Application of Classical Psychological Theory to Terminological Ontology Alignment

Fumiko Kano Glückstad

Copenhagen Business School
Dept. of International Language Studies and Computational Linguistics
Dalgas Have 15, DK-2000, Frederiksberg, Denmark
fkg.isv@cbs.dk

Abstract. Terminological Ontology [1] is a method used for knowledge sharing and domain-specific translation practices, and could potentially be suitable to apply and for simulating the cognitive theories/models explaining real-world inter-cultural communication scenarios. In this paper, we investigate - as a preliminary study - whether Tversky´s contrast model [2] is applicable to data-sets obtained from the Terminological Ontology method. The eventual purpose of this study is to propose an approach for identifying potential translation candidates by optimizing relevance between concepts in two remote languages such as European and Asian languages.

Keywords: Terminology, Ontology, Translation, Inter-Cultural Communication, Set-theory, Similarity, Generalization, Relevance

1 Introduction

The role of ontology in a multilingual setting is considered an emerging challenge to Semantic Web development. As a consequence, there are several major ongoing projects, such as the MONNET project on Multilingual Ontologies for Networked Knowledge [3] and the KYOTO project on Knowledge-Yielding Ontologies for Transition-Based Organization [4]. Though both projects deal with translation of terms from a Source Language (SL) to a Target Language (TL), they focus on linking lexical data through an interoperable common ontology rather than on optimizing relevance between concepts that are potentially measurable based on diverse models derived from cognitive theory.

Terminological Ontology (TO) is a domain-specific ontology used for knowledge sharing [1], which normally is applied in terminology work, cf. for example [5]. The unique characteristics of TO that differentiate it from other types of ontologies are feature specifications and subdivision criteria [6]. A feature specification consists of a feature dimension and its value. Hence, a representation of a whole concept is a feature structure, i.e. a set of feature specifications corresponding to the unique set of characteristics that constitutes that particular concept [1][6]. Terminological ontologists argue that concepts are defined in a language-dependent context, and therefore TO is language-dependent. TO is developed within a knowledge sharing community, then dynamically updated and validated. If it is necessary to share knowledge with other communities, TOs developed in different communities should be compared, aligned and merged upon necessity. While the aforementioned two mainstream projects, MONNET and KYOTO, both deal with complex ontologies involving huge data-sets, TO usually handles smaller amounts of concepts.

Considering this, a point that should be emphasized in this work is that TO could potentially be a suitable tool to apply and for simulating cognitive theories/models explaining a real-world inter-cultural communication scenario. Thus, in the next chapter a problem existing in the real-world is explained by use of the Relevance Theory of Communication [7]. Next, the concept of similarity, which has a long history in psychological theory, is reviewed in Chapter 3. We consider how the theory proposed by Tversky [2] is applied in the context of the Relevance Theory of Communication. In Chapter 4, the empirical analysis is performed to assess the potential of applying the models based on [2] to TOs. Chapter 5 discusses findings and future work followed by conclusions in Chapter 6.

2 Real-World Problem

Imagine a situation where a non-native English speaking European and an Asian are debating in English about the issue of academic degree systems in their respective cultures. While a German might be explaining about the Doctor of Science (Habilitation) degree (the highest achievable academic degree in Germany after obtaining a Doctor of Philosophy degree), a Japanese might be having the highest possible academic degree in Japan in his mind which is a Doctor of Philosophy degree (also frequently referred to as Doctor of Science in Japan). This imagined conversation shows a typical scenery revealing a deep inherent misconception between the two communicating parties since each of them have their own conceptual - and correct - understanding of the highest obtainable academic degree in their respective cultures.

This example may further create problems for a translator who is going to translate academic titles into the language of the other party. When a translator translates the term of the German Doctor of Science Degree into Japanese, the first condition he/she has to fulfill is that his/her translation should convey the same meaning as the original German meaning. Gutt [8] explains that this requires *the receptors to familiarize themselves with the context envisaged for the original text.* Now the question is, when a Japanese receptor is not familiar with the German language and its academic culture, how should this particular German academic title then be translated into Japanese?

The proposal [8] of applying the Relevance Theory of Communication [7] is a key to address this issue. This theory focuses on how people share thoughts with one another and views communication as principally an inferential process. It means that the essential task of the communicator is to produce a *stimulus* from which an audience can infer what set of thoughts or assumptions the communicator intends to convey [8]. Hence the second condition the translator has to fulfill is that his/her translation should explicitly provide a set of assumptions that are adequately relevant to the audience. The issue here is how the translator should create the *stimulus* (that is *translