ACM’s Special Interest Group on Computer-Human Interaction) http://sigchi.org

Inclusive Design Workshop by Regional Cooperation Between an NPO and a University

Takayuki Shiose[1], Kentaro Toda[2], Hiroshi Kawakami[1], and Osamu Katai[1]

[1] Kyoto University
[2] Osaka Gakuin University
{shiose,kawakami,katai}@i.kyoto-u.ac.jp, toda@upt.osaka-gu.ac.jp

Abstract. This is a case of running inclusive design workshops. Communication in an inclusive design workshop is not intercultural collaboration itself. However, as designers, engineers, and users who have individual needs, such as aged and disabled people, gather in the workshop, participants may become aware of differences in individual creation of meanings by communicating with other participants. Because as individuals our imaginative power is sometimes quite insufficient, it is not possible to sufficiently analyze such individual needs without communicating directly. In the inclusive design workshop, participants cooperate with each other during the design process, revealing differences in individual creation of meanings. In this paper, we describe some activities of social-academic cooperation, specifically cases of workshops that were run by an nonprofit organization (NPO) and a laboratory at Kyoto University. The former supports aged and disabled people's participation in society and the latter researches communication by the system theory.

Keywords: Inclusive Design, Case Study, Collaboration, University Social Contribution.

1 Overview

1.1 Aim of the Case

This case is a record of experimental social-academic cooperation activities between an nonprofit organization (NPO) and a university. In Japan, the government passed legislation allowing the existence of NPOs in 1998, and since then NPOs have completed many missions in economic and physical support of industries and the government. As the number of actual activities increases, generalization of accumulated knowledge is required to expand activities further. On the other hand, universities have led society by developing human resources and engaging in the pursuit of truth. However, a more practical contribution is required for universities these days. Many universities are increasing their activities in cooperation with industries now, but cooperation with NPOs and citizens' groups is also needed.

In this case, we show the process of cooperation between the Tanpopo no Ye Foundation, which supports artistic and cultural activities of disabled people,

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 355–367, 2007.

and the Kyoto University Symbiotic Systems Laboratory, which researches relationships between human, artifacts and environments based on system theory. The main subject of this cooperation is holding an inclusive design workshop. We propose a concrete model for social-academic cooperation through investigating how to hold constantly inclusive design workshops in which participants create designs by carefully observing the individual needs of aged and disabled people who have been excluded from mainstream design.

We describe mainly management of the workshops and communication between participants in the workshops, and actual design processes are reported in outline only. The essence of inclusive design method is in communication between people who have completely different points of view, and creating actual design results is a different problem.

1.2 Partnership

Tanpopo no Ye (Dandelion House) The "Tanpopo no Ye" Foundation was established in 1976 as a work center that supports the independence of disabled people based on a network between many people who sympathize with mothers of disabled children. Currently, it carries out projects to support artistic and cultural activities of disabled people and exchange projects between disabled people in Asian countries through artistic and cultural activities. The center consists of ten staff 30 disabled people.

Although people generally want to actively participate in their local community, in reality there are many barriers depending on individual characteristics such as disabilities. The purpose of Tanpopo no Ye is to promote a range of projects in order to shrink these barriers, and works by helping disabled people to find lifework in which they can make the most of their disabilities and provide opportunities to test their limits. Disabled people can obtain advice about life planning according to their particular characteristics and are able to experience independent life for a short time. Instead of forcing disabled people to fit into existing situations, Tanpopo no Ye creates work in which they can make the most of their own characteristics and provides opportunities to learn appropriate for them to achieve their potential. Tanpopo no Ye is open to the local community. By communicating with all kinds of people, disabled people can build rich human relationships and organize their social lives.

Symbiotic Systems Laboratory, Kyoto University Kyoto University, established in 1897, is known to have achieved satisfactory results in fieldwork especially in biology and anthropology. The Symbiotic Systems Laboratory (SSL) researches in new fields based on system science and fuzzy theory. In order to research design theory focused on relationships between humans, artifacts and environments, a field-oriented methodology of system science is required. Until now, the SSL has conducted system science research into various subjects, not restricted to hard system approaches centered on rationality and efficiency nor on soft system approaches centered on relationships and affinity. For example, the SSL is researching a communication support system between users and designers in universal design, and is researching about communication design in cooperative work between industrial designers and Kyoto traditional craftspeople based on Soft Systems Methodology. Recently, it pointed out a limit of efficiency- and economy- based manufacturing system methodology and carried out research on a community-business model based on complementary care communication with aged and disabled people to develop new system theory that can handle people's life in a satisfactory manner. Moreover, since the SSL considers that mission-based organizations such as NPOs are an advanced models of coexistence in individual local communities, it is developing and implementing human resource development programs by holding workshops to find potential needs in cooperations with NPOs.

2 Focusing on Inclusive Design

2.1 Disabilities Due to Bad Design

Disabilities include not only congenital ones but also those forced on aging or by an accident. Therefore, there is the possibility that anyone may face inconveniences brought on by a disability. Nevertheless, if we see a child jumping up to throw coins into a vending machine, we think it is just a heartwarming scene and overlook it: we do not notice that putting a ticket into an automatic ticket gate with the left hand is difficult until we brake our right hand and it is in plaster. Most of these inconveniences are imposed by thoughtless design and are realized only when we ourselves actually suffer. Disabled people do not need to apologize for their disabilities and body type. We can change the definition of "disabled people" not by ascribing inconveniences to disabilities, but by considering inconveniences as a challenge of design. Thinking deeply about design is thinking profoundly about what is aging and what are disabilities.

2.2 Beyond Universal Design

Many universally designed living goods are becoming increasingly available. These include tools such as scissors that can be used with both the left and right hands, as well as home electrical appliances like oblique drum washing machines, which are easy to use regardless of the user's body type. Universal design is a design criterion by which we try to make products and space usable for everyone despite a person's ages, body types or disability. However, the word "universal" may arouse the thought that the uniform design will replace other designs. From the perspective of disabled people, it is pointed out that it is difficult for them not to feel imposed on by "the standard goods" that are considered usable by everyone.

In fact, not a few people feel uncomfortable with universal design as the unified solution in Europe, where various communities coexist. It is fair to say that the idea of inclusive design meaning that deeply individual needs are observed and considered in design receive strong support in Europe[1].

If there is a good that is easy for everyone to use, it is natural that people will want it. However, trying to obtain the one universal design solution is impractical: it is more effective to experience directly the process in which the individual needs of aged and disabled people are observed and included in society through creating designs. Inclusive design policies, such as "start from careful observation of an individual" and "extract user's inconveniences without preconceptions" are not special: they merely outline ordinary processes. Instead, it is important to actually go out into the field to think and act from this perspective. There must not be reluctance to communicate directly with excluded people to understand their circumstances.

2.3 Communication Reality in Design Workshops

The main feature of inclusive design is that disabled people not only participate as test users, and are surveyed on usability, but also participate in upstream operations such as a basic design and survey analysis[1]. For example, when conducting surveys in the field, participants not only interview a disabled person but can also spend a time with a disabled person in several situations of everyday life such as making an ATM withdrawal and buying a ticket at a railway station. If disabled people participate in a focus group or a user group for basic design, or join in a workshop as design partners, their contribution can become greater. For example, they might propose concrete new designs in cooperation with designers. At the Royal College of Art, which is famous as the center of inclusive design, disabled people members have formed the regular focus group to assist designers in five years, highlighting their excellent communication abilities.

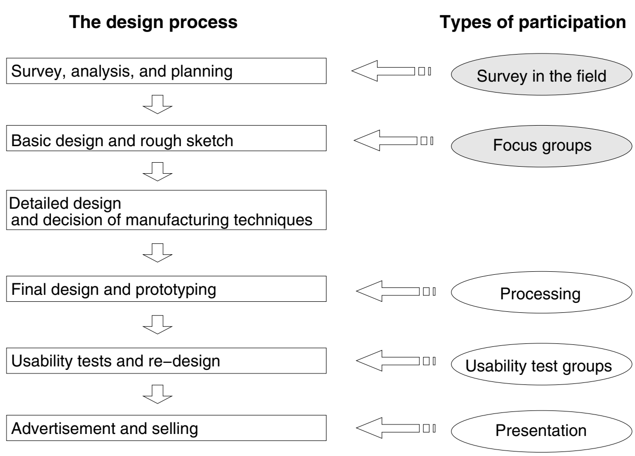


Fig. 1. Various types of participation in design processes

A workshop is not a one-way method of transmitting information from the chairperson to participants, but a method of mutual communication between participants in which they act on their own initiative and gain first-hand experience[2]. Nowadays, there are various workshops such ones on school education, and business and town development with the participation of residents. The most important point is that reliable information is exchanged by participation of the parties concerned on their own initiative.

2.4 Intersection of Motivations

Disabled people sometimes despair about society, making them despair about themselves and lose the opportunity to engage in self-affirmation. This is difficult to understand for people who are not aware of forced inconvenience due to bad design. Education these days has enlarged the chasm between specialists and non-specialists, thus creating a trend toward placing great value in the words of specialists than those of the parties concerned. Reaching beyond the simple transmission of words from aged and disabled people, it is important to experience the verbalization process of what the parties concerned feel and need together with them. Tanpopo no Ye, as one partner of this social-academic cooperation, hopes that participation by disabled people in inclusive design workshops will become one option of societal contribution for them, and adopt regular running of workshop as one of their projects. Kyoto University's Symbiotic Systems Laboratory, which is another partner in this project, regards regular holding of the workshop as the concrete field which requires new system theory, not conventional system theory based on efficiency and economy.

Until now, aged and disabled people have not been recognized as mainstream users even in product design based on universal design. To obtain common user needs, we tend to conduct surveys based on the hypothesis that mainstream users are adults with no disabilities. However, as many countries have a rapidly aging population these days, aged people will become the majority and an ever-increasing number of people will realize loss of body functions due to aging. It is important, therefore, also for economic reasons to establish a method for running inclusive design workshops. In addition to recognizing of aged and disabled people as mainstream targets, the opportunity for designers, engineers, social welfare workers and public servants to adopt concrete practice. Moreover, everyone must realize that it is possible for all of us to face inconvenience imposed by bad design.

3 Learning in Inclusive Design Workshops

3.1 Variations of Trials

Kyoto University's SSL hosted six inclusive design workshops in December 2005 to August 2006. Analysis of the activities revealed that the concentration of decision making onto one key person causes problems about the continuation and broadening of the activities. Accordingly, we considered it necessary to develop a system for fostering human resources capable of running workshops and accumulating know-how on them. That is to say, there is a requirement for a graduate school education curriculum on searching social needs on research of informatics through co-creation with NPOs. Recently, there has arisen the need to develop people with leadership skills in graduate school education, and we expect coordination and management of design workshops, which often include much conflict, to provide valuable experience for graduate students.

Most of the workshop participants are undergraduate and graduate students in the Kansai area, mainly Kyoto. Their majors cover a lot of fields such as mechanical engineering, information systems, product design, and the study of disabilities. Participants are arranged to groups in a way in which at least one person who majors in one of these four fields is included in each group. This is because we believe that there is a need to specify specialties necessary for creating original designs before starting communication between people from various fields. Previously, we had difficulty in ensuring the quality of outputs when we arranged communication between people from desultory various fields. Thus our previous experience causes us to believe what we described above. Moreover, inviting participants from many fields is a request of Tanpopo no Ye. They wants to provide opportunities for profound communication with disabled people to students from as many fields as possible.

Disabled people are invited by Tanpopo no Ye, and it is the major objective for Tanpopo no Ye to provide opportunities for disabled people to join a design group as new work for them.

In addition to students and disabled people, we invited professional designers, without concern for whether they were freelance or employed with a manufacturer. There are two reasons for this: one is to improve the quality of design in the workshops; the other is for professional designers to be available to give their opinions about what is necessary to regularly run inclusive design workshop in cooperation with various enterprises.

The inclusive design workshop provided a learning environment with two aspects for students. First, they could learn communication skills to collaborate with all kinds of participants in the workshop. Second, they could learn communication skills to manage the workshop itself through negotiations with NPOs. It is important that they became capable of finding a problem from the direct communication in the field, not being satisfied with finding a problem in learning for its own sake.

3.2 Details of Workshops

To accumulate know-how on running the inclusive design workshops in which disabled people participate on their own initiative, we attempted to run six workshops in various styles. For example, the time for a workshop was short in WS-I, II, and V, and a long time in WS-III, IV, and VI. Furthermore, the aim of a workshops was not same. In WS-I, II, V, and VI, the type of products in which participants contributed to the design of was set beforehand, whereas in WS-III and IV the participants decided on the product themselves. It is expected that disabled people who participate in the inclusive design workshops are diverse in their motivation and standpoints, and the needs of enterprises that can cooperate with us are also diverse. Therefore, we consider that running various workshop patterns experimentally enables us to hold workshops respond to the diverse requests of enterprises and society.

On holding various workshops, we decided on venue size and relative proportions of participants from each field in advance because these factors have nothing to do with requests of enterprises and society. We adopted three control variables for running workshops: a time, a target for design, and a method for governing how users participate.

Table 1. 2005-2006 Inclusive Design Workshops in Kyoto, Nara

	WS-I	WS-II	WS-III
Date	9 Dec. 2005	28 Feb. 2006	1-3 Mar. 2006
Time	120 min.	150 min.	1260 min.(3 days)
Target for design	Fixed: Electric Pot	CyARM	Not fixed
Participants	Disabled: 4 Student:1 Others: 5	Disabled: 5 Student:3 Others: 4	Disabled: 4 Student:17 Others: 5
Facilitator	Prof. Hirai	Mr. Nishitani	Ms. Cassim
Participating as	User group	User group	Research group

	WS-IV	WS-V	WS-VI
Date	6-8 Mar. 2006	29 Jun. 2006	24-25 Aug. 2006
Time	1260 min.(3 days)	180 min.	720 min. (2 days)
Target for design	Not fixed	Picture Books	Fixed: Electric Pot
Participants	Disabled: 2 Student:26 Others: 7	Disabled: 4 Student:4 Others: 3	Disabled: 4 Student:6 Others: 10
Facilitator	Ms. Cassim	Mr. Nishitani and Mr. Nishikawa	Prof. Arai
Participating as	Research group	Research group	User group

In WS-I, people with cerebral paralysis and wheelchair users participated. The target for design was set as an electric pot. In the workshop, disabled people told of their experiences whereby buttons on a pot are placed too high for wheelchair users and that the handle of a pot is too unstable for carrying to draw water.

Fig. 2. A scene of WS-I

In WS-II, visually impaired people and wheelchair users participated. The target for design was "CyARM," which is new device developed at Future-University Hakodate. CyARM is a device for visually impaired people that employs an infrared sensor to perceive a distance between themselves and a wall ahead. Because it was an unfamiliar device to the participants, they could not tell of their experiences or utilize it at all.

In WS-III and IV, various types of disabled people participated, and Ms. Julia Cassim worked as a facilitator. The target for design was not fixed in advance: it was decided on the three-day workshop through communication with disabled people. For example, members of one group in WS-III went to a video game arcade. Although it is difficult for disabled people to enjoy games in which the intensity of motion is essential, such as fighting games and racing games, games in which the fun comes from communication with other players itself, such as purikura (print club, which provides sticker sheets of 16 miniature photographs of the player and players can exchange them with each other) can be enjoyable for disabled people. Based on this observation, they proposed a rhythm game that is playable by anyone, even if a player is visually impaired or wheelchair-bound. Members of another group proposed a new system for making individual cups of filtered coffee, which makes easy for the visually impaired or those with a disabled arm to make coffee.

In the long-term workshops (WS-III, IV, and VI), we provided enough time to give a lecture on inclusive design and an orientation session for the design workshop. Presentations of each group were evaluated with respect to the main idea and the presentation itself separately. In the short-term workshops (WS-I, II, and V), the lecture and orientation session were omitted and presentations were given only briefly. In all workshops, the number of group members was set to six to eight.

In order to learn the inclusive design method, naturally lectures are important, actual experiences of communication with users in a workshop are more important. These experience that they never obtain in a classroom lecture improve their communication skills.

3.3 Qualitative Research Approaches to Unconscious Needs

A characteristic of the focus group method is that we do not need to prepare structured queries because the existence of other members restricts a member's answers[3] [4]. However, systematization of the method as a needs extraction method is still insufficient. For example, it is unclear who becomes an interviewer and who becomes an interviewee. We consider it important for participants to listen to a lecture to acquire practical interviewing methods such as the Focused Interview[3] and the Ethnographic Interview[5]. Using the Focused Interview method, we can restrict an interviewee's answers using unstructured queries. For example, at first we ask, "Was there any impressive event in this system?" and then ask, "By this presentation of information, could you suppose a next operation clearly?" to structuralize it. Using the Ethnographic Interview, we can find what an informant sees, enjoys, and how the informant acts in the world at large as well as in his or her culture and daily life through natural conversations via descriptive queries and contrastive queries.

In WS-I, the target for design was set as an electric pot, with participants holding a discussion as mentioned below.

We show here a typical conversation.

[Conversation 1]

(Participants observe C pouring hot water. Because C's fingers are trembling, a pot totters.)

F: Ah, it's really tottering!

A: If it were a cup, it might jump up.

D: Yes, if it were a cup, it would.

H: *I want it to be more stable*, to be fixed tightly. It turns too easily. D: It moves too much.

F: It turns only if you touch it. It's not necessary to turn so much.

(Participants observe A pouring hot water. Similarly to the case of C, a pot totters dangerously.)

G: Just as we saw, loose fixing makes ...

F: Loose fixing is pretty troublesome.

H: The same thing happens when I use a PC. (Showing trembling fingers) I mean that I might hit the same key twice.

C: How about adding a lever below to fix it?

Most electric pots can be rotated left and right without lifting. However, C and A suffer trembling fingers due to cerebral paralysis. Consequently, when they try to push a button on the electric pot, they rotate the pot unwillingly. This seems to highlight certain individual needs of disabled people at first glance: in fact, people having no such disability control this movement by applying more strength with their fingers. In reality, not a few people have spilled hot water regardless of a disability. This function of turning the pot easily may be undesirable for all people when pouring hot water, and discovery of this possibility is a good example of the discovery of potential users' needs.

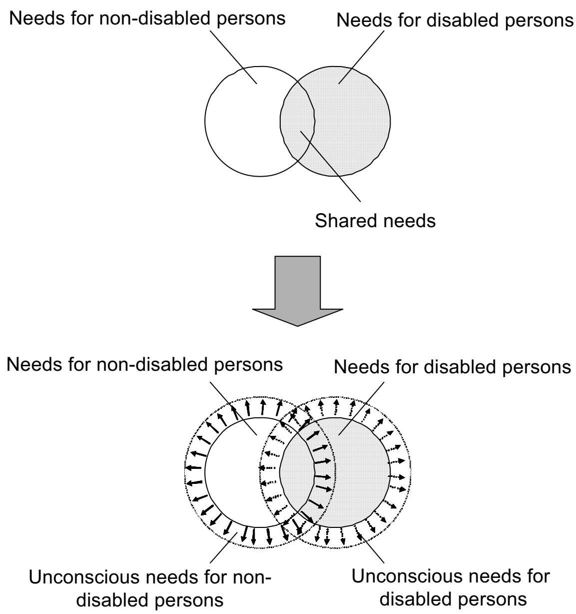


Fig. 3. From discovery of common needs to discovery of potential needs

Fig. 3 shows change of needs which results from process of Inclusive Design. We don't aim to find shared needs of disabled and non-disabled people by Inclusive Design. By focusing on actual scenes of daily lives, we can discover unconscious needs of individual people. Moreover, we can obtain persuasive explanation of the needs based on environmental factors and constraint.

The following discussion is also suggestive.

[Conversation 2]

I: *It is difficult for aged or disabled people to use fire, so we use a pot at all times. All year round.* So the inside of it is often covered by something like lime. This is why I'd like an auto-cleaning pot... Besides, once the lime sticks, it's difficult to remove. Though it says water purification by minerals, I'm afraid that it may worse for health than without it.

H: It seems bad for health, rather.

F: Plus, many people don't replace the filter of their water purifier.

G: Right, using it seems rather dirty.

At first glance, a self-cleaning electric water pot is something wanted by everyone, regardless of whether they have a disability. Nevertheless, the comment that, “It is difficult for aged or disabled people to use fire, so we use a pot at all times. All year round.” shows a critical difference between the needs of disabled and able-bodied people. It is actually impossible for seriously disabled people to boil water because of the danger involved. Therefore, an electric pot must be filled at all times, and their true need is that a pot be ready to pour hot water at all times and clean itself regularly, too. What is important in the inclusive design workshop is to realize a subjective aspect of a relationship between disabled people and products, such as anxiety and joy, not only an aspect which can be analyzed ergonomically.

3.4 Mutual Understanding Through the Workshop

The objective of the inclusive design workshop is not only to determine the needs of aged and disabled people. Needless to say, an abstract query such as, “What causes trouble for you?” produces no creative results. Needs are revealed through conversations with aged and disabled people through conversations about their daily lives. Moreover, it is expected that this process will lead to the discovery of potential needs regardless of disabilities, something that often tends to be overlooked.

Participating in the inclusive design workshop is completely different from thinking about design through simulated experience, such as by using a pregnancy simulator and a senior simulator (for example, using glasses to simulate glaucoma). Disabled people’s internal movement of mind – what they feel at that time – is revealed only through actual conversation with them. Comments such as, “I’m afraid of splashing hot water”; “What can I do when I fall over?”; or “I was happy when it was like this.” do not appear in ergonomic analyses or cognitive engineering. It is not difficult to imagine that we would be happy if an electric pot had an automated cleaning function, yet we cannot imagine an earnest wish of disabled people who use electric pots, without a workshop. Professor Arai, who facilitated WS-VI, emphasized that it is important that disabled people talk about their experience in reference to their own specific cases, not in general terms.

“I was at a loss when they said that there is no problem”. The results of questionnaires distributed to participants revealed symbolic impressions of students about the workshop. At first, students participating in a workshop would ask “Are there any problems with this?” Unexpectedly, Answers to such a question often showed that disabled people had already overcome a problem using their ingenuity. Disabled people also taught us that it is not true that there is no system to support disabled people, but that such systems are often wasted because the people who supports them do not sufficiently understand the situation. People want to dress up and enjoy themselves regardless of disabilities or age. Thus the inclusive design workshop is an opportunity for teaching afresh that this just a matter of course. These subjects are hard to handle from an engineering aspect alone.

4 Community Business Model for Inclusive Design

We ran an inclusive design workshop to provide an opportunity to experience the verbalization process of what the parties concerned feel and need together with them, not to simply transmit words from aged and disabled people. In the inclusive design workshop, we could obtain a renewed recognition of disabled people as sensitive users by communication with disabled people through working closely with them. Their suggestions clearly revealed inconveniences that able-bodied people avoid unconsciously. It is difficult for designers to understand that users do not comprehend what is obvious to designers, and vice versa. Therefore, suggestions from these sensitive users are extremely valuable.

For the Tanpopo no Ye Foundation, a new type of learning environment was provided through this workshop, and for Kyoto University's Symbiotic Systems Laboratory, a research environment to establish new design theory of systems considering life itself was created. Nevertheless, this social-academic cooperation cannot prove its worth if this cooperation is only short-term. To ensure that this project is self-sustaining, there are many problems to be solved such as establishing the identity of this project in wider society considering needs of enterprises, not to speak of economic problems. For this reason, we must analyze many factors such as lead time, costs, and a form of qualitative research for publishing.

Especially, securing the right people to manage this project is a difficult issue. In many cases, most decision-making is concentrated on one key person, making long-term continuation and widening of the project difficult. In this case, we tried to develop human resources and run the workshops simultaneously by leaving negotiation and decision-making tasks to three graduate students. Although in the early stages of the project they could only act as instructed by teachers, after several workshops they displayed leadership often.

Furthermore, it is important to obtain feedback from participants when we run the workshops regularly. Normally we distributed a questionnaire to participants to obtain feedback. However, whether the feedback is valuable depends mostly on the intuition and experience of the questionnaire writer. In future, we must establish a methodology for including participants' suggestions in workshop management.

The Tanpopo no Ye and Kyoto University hopes that the inclusive design workshop provides the new work for disabled people. Nevertheless, the present system is inadequate to accept a consultation of testing from enterprises. There are three reasons for this. First, we cannot warrant the quality of design results. Participation of disabled people does not necessarily provide the new idea. In the future, we will have to clarify effects of communication with disabled people on design. Second, it is difficult to protect intellectual properties. For enterprises, participation of people outside business activities such as students and disabled people means that there is a possibility that important information about product development leaks out. To avoid a leakage, students and disabled people must be well aware of intellectual property right, and treat carefully ideas in the workshop. Third, decision of a schedule is also problematic. Although enterprises want to work on weekdays, disabled people who are enthusiastic about participation and expression in the workshops often have their regular occupation. Consequently, it is difficult for them to spare a time on weekdays, and not a few of them want the workshop on a holiday. These three problems prevents us from establishing a community-business model by which we can collaborate with enterprises.

In the future, first we intend to systematize communication in the workshops. For that purpose, it is important to announce that an inclusive design workshop is a learning environment of communication for students, designers and disabled people rather than an environment for production of new practical ideas. It is also desirable that disabled people and students understand the workshop as a learning environment with an admission fee, rather than as a place where they work as test users and are paid.

References

1. John, C., Roger, C., Simeon, K., Cherie, L.: Inclusive Design -Design for the whole population. Springer, Great Britain (2003)
2. Robert, C.: Participatory Workshops: A Sourcebook of 21 Sets of Ideas and Activities. Earthscan Pubns Ltd. (2002)
3. Merton, R.K.: The Focused Interview and Focus Groups: Continuities and Discontinuities. Public Opinion Quarterly. 51, 550–556 (1987)
4. Richard, A.K.: Focus Groups: A Practical Guide for Applied Research, 3rd edn. Sage Pubns, Thousand Oaks (2000)
5. Spradley, J.P.: Participant Observation. Rinehart and Winston, New York. Holt (1980)

How Intercultural Disaster Reduction Education Change Students: A Case Study of an Evening Course Senior High School in Hyogo, Japan

Yoshie Naya

Hyogo Prefectural Institute for Educational Research and In-service Training,
2006-107, Yamakuni, Kato-shi, Hyogo,673-1421, Japan
yoshie.naya@gmail.com

Abstract. The year 2005 was the 10th commemorative year of the Great Hanshin-Awaji Earthquake. To commemorate the 10th anniversary, one of the evening senior high schools in Hyogo, Japan tried an on-line intercultural project in the period of integrated study. Students introduced their own experience of the earthquake on the web so that those who read their experience may prepare for future disasters. Six months study of disaster reduction and the participation in the international conference, 'Natural Disaster Youth Summit' changed the students' attitude to natural disasters and view to global problems. In this report, the process of the project and the change of the students' attitude will be described. Also, theoretical hypothesis of on-line international collaborative learning will be considered.

Keywords: On-line intercultural project, Disaster reduction education, Period of integrated study, Natural Disaster Youth Summit.

1 Introduction

The year 2005 was the 10th commemorative year of the Hanshin-Awaji Great Earthquake. To commemorate the 10th anniversary, one evening senior high school in Hyogo, Japan tried a new study aiming at students to introduce their own experience of the earthquake on the web so that those who read their experiences may prepare for a future disaster.

By participating in the project "Natural Disaster Youth Summit" led by the International Education and Resource Network (iEARN), an international education network linking schools of the world online by using Information and Communication Technology (ICT), and by exchanging information and views on the electronic forum on the Web, the students were able to learn about disaster reduction through international/intercultural exchanges and in cooperation with students from fifteen other countries. The students published their blog covering their exchange of disaster

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 368–381, 2007.

bears (stuffed toys) with overseas schools as a part of disaster-reduction education, and received comments from the exchange counterparts, enjoying the interactive activities of the project. By reflecting on, writing about, and Web-publishing their own earthquake experience, the students developed their ability to communicate information. They also searched the Internet to acquire disaster information. The students obtained information directly from overseas or the Internet via the iEARN based on ICT, to develop their inter-cultural communication and info-communication skills as well as feel a bond with other children in the world. ICT also helped the teacher manage the progress of the project in cooperation with other teachers in Japan and other parts of the world.

The project was done as a period of integrated study. This subject area started from 2003 in upper secondary education in Japan. Since a period of integrated study does not have its limit to the theme, project based learning is easily applied to this class. Although the need for disaster reduction education is widely recognized and the environment of applying project-based learning in the school is settled, most schools are satisfied with doing emergency drills and evacuation drills only. To spread the practice of disaster reduction education in many schools, we may need a model which students participate in and enjoy the learning.

This case study will be one of the models for disaster reduction education for other schools. Also theoretical hypothesis of on-line international collaborative learning will be considered.

2 The On-Line Intercultural Disaster Reduction Education

As Abhyankar (2006, p.13) said "In any calamity whether natural or manmade, children are worst affected since they are most vulnerable to such adversities", children are the most needed for disaster reduction education. In this chapter the need of disaster reduction education and intercultural education is examined. Then, the characteristics of on-line disaster reduction education and its structure are considered. Difficulties of practicing the on-line international disaster reduction at school and possible solutions to the difficulties will also suggested.

2.1 The Need of Disaster Reduction Education

Many natural disasters have occurred in the world and people from where disaster occurs suffer large damages. Sadly, children are the worst victims of natural disasters. Acquiring the knowledge and preparing against disasters are effective in reducing the damages from disasters. However, people have not paid much attention to take measures before disasters (disaster prevention) so far. This is a significant issue for disaster management. Disaster education for students is one of the solutions. Education can let students take measures not only at present but also in the future.

2.2 The Need of International / Intercultural Exchange Learning

The main objective of international / intercultural exchange learning is often thought as the understanding of different cultures. Therefore, many people have thought that international/intercultural exchange learning is just to show their cultures to each other. But the important objective of international cultural exchange is to raise awareness of being a global citizen through understanding different culture and people in different areas. Globalization is taking off. In the world of the 21 century, problems such as natural disasters, environmental problems and conflicts between different cultures occur in various places. Many of the problems can not be solved domestically. Therefore, people who have global view points and competence to collaborate with people from different cultures are in need.

2.3 The On-Line Global Collaborative Learning

To foster global citizenship among children, intercultural collaborative learning is expected for education, because we may expect that those who experienced such education grow as global citizens and leaders of their communities. It used to be difficult to have such a learning environment at school. However, the improvement of technology enabled the on-line global collaborative learning environment. It made possible to have global collaborative learning environment by utilizing ICTs at school. Students can work and learn together with students from different schools in different countries through ICT at classroom. Since the on-line global collaborative learning environment is still innovative, we need theoretical background and more practices of utilizing it in education.

2.4 Characteristics of On-Line Intercultural Disaster Reduction Project

Ordinary classroom lessons are focused on their own subjects. For example, English class is focused on mainly language skills. However, different from traditional subjects, an on-line intercultural collaborative project basically consists of three factors. They are a) Thematic subject, b) ICT skills and c) Intercultural communication / collaboration. We can illustrate the concept of on-line intercultural project in figure 1. The area that all three factors are overlapped is on-line intercultural collaborative project.

These three factors are closely linked and none of the factors can be ignored. The problem is that these three factors are based on different educational fields. ICT skills are taught in information class. Intercultural communication is taught in foreign language class. Thematic subject is taught in different subject class.

This means when a teacher decides to do the on-line intercultural project, he or she has to pay attention to those fields. This multiple tasks make teachers feel difficult to practice the on-line global project, because teachers usually have special skills on one field, not multiple fields. As a practitioner, the author had to face lack of confidence in some fields. To solve this problem, it may be ideal to do the on-line intercultural

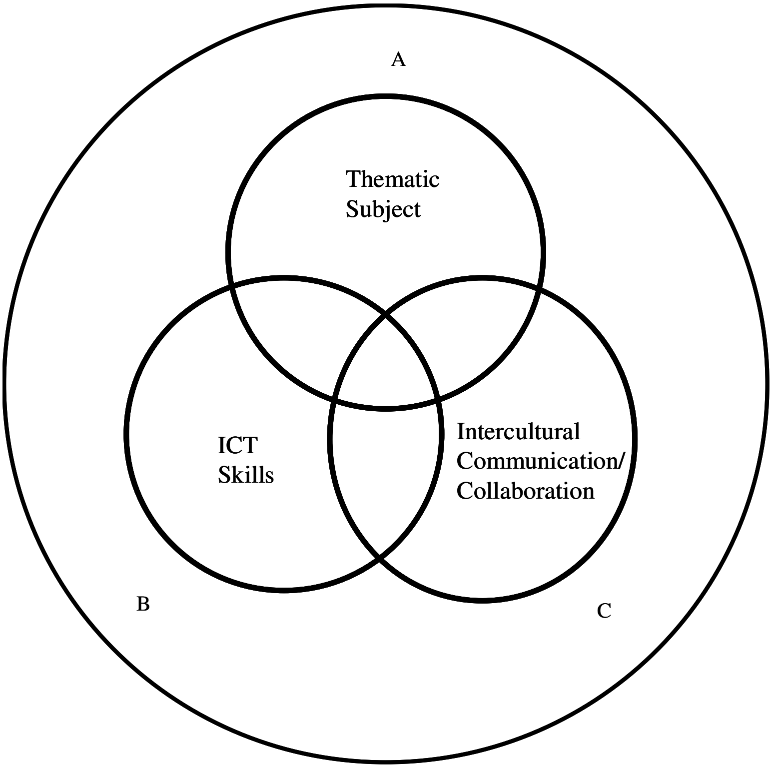


Fig. 1. Concept of on-line intercultural project

project with multiple teachers whose specialties are in different fields. Teachers can ask for some support from schools, such as specialists or NPOs. Collaboration with other agencies will be one of the solutions as well.

3 Case Study

In this chapter, the practice of one senior high school in Hyogo, Japan will be described.

3.1 Overview of the 'My 1.17' Project

Title: My 1.17 ("Natural Disaster Youth Summit" project to commemorate the 10th anniversary of the Great Hanshin-Awaji Earthquake)

Subject: Period of integrated study

Targeted Students: Senior high school year 3, grade 12, Number of students 32
Unit: Disaster Reduction Education

Research Period: September 1, 2004 to March 28 ,2005

Period of Conference: March 25 to 28, 2005
Research Method: Participating observation
Research Question:
Q1: How does the Intercultural disaster reduction education change students' attitude and awareness toward natural disasters and the global problems?

Applicable School Curriculum Guidelines:

(1) Develop students' qualities and skills to identify issues, learn and think about the issues on their own, and self-reliantly make decisions and find better solutions.
(2) Help students acquire how to learn and how to think, develop their mindset to independently and creatively undertake problem-solving and exploring activities, and develop their ability to think about their way of living and being.

Objectives:

(1) Tell the experience of the Great Hanshin-Awaji Earthquake to the world and future generations by reflecting on the students' own experience and expressing it in words to be published on the Web, on the 10th anniversary of the earthquake. Compile the students' earthquake disaster experience in the form of a booklet as the outcome of the activity.
(2) Have the students reflect on the preciousness of life and importance of helping each other, to raise their awareness about disaster reduction.
(3) Utilize ICT such as Internet to self-reliantly collect, send, and share information, and deepen understanding of copyright issues through the activity to produce information content.
(4) Participate in the international project, "Natural Disaster Youth Summit," use the Web and the electronic forum as a tool of inter-cultural communications, and undertake project-based learning together with students overseas.
(5) Become a person who can act independently from the global viewpoint, as a member of the global community.

3.2 General School Curriculum Guidelines

Curriculum by unit (10 hours in total)

(1) Confirm objectives of the lessons…………………………………………1 hr
(2) Talk about your earthquake disaster experience…………………………..1 hr
(3) Read materials prepared by Takarazuka Fire Department to capture the big picture of the damages in your area………………………………………1 hr
(4) Read *Asu ni Ikiru* (Live for Tomorrow) to learn about earthquake experience in other areas……………………………………………………………1 hr

(5) Think of what others can learn from your experience……………………1 hr
(6) Put your earthquake disaster experience in writing………………………1 hr
(7) Prepare Web publication…………………………………………………1 hr
(8) Listen to the reports of students who participated in an affiliated program of The United Nations World Conference on Disaster Reduction to learn about disaster-reduction learning at other schools………………………………1 hr
(9) Learn about earthquake damages in Iran and international rescue effort coordinated using ICT……………………………………………………1 hr
(10) Prepare PowerPoint presentation for Natural Disaster Youth Summit…...1 hr
(11) Give a presentation at Natural Disaster Youth Summit 2005 in Hyogo on the outcome of the project (by students representing the school)

Other: Independently use the after-school time to collect materials, gather information, and submit posts to the electronic forum.

Table 1. Activities, approach and planned activities and the reason for using ICT

Activities	Approach and Planned Activities	Reason for ICT use
Use earthquake disaster experience for the future	Reflect on and write about earthquake experience. Publish the writing on the Web.	Send information
Learn about earthquakes in the world	Introduce earthquake information sites on the Internet. Research on earthquake damages in other countries.	Research information
Exchange information on learning about disasters	Submit posts to the electronic forum for international and cooperative learning and read posts by others.	Exchange views
Exchange with other schools	Exchange the disaster bears and introduce the content of exchange activities on a blog.	Share information
Give presentation on the learning outcome	Prepare the presentation for the "Natural Disaster Youth Summit 2005 in Hyogo"	Prepare presentation materials

3.3 Process of the Project

In the project learning, the teaching material and planned activities are closely linked with students' learning activities. The table 2 clearly shows the relation between teaching material and students learning activities by putting teaching materials and students' learning activities side by side.

Table 2. Teaching material and planned activities with students' learning activities

	Teaching Materials and Planned Activities	Students' Learning Activities
S E P T E M B E R to D E C E M B E R	The class starts in September. Hyogo Prefectural Board of Education's Supplementary Reader for Disaster Reduction Education, *Asu ni Ikiru.* Materials by Takarazuka city Fire Department Iris: Seismic Monitor Introduction of a Website where you can research on the earthquakes in the world http://www.iris.edu/seismon/ Natural Disaster Youth Summit electronic forum for students of the project-participating schools to talk with each other.  **Fig. 3.** Webpage of electronic forum. http://ndys.jearn.jp/2005/ja/index.html Our "Sakura" disaster bears were sent to an exchange counterpart as the representatives of Kawanishi SHS Takarazuka Ryogen Campus. An exchange counterpart in Taiwan sent us the photograph of their activity.	The students asked their ALT to support their activity by proofreading their English publication materials and the ALT gladly agreed to do that.  **Fig. 2.** Students and teachers of the project The students wrote about their earthquake experience. After the class discussion on each one's experience, the students put their experience in writing to let many more people know about it.  **Fig. 4.** Writing stories in the classroom The students sent their posts to the electronic forum and searched the Internet to obtain earthquake information.

Table 2. (*continued*)

<table>
<tr>
<td></td>
<td>


Fig. 5. Partner students in Taiwan

Introducing disaster bears sent from overseas as an exchange

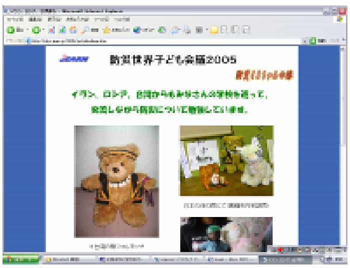


Fig. 7. Webpage of litter ambassadors

The disaster bear “Sakura” blog covered the students’ exchange with others in other countries.

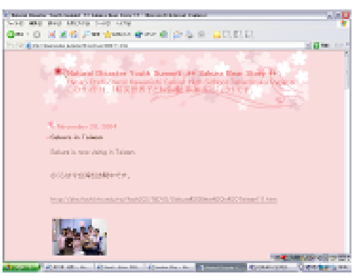


Fig. 9. Blog for communication. http://blog.livedoor.jp/epiker78/
</td>
<td>


Fig. 6. Activities in the computer classroom

The students published their experience on their Website, “My 1.17”

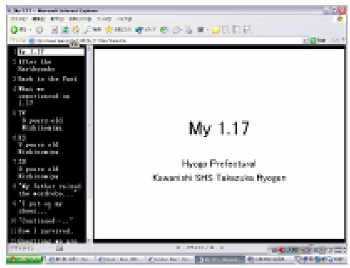


Fig. 8. Webpage of ‘My 1.17’

To reproduce their experience of the earthquake disaster, the students decided to write about their experience from the viewpoint of the children that they were at the time of the earthquake.

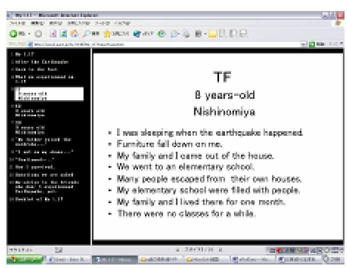


Fig. 10. Webpage of students’ message
</td>
</tr>
</table>

Table 2. (*continued*)

	An exchange counterpart in Iran sent us the photograph of an activity of Iranian counterpart.  **Fig. 11.** Photo from Iran	The students created a Web page to express their condolences and moral support for the people affected by the Niigata-Chuetsu Earthquake.  **Fig. 12.** Student's work
J A N U A R Y 17	"Hopes for revival prevail on earth —Painting, Photo Exhibition and Video Conference" (an affiliated program of The United Nations World Conference on Disaster Reduction) Video conference with Iran and commemorative tree planting The students learn about practices of other schools.  **Fig. 14.** Video conference between Japan and Iran	The students representing the school participated in the commemorative tree planting.  **Fig. 13.** Tree planting activities The students made friends with other students from Japanese schools that participated in the project.  **Fig. 15.** Group photo of tree planting

Table 2. (*continued*)

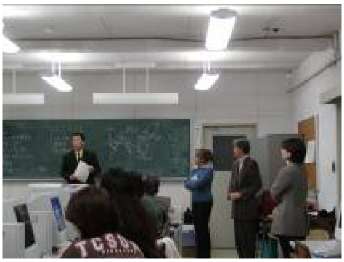
JANUARY to MARCH	Special Lecture at the school to Commemorate the 10th Anniversary of the Great Hanshin-Awaji Earthquake Visiting lecturer from Iran was invited to talk about the earthquake damages in Bam, Iran, and the international rescue effort coordinated by children using ICT in the aftermath.  **Fig. 16.** Special lecture of Bam Earth-quake Making a booklet for introducing the students' experiences of the earthquake.	The students watched the exhibition of paintings by children affected by the earthquake in Bam, listened to the lecturer, and learned about how ICT helped the coordination of the international rescue effort.  **Fig. 17.** Group photo with pictures painted by children in Bam The students printed the PowerPoint presentation compiling their earthquake disaster experience and created a booklet entitled "My 1.17," to be used as teaching materials for the next school year.
MARCH 25 to 28	**"Natural Disaster Youth Summit 2005 in Hyogo"** As the final phase of the project, students got together in Hyogo, Japan and had an international conference.  **Fig. 19.** Poster session of NDYS	Two students representing the school participated in the conference.  **Fig. 18.** Group discussion The students presented their learning outcome in posters at the Poster Session Booth.

Table 2. (*continued*)

<table>
<tr>
<td></td>
<td>


Fig. 20. Poster session of Iranian booth

Oral presentation

Joint presentation with exchange counterparts overseas

Fig. 22. Oral presentation with Russian students

Review by experts on disaster-reduction education

Fig. 24. Review by one of experts

All participating students announced their Declaration at the Natural Disaster Youth Summit 2005.
</td>
<td>
The booth was set up as a poster gallery, where the students answered questions by the visitors, and distributed the booklet on their earthquake experience to make use of their experience in the future.

Fig. 21. Poster session booth presented by students

The students representing the school gave the presentation on the exchange activity involving the Little Ambassador (Stuffed bear).

Fig. 23. Oral presentation by the representative students

The representative students joined and announced the Natural Disaster Youth Summit Declaration together.

Fig. 25. NDYS Declaration
</td>
</tr>
</table>

3.4 Evaluation Planning

Most students experienced the Great Hanshin-Awaji Earthquake in 1995. It was possible that some students remained traumatized by what they had experienced. Therefore, when the author asked them to talk about their experiences, she did not force them to talk, and told them to limit the subjects to what they can comfortably talk about in public. Students' class participation was assessed by a three-level assessment system for each lesson. Table 3. shows the assessment points, method, and tools for the evaluation.

Table 3. Assessment points, method, and tools

Assessment Points	The overall assessment looked at the students' ability to 1) tell their story, 2) listen carefully to stories of others, 3) ask questions, 4) complete the deliverables, 5) conduct research independently using the Internet, and 6) submit their posts to the electronic forum.
Assessment Method	Observation by teacher; assessment of the work submitted; and peer assessment by other students
Assessment Tools	Assessment of the portfolio Assessment of the essay Assessment of statements and questions made during class Assessment of Web content creation

4 Findings

In this section, findings through the six months practice will be discussed. Students change is explained as the result of the project.

Compared with the time before the course started, the students' awareness and attitude about disaster reduction have changed in many ways.

Before this course, most students had never related their earthquake experience to others. Close friends in the same class knew very little or nothing at all about other students' experience of the earthquake. The aim of this project was to publish on the Web their stories of being hit by an earthquake and to inform those who have never had such an experience, so that they could be better prepared. By talking about their own experience, the students came to know that some of them had made a very narrow escape from death, some had been rescued by their families, and some had been rescued by volunteers and were very grateful for it. The students came to realize that all of them had profound feelings about the earthquake. After listening to other students, one student who had in the beginning pretended to have total memory blackout about the earthquake gradually opened up and started relating his experiences over a long period of time as a refugee in a primary-school-turned-shelter. It turned out that all the students in the class had taken shelter at their respective primary school in the immediate aftermath of the earthquake. Through this discussion, the students were made aware of the roles that schools played at the time of a major disaster.

When the teacher/author told the students about a plan to exchange information with a school in Taiwan, some students researched about the major earthquake that hit Taiwan in 1999 on the Internet, without being told to do so. After starting the exchange with students overseas on the electronic forum, our students demonstrated a high level of concern for disasters in other countries that hit people who are strangers to them. Some students started collecting earthquake-related information from the Internet, using their spare time after school hours. Thus, the students published their own experiences on the Web, obtained disaster information from the Internet, and acquired general skills and aptitude to actively utilize ICT.

The school's unique activity of asking the students to reflect on their own experiences and publishing them in such a way that would be useful for future earthquake disasters, in combination with holding exchanges with schools in Japan and other countries, helped the students develop their ability to relate to disasters in other parts of the world.

Before the course started, there was some concern about the psychological impact on the students arising from having to talk about their earthquake experience. However, once the course started, such concern proved to be unnecessary, as the students seemed to enjoy their learning process. This was perhaps because they had classmates and other exchange counterparts who listened to their stories with empathy. Students who had never related their experience to others put it in writing, translated it into English for people overseas to read, and published it on the Web. This may be described as the first sign of their awareness as a member of the global community.

These changes observed in the students were partly achieved by the acquisition of the skills to utilize ICT. However, the bigger contributor seemed to be the learning environment newly available to us, i.e., the electronically networked global community born out of the development of ICT.

Last but not least, the students were able to nurture friendship and trust with other students overseas by participating in the "Natural Disaster Youth Summit" project, in which 23 schools from Japan and 28 schools from 15 other countries participated. Two students represented our school at the "Natural Disaster Youth Summit 2005 in Hyogo" held in March 2005, delivered a presentation on our undertakings, and were engaged in direct exchange with students from overseas.

This project helped develop a trusting relationship between teachers in Japan and other countries as well. A network that can be described as a world educational network for disaster reduction has been formed as one of the many positive outcomes of the project.

5 Conclusion

In this report, practice of one high school is examined. Through the practice at the high school, it became obvious that students' attitude and awareness toward disasters and global problems are changed through the project. It may be said that the potentiality of changing student attitude and conscious through the project is high.

Through the six months observation of students learning, we can conclude that the on-line intercultural project changed the students' view toward their own experience and to the world. However, as the practice was conducted by the author, we can not deny the possibility that the observation may include the subjective bias. Therefore, one case study is not enough to generalize the effectiveness of on-line intercultural project. We need more proofs for the generalization, therefore, the research should be continued.

Acknowledgments. This paper is re-written and revised based on the report handed in to Microsoft Innovative Teachers' conference. This project is conducted with the cooperation of iEARN, JEARN, Microsoft and Hyogo Prefecture. I would like to show my appreciation for supports.

References

1. Abhyankar, M.: Education, Knowledge, Innovation – Building A Culture Of Safety And Resilience. An Indian Experience. In: IDRC DAVOS 2006 2 Extended Abstracts (2006)
2. Anderson, J.D.: The International Education And Resource Network: Building a Sustainable Virtual School, PhD Dissertation, University of Washington (1995)
3. iEARN: iEARN Toppage, accessed (December 1, 2006), URL: http://www.iearn.org/
4. JEARN Natural Disaster Youth Summit Committee: Bosai Sekai Kodomo Kaigi 2005 in Hyogo final report, Hyogo (2005)
5. Natural Disaster Youth Summit Committee: NDYS, accessed (December 1, 2006), URL: http://ndys.jearn.jp/
6. Samovar, L.A., Porter, R.E.: Intercultural Communication: A Reader, Wadsworth (1976)
7. Sawyer, R.K.: The Cambridge Handbook of the Learning Sciences. CUP, Baldonado (2006)
8. Yoshie Naya: 'My 1.17'in Innovative Teachers Network, accessed (December 1, 2006), URL: https://member.innovativeteachers.jp/resource/resource.aspx?id=424/

Intercultural Collaboration: Instrumentality of the GLOBE Study

Mikhail V. Grachev[1] and Mariya A. Bobina[2]

[1] Associate Professor of Management, Western Illinois University, 3561 60th Street, Moline, Il 61265, USA
[2] Ph.D. Candidate, University of Illinois at Chicago, 601 South Morgan Street, Chicago, Il 60607, USA

Abstract. The paper focuses on and further develops the social attributes of intercultural collaboration. It introduces the findings of the Russian component of the GLOBE study (conducted in 62 countries) of manifestation of culture and leadership in business settings. In reflecting on the findings, the authors rely on interviewing managers of leading MNCs about cultural dimensions of Russian business environment, and on reviewing the case of 3M Company that demonstrates tactics to achieve successful intercultural collaboration.

Keywords: Cross-cultural management, multinational companies, intercultural collaboration, emerging markets, Russia.

1 Introduction

The complexity of today's business environment calls for search for non-traditional, less tangible sources of competitive advantage, such as flexible organizational practices, leadership development, and intercultural collaboration. This notion is recognized by the intellectual champions of the business community who take into consideration cultural dimensions in determining their choice of organizational practices in foreign operations.

Cross-cultural problems in society, and in particular, in international business organizations, for decades have been attracting attention of academic scholars. It has been shown in the literature that variations in appropriateness and efficiency of organizational practices in the multinational companies are partly explained by dimensions of societal and corporate cultures. At the societal level political scientists created typologies of civilizations [15] and linked channels for effective collaborations to the structure of social capital [4]. Anthropology and behavioral scholars generalized empirical findings on values, beliefs, and behaviors to the level of society; and compared and clustered national cultural profiles accordingly [9, 11, 22]. They also explained how participants from different cultures or cultural clusters may display visible cultural asymmetries; and set the broad stage for culture distance as predictor for multinational companies' (MNCs') success in foreign direct investment and international operations [1, 20, 21]. At the organizational level, strategy and behavioral scholars explored companies' approach to host-country cultural environments and linked spectrums in corporate behaviors (from "cultural

T. Ishida, S.R. Fussell, and P.T.J.M. Vossen (Eds.): IWIC 2007, LNCS 4568, pp. 382–393, 2007.

relativism" to "ethical imperialism" [3]; or from "ethnocentric" to "polycentric" and "geocentric" [2]) to internal values and strategic predispositions the organization.

The purpose of this paper is (a) to present the value of the cross-cultural project GLOBE (Global Leadership and Organizational Behavior Effectiveness) conducted in 62 countries and (b) to display the instrumentality of these findings in predicting critical issues in intercultural collaboration in MNCs. We present our GLOBE results (as GLOBE County Co-Investigators) based on interviewing 450 managers in three industries and test these results by interviewing managers of nine large multinational companies doing business in Russia and by focusing on intercultural practical solutions in one of these companies. While there has been growing interest to MNCs' intercultural collaboration in Russia [5, 6, 10, 18, 19] the country was not a part of previous large-scale multi-nation cross-cultural studies. Hence comparative GLOBE findings can be of a substantial help for those business organizations that are trying to find out which organizational practices should be used in collaborating with their employees and with local supply chain partners in emerging markets.

2 GLOBE Study and Manifestations of Culture

2.1 GLOBE Project Design

Cross-cultural problems in society, and in particular, in business organizations, for decades have been attracting attention of academic scholars in different disciplines. Recent advanced research conducted by the multidisciplinary GLOBE project team helps to further understand cultural configuration of societies and to predict cultural and behavioral discrepancies when representatives of these societies interact in the organizations [12, 13, 14, 23]. The central GLOBE proposition is that attributes and entities that distinguish a given culture from other cultures are predictive of the practices of organizations of that culture and predictive of the leader attributes and behaviors that are most frequently enacted, acceptable, and effective in that culture. The theoretical base that guides the GLOBE research integrates implicit leadership theory, value/belief theory of culture, implicit motivation theory, structural contingency theory of organizational form and effectiveness, and integrated leadership theory.

Within the GLOBE research the societal and organizational cultures were operationally measured by assessing questionnaire responses from middle managers in three industries (telecommunications, food processing, and financial services) with respect to (1) the values they endorse and (2) reports of practices of entities in their societies. The value questionnaire responses concerned respondents' reports of their values with respect to nine cultural attributes (responses in "Should Be" mode). The entity questionnaire responses concerned the respondents' reports of societal and organizational practices (responses in "As Is" mode).

The GLOBE cultural dimensions were based on previous fundamental works in anthropology and behavior science [11, 16, 17]. Cultural values and practices were measured on a 7-point response scale with respect to nine cultural dimensions that display high within-culture and within-organization agreement and high between-culture and between-organization differentiation: Institutional Collectivism (degree to which organizational and societal norms and practices encourage and reward collective distribution of resources and collective action), Group Collectivism (degree to which individuals express pride, loyalty, and cohesiveness in their organizations or families), Gender Egalitarianism (extent to which an organization or society minimize gender role differences), Assertiveness (degree to which individuals in organizations or societies are assertive, confrontational, and aggressive in social relationships), Power Distance (degree to which members of an organization or society expect and agree that power should be unequally shared), Performance Orientation (extent to which an organization or society encourages or rewards group members for performance involvement and excellence), Future Orientation (degree to which individuals in organizations or society engage in future-oriented behaviors such as planning, investing in the future, and delaying gratification), Uncertainty Avoidance (extent to which members of the organization or society strive to avoid uncertainty by relying on social norms, rituals, and bureaucratic practices to alleviate the unpredictability of future events), and Humane Orientation (degree to which individuals in organizations or societies encourage and reward individuals for being fair, friendly, generous, caring, and kind to others). The GLOBE team reported on rankings of 62 societies with respect to these nine attributes [12].

Several statistical procedures were applied to define the properties of the GLOBE cultural scales. The determination of cultural aggregation was justified by measures that compare the observed variance within a society to the variance expected if there is no within-society agreement. The ICC(1) statistic provided information on the appropriateness of aggregation, comparing the variance between societies with the variance within societies. The reliability of scales was assessed with respect to two random error sources (internal consistency and interrater reliability).

GLOBE project in itself was a great example of successful intercultural collaboration of scholars from different countries. It relied on multiple cross-border interactions of professionals, on overcoming language barriers by translating questionnaires from and back to English language; on discussing and eliminating ambiguities in interpretations of the terms; and on active collaboration in designing shared definitions and concepts.

2.2 GLOBE Project: Russian Data Collection

The GLOBE data collection in Russia was performed by the authors of this paper in 1994-2001 in several stages via pre-pilot study, focused group interviews, main survey, and analysis of media's reports on critical cultural issues.

The main GLOBE data on Russia was collected in 1996-1998 from 450 managers representing the main regions of the country (Central, North-West, Urals, Siberia, Far East). The sample has the following composition. The average age of respondents was 38.8 years. 61.7 percent of the sample were men and 38.3 percent women. Ethnic composition of the sample reflected the percentage of ethnic groups in Russia with the Russians as predominating one. In the GLOBE sample, the share of Russian respondents was 69 percent, Ukrainians – 10 percent, Tatars – 5 percent, other nationalities' percentage was under 2 percent. The average employment history of respondents was 16.8 years, average management experience – 7.4 years, employment in their current organization – 8.6 years. 40 percent were members of professional organizations, and 15 percent were actively involved in trade and industry associations. In the past, 5 percent of respondents worked for multinational corporations. Surveyed managers worked in production and engineering (42 percent), administration (28 percent), sales and marketing (15 percent), human resource management (8 percent), R&D (5 percent), and 2 percent were in planning and other functions. Educational level of respondents was very high – total number of years of education was 15.5. The university/college background of 61 percent was technical and 39 percent in economics, planning and finance.

2.3 Russian Cultural Profile

GLOBE indicators and rankings reflect the realities of economic and social transformation in Russia. Table 1 summarizes Russian scores and all-country average behavioral scores.

Table 1. Russian GLOBE scores and all-country average scores

GLOBE dimension	Russian behavior score ("As Is")	Russian values score ("Should Be")	62 countries behavior ("As Is") average
Institutional Collectivism	4.50	3.89	4.235
Group Collectivism	5.36	5.79	5.118
Gender Egalitarianism	4.07	4.18	3.397
Assertiveness	3.68	2.83	3.858
Power Distance	5.52	2.62	5.147
Performance Orientation	3.39	5.54	4.094
Future Orientation	2.88	5.48	3.837
Uncertainty Avoidance	2.88	5.07	4.158
Humane Orientation	3.94	5.59	4.092

In particular, on the Russian "As Is" scale, the level of Institutional Collectivism is 4.5 displaying traditional group-oriented behavior rooted in historic traditions and Communist indoctrination of collectivist behavior. At the same time, "Should Be" scores place Russia at the end of spectrum, with low collectivism of 3.89, thus indicating the process of fragmentation of society, and break-down of traditional group values. "As Is" score on Group Collectivism is high (5.63) placing Russia high on this dimension. While the "Should Be" results are still high - 5.79, they are relative, as the other countries' respondents consider their "Should Be" values in Group Collectivism at a higher level.

The "As Is" (4.07) and "Should Be" (4.18) scores on Gender Egalitarianism for Russia are quite close, thus displaying respondents' relatively low concern for the difference between values and behaviors on this dimension. Current behaviors could be considered highly egalitarian. However, when compared with the general trend, expressed by respondents in the other countries, Russia is less concerned with egalitarian values.

The "As Is" score on Assertiveness 3.68 and the "Should Be" score 2.83 place Russian managers' behaviors and values low on GLOBE scale. While the current transitional society demands more assertive behavior with tough measures to survive and transform businesses and society at large, the historic heritage of caring for the other people and social assistance limits such assertive behavior for many managers. Also, interpersonal networks (family ties, nepotism, or criminal structures intertwined with business) and collective obligations often suggest forms that are different from open and direct assertive management.

On Power Distance the gap between "As Is" and "Should Be" scores is enormous. The "As Is" behavioral score 5.52 is high. When asked about the values, respondents indicated their preferences for lower power stratification with the "Should Be" score 2.62. Current common belief in democratic reforms may eliminate political power over economic behavior, but this option, to a great extent, depends on Russia's overall ability to balance democracy with establishing order in society.

On Performance Orientation, quite poor performance orientation is reported with the "As Is" score of 3.39. This reflects the heritage of the previous command system with limited encouragement for exceeding the planned targets. However, the "Should Be" score is quite high (5.54). This may be interpreted as competition orientation among respondents.

On Future Orientation the "As Is" score 2.88 is extremely low. It is far below the other countries' average responses, and unmistakably characterizes the nature of the current transformational business climate. However, as in the Performance Orientation case, the Russian respondents believe that economic stabilization is contingent on changes in value orientation. The high "Should Be" score of 5.48 brings Russian to the group of future-oriented countries.

On Uncertainty Avoidance with an "As Is" score of 2.88, Russia has the lowest among all countries on avoidance practices. To a certain extent, this indicates the entrepreneurial and risk-oriented behavior of Russian managers. At the same time, managers' responses to the "Should Be" questions show a large gap between reality, on the one hand, and values and expectations on the other. The "Should Be" score of 5.07 positions Russia among the countries with high uncertainty avoidance preferences. That is the preference for better planning system the Russian managers feel comfortable with. In the current situation, managers quickly and creatively adjust to rapidly changing situations in the environment.

On Humane Orientation the relations between the current behavior and the values of the Russian managers look encouraging for the prospects of the country. While the "As Is" score of 3.94 demonstrated relatively low level of humane orientation in behavior, the "Should Be" score is high - 5.59 that placed Russia among the most humane oriented countries.

Summarizing these findings, our research positioned Russia as having an extreme profile when compared to the other countries on GLOBE dimensions: very low in Uncertainty Avoidance, Future Orientation, Performance Orientation, and Humane Orientation, very high on Power Distance. Some dimensions display more agreement between behaviors ("As Is") and values ("Should Be") - Institutional and In-group Collectivism, Egalitarianism and Assertiveness; while the others present a visible gap between behaviors ("As Is") and values ("Should Be") - Power Distance, Performance Orientation, Future Orientation, Uncertainty Avoidance, Humane Orientation.

In particular, in behavioral set of findings, the extreme low Uncertainty Avoidance score and rank could be considered favorable for entrepreneurship activities unless one links it to very low Future Orientation. That can be interpreted as a lack of vision in management and entrepreneurship, as primary focus on survival and short-term business development. Low Performance Orientation makes it difficult to encourage managers to focus on continuous improvement and learning. Low ranking on Humane Orientation raises doubts about long-term investments in human resources. High Power Distance scores explain the tough bureaucratic measures in crisis management and in restructuring enterprises and industries.

There is a wide gap between “As Is” and “Should Be” data on the dimensions linked directly to economic reforms. The “Should Be” model displays the deficit in and the preference for a more humanistic, ethical, democratic and values. At the same time, there is no serious gap on dimensions which are strongly linked to historical cultural roots, such as In-group Collectivism. The Gender Egalitarianism “As Is” score is positive and is not much different from “Should Be”. But when compared to the overall democratic trend among the other countries, a relative decline in egalitarian values may create potential problem for Russia in the future.

3 MNCs’ Response to the Host-Country Culture

3.1 Survey of MNCs’ Managers on GLOBE Data

GLOBE study has found some distinctive positive features of the Russian management culture that can be considered as sources of optimism for international companies considering growth strategy in Russia. Several deeply rooted cultural characteristics – a willingness to work cooperatively, a downplaying of traditional gender roles in the workplace translate into potential benefits for international businesses. In addition the Russian managers tend to be educated, cosmopolitan in outlook, and comfortable reacting to the rapid changes of an unstable environment.

In 2002 to test the behavioral profile of Russian management and to better interpret, examine, and enrich the GLOBE results presented herein the authors have surveyed a group of top managers of the leading multinational companies that were successful in doing business in Russia [6]. In particular, they focused on collaboration between Western and Russian managers and on the willingness of these companies to incorporate cultural components into policy and decision-making process. While the authors understood the limitations of this follow-up study, such as focus on perceptions or its qualitative nature, they extracted valuable data on the channels for improving interaction with employees internally and with local supply chain externally. The other visible differentiator of this survey was the authors’ interest in policy issues beyond traditional discussion on correcting cross-cultural mistakes – to capitalizing on cultural differences.

Managers surveyed on Russia’s GLOBE profile represented the following companies: Shell, Unilever, 3M Company, J.P. Morgan, Motorola, Golden Telecom, KPMG, Hines International, and Sun Intrebrew. In the sample of 9 top managers, 3 managers were expatriates, and the rest were Russian citizens working for their companies for more than 3 years. Out of 9 respondents 6 were female managers. The average age of surveyed managers was 39.

On each GLOBE dimension, respondents were asked two sets of questions. First set of questions asked if the Russian score on selected dimensions had provided strategic advantage or strategic disadvantage to their company. The second set of questions was linked with the first one: if it provided an advantage, did the company capitalize on this or not? If it provided a disadvantage, did the company try to correct this or not?

3.2 Cultural Advantages and Disadvantages of the Russian Profile

The authors had identified areas of high agreement among MNCs' managers on some of the GLOBE-generated cultural dimensions.

The analysis displayed agreement on the *advantages* of the Russian behavioral configuration. One of such areas was high scores on Collectivism. The company respondents positively assessed the ability of Russian managers and employees to work in teams and to follow group norms. Each company in the sample has developed a sophisticated system to exploit this factor and motivate the high loyalty of its Russian managers, sometimes pushing them to sacrifice individual interests. Environment with high Family Collectivism helps multinationals to combine innovative efforts within the company and target specific groups in the market. However, some respondents indicated an unusually high influence of trendsetters within collectivist environment.

The respondents shared high level of agreement on selected cultural *disadvantages*. One such area was low score on Uncertainty Avoidance. The companies tried to correct the influence of this factor by providing clear corporate guidelines and codes of conduct. Low score on Assertiveness was also considered significant. The respondents explained this score referring to conformism rooted in the Soviet system and lack of leadership initiative. To balance the negative impact of this factor, the companies designed specific programs encouraging initiative and entrepreneurial assertive behavior for managers.

Most respondents mentioned the companies' serious consideration of low Performance Orientation and Future Orientation and impact on company policies. These scores reflected the heritage of the previous command system and painful realities of economic transition. They were deeply cemented in the current society and difficult to change in the short period of time. However, the responses whether the situation with low Performance Orientation and low Future Orientation provided advantage or disadvantage for the company were mixed. Some respondents viewed this as a positive factor since companies could build competitive advantage by creating vision and encouraging leadership behavior. The others considered this a disadvantage when working with more future-oriented headquarters and other international subsidiaries.

4 "Role Model" in Intercultural Collaboration

4.1 Focus on 3M Company

The U.S.-based multinational company 3M participated in the abovementioned survey. It was chosen for our in-depth study of the "role model" in intercultural collaboration in the Russian business environment. 3M is known for its careful examination of local conditions and the ability to transfer organizational policies and practices across the national borders. It successfully handles cross-cultural interaction internally (among employees with different ethnic backgrounds) and externally (with local partners, suppliers, and stakeholder groups).

3M Company has had a long history of business relations with Russia, with true expansion into the Russian market within 1990s' economic reforms. For the past decade, the company has enjoyed double-digit sales growth in Russia. The company has helped launch more than a hundred new local businesses by sharing its technology with them and converting 3M products to the Russian market. These startups served as a nationwide network of manufacturing facilities for the company as well as dealers, distributors and partners who shared the 3M values of reliability and quality of products and services.

The authors collected data on the company, visited corporate head office, industrial facilities and offices in four countries, surveyed 18 corporate managers in 8 countries beyond Russia, and interviewed 10 managers in 3M's Russian subsidiary [5, 7, 8].

In the process of extracting and analyzing data, the authors perceived 3M as the global company, balancing home-based component with local sensitivity. However, the authors were also aware of the U.S.-based source of key corporate values, norms, and behaviors, and considered Russian-American cultural distance in the interview process. In particular, GLOBE findings suggested that in Russia Performance Orientation was much lower than in the U.S. American managers strived to avoid uncertainty while the Russian managers widely accepted uncertainty. The American managers engaged in future-oriented behaviors such while the Russian managers were scored extremely low on Future Orientation. American managers were more assertive in social relationships. And compared to behaviors in more democratic American society the Russians performed with much higher degree of Power Distance.

However, GLOBE results displayed Russian organizational and societal norms and practices encourage collective action at a higher extent than in the U.S. The Russian managers could express higher pride in and loyalty to their organizational. Also in Russia the organizations or society still minimized gender role differences.

4.2 Intercultural Factors in Company Policies

The process of achieving compatibility between current Russian business environments and 3M corporate culture started through careful recruitment and selection procedures. At the start-up phase 3M Company did not insist on sending Americans to be in charge. Nationals of Finland, Germany, and Belgium served as country managers. 3M believed that European managers tend to be more familiar with Russian culture and business norms. This reflected a network philosophy of 3M, which values local expertise over corporate control. European nationals were also 'cheaper' for the company than the American expatriate managers, due to relocation costs and cross-country differences in personal income taxation systems. In the end the company accepted the professionalism of locals and their ability to take responsibility for decisions and implement them effectively. Only one expatriate currently works at 3M's Russian subsidiary.

In Russia hiring procedures were much stricter here than in most other countries of operation. On average, a local job applicant had to go through six to eight interviews with human resources specialists, future peers, and line managers. This complex procedure removed not only those candidates whose professional abilities were not sufficient for the company, but also those, who demonstrated ambiguous ethical perceptions. It also allowed testing whether an applicant's values and expectation did fit into a set of values, proclaimed by 3M Company. The mandatory probation period of three months served the same goals.

Russian job applicants competing for the job at 3M, demonstrated high level of education and adequate English language skills. However, a lot of training was required to develop such deficient skills as an ability to make autonomous decisions, communication skills, and time management skills. A lot of time was also devoted to get the new employees acquainted with the ways things work at 3M, which required cross-cultural training and classes in history, strategy and culture of the corporation.

3M's employees were motivated by a wide variety of methods, which included pay-for-performance, medical insurance, meal plans, housing allowances, gain sharing, etc. Many benefits offered by 3M referred to country-specific needs of the Russian employees. The withdrawal of the state from the market and economic reforms carried out in a monetarist way, led to the full collapse of the previous social infrastructure in Russia. This situation forced international companies to provide their employees with generous benefit packages. 3M pioneered some of these practices. For example, it was the first American company in Russia to offer its best employees financial assistance in purchasing apartments. The company also took the leadership role in offering attractive pension schemes for its employees.

4.3 Facilitating Cultural Sensitivity

3M policies considered such characteristics of the Russian organizational culture as Institutional and Group collectivism and employees' ability to sacrifice personal interests for corporate success.

Managers practiced multi-functional team building, re-assessment of job functions, job redesign, and devised gain-sharing plans aimed at encouraging innovations and cost-reduction. They allowed employees to spend some working time to pursue their own innovative projects. The development of formal and informal networks within organizations was important to an effective innovation process. Networks were believed to encourage knowledge sharing and help team building.

3M defined a set of key values, shared by company employees. These norms were conveyed to the Russian employees through circulation of written materials, orientation seminars, hot lines, and other practices. The other important vehicle for promoting corporate values was detailed and clearly defined corporate guide of business conduct (such guides are still a rare practice in Russia). Serious attention was paid to eliminating cheating, lying, and stealing in - or on behalf of - the company. Among the reasons for employees being fired were: overstated use of expenses on corporate business trips, or extensive use or misuse of a company car for private purposes (type of behaviors quite typical in post-Communist countries).

3M was emphasizing leadership competencies and was actively looking for the new ways of cooperation and collaboration, mentoring through the internal network.

When compared to the Western leadership practices that focused on teamwork and communication, 3M in Russia, however, underlined other priorities such as visibility, status, and personal responsibility of leaders.

The company understood that Russian societal culture was not homogeneous, rather a kaleidoscope of technocratic, predatory, and socially responsible business cultures. 3M focused on the last one and strongly protected its ethics. The company went to clients and taught them business ethics at no charge. The company had set "open schools" for customers, and its own sales representatives were trained to teach ethics courses on site. As a consequence, more and more Russian companies approached 3M Russia for advanced ethical know-how. An important part of company policy was telling its sales people that cheating and bribes shoot back with personal insecurity, making such people the target for criminal structures. So employees' personal safety really concerned the human resource management department.

Russia was known for its charity traditions. Interaction with the local community combined with charity initiatives was highly important for the long-term strategy of 3M presence in the country as corporate citizen. The examples of such a behavior were: distributing medical products to leading Moscow hospitals, participating in the program of Ministry of Education providing schoolchildren with reflective safety labels, donating respirators to workers constructing the Cathedral in Moscow, assisting safety programs in traffic control on Moscow belt highway, bringing the leading corporate inventors (such as inventor of "post-it' notes Art Fry) to local high schools.

To sum up. 3M not only incorporated intercultural component into collaborative practices in Russia, but also positioned itself in this emerging economy as the leader in setting positive examples and promoting culture-sensitive and ethical behaviors.

5 Conclusions

This paper shed light on the further integration of cross-cultural academic findings and their practical implications. GLOBE project was not only the comprehensive global study of cultures and organizations, but also one of the first attempts to collect research-oriented data set on Russia which, (a) was created by using internationally recognized and reliable research methods and (b) was ready for future cross-cultural comparative research among 62 nations within a common scholarly framework. Three basic lessons should be underlined.

First, the gap in values and behaviors related to universal cultural dimensions may be critical in interactions between the Western managers and those from emerging economies. It may also generate obstacles to developing business collaborative practices – within the international company and with outside stakeholders. Cultural reasons may put constraints on partnerships with locals and support international companies' preference to wholly-owned start-ups compared to acquisitions of and joint ventures with the local (in our project – Russian) businesses.

Second, the leading multinationals try to better understand how culture moderates relationships between organizational processes and organizational effectiveness, and frame their strategies by taking into consideration positive component of the country's cultural configuration and benefiting from it. And the leading multinationals may serve as “the role model” for the local businesses by transferring management know-how, as well as internationally recognized practices in intercultural collaboration.

Third, the authors underline the need to further integrating major streams of research and interpretation of corporate tactics of intercultural collaboration grounded in measurements such as the GLOBE study.

References

1. Bobina, M., Grachev, M.: “Passport of Alliance”: from Strategic to Cultural Fit. In: Trick, M. (ed.) Global Corporate Evolution: Looking Inward or Looking Outward? International Management Series, vol. 4, pp. 261–268. Carnegie Mellon University Press, Pittsburg (2004)
2. Chakravarthy B., Perlmutter H.: Strategic Planning for a Global Business. Columbia J. World Bus 3–10 (Summer 1985)
3. Donaldson Th.: Values in Tension: Ethics Away from Home. Harvard Bus. Rev. pp. 48–56 (September-October 1996)
4. Fukuyama, F.: Trust. The Social Virtues and the Creation of Prosperity. Free Press, New York (1995)
5. Grachev, M.: Making the Most of Cultural Differences. Harvard Bus. Rev. pp. 28–30 (October 2001)
6. Grachev, M.: Cultural Attributes of Russian Management. In: Cheng, J., Hitt, M. (eds.) Managing Multinationals in a Knowledge Economy: Economics, Culture, and Human resources. Advances in International Management, vol. 15, pp. 159–177. Elsevier, Amsterdam (2004)
7. Grachev, M., Rogovsky, N., Bobina, M.: 3M: Role Model for Emerging Markets. Thunderbird Intl Bus. Rev. pp. 803–821(2006)
8. Grachev, M., Bobina, M.: Corporate Strategies in Transitional Economies: Capitalizing on Cultural Differences. In: Trick, M. (ed.) Growing the International Firm: Success in Mergers, Acquisitions, Networks, and Alliances, pp. 119–128. Carnegie Bosch Institute, Pittsburg (2002)
9. Hampden-Turner, C., Trompenaars, F.: Building Cross-cultural Competence: How to Create Wealth from Conflicting Values. Yale University Press, New Haven (2000)
10. Hitt, M., Ahlstrom, D., Dacin, T., Levitas, E., Svobodina, L.: The Institutional Effects on Strategic Alliance Partner Selection in Transitional Economies: China vs. Russia. Org. Sci. 15(2), 173–185 (2004)
11. Hofstede, G.: Cultures Consequences. International Differences in Work-related Values. Sage Publications, Newbury Park (1984)
12. House, R.J., Hanges, P.J., Javidan, M., Dorfman, P.W., Gupta, V. (eds.): Culture, Leadership, and Organizations: The GLOBE Study of 62 Societies. Sage Publications, Thousand Oaks (2004)
13. Chhokar, J., Brodbeck, F., House, R. (eds.): Culture and Leadership Across the World. The GLOBE Book of In-depth Studies of 25 Societies. Lawrence Erlbaum: Mahwah NJ (2007)
14. House, R., Hanges, P., Ruiz-Quintanilla, S., Dorfman, P., Javidan, M., Dickson, M., Gupta, V., Grachev, M., et al.: Cultural Influences on Leadership and Organizations: Project GLOBE. In: Mobley, W., Gessner, M., Arnold, V. (eds.) Advances in global leadership, pp. 171–233. JAI Press, Stamford, CT (1999)

15. Huntington, S.: The clash of civilizations? Foreign Affairs 72(3), 22–49 (1993)
16. Kluckhohn, F., Strodtbeck, F.: Variations in value orientation. HarperCollins, New York (1961)
17. McClelland, D.: Human motivation. Scott, Foresman Glenview (1985)
18. Michailova, S.: Contrasts in Culture: Russian and Western Perspectives on Organizational Change. Acad. Management Executive 14(4), 99–112 (2000)
19. Naumov, A., Puffer, S.: Measuring Russian Culture Using Hofstede's Dimensions, Applied Psych. An Intl. Rev. 49(4), 709–718 (2000)
20. Rogovsky, N., Schuler, R.: Managing Human Resources Across Cultures. Bus. and Contemporary World 1X(1), 63–75 (1997)
21. Shenkar, O.: The role of cultural distance in international expansion. In: Trick, M. (ed.) Global Corporate Evolution: Looking Inward or Looking Outward? International Management Series, vol. 4, pp. 363–371. Carnegie Mellon University Press, Pittsburg (2004)
22. Triandis, H.: Individualism and Collectivism. Westview Press, Boulder (1995)
23. Waldman, D., de Luque, M.S., Washburn, N., House, R., Bobina, M., Grachev, M., et al.: Cultural and Leadership Predictors of Corporate Social Responsibility Values of Top Management: A GLOBE Study of 15 countries. J. Intl. Bus, Studies 37, 823–837 (2006)

Author Index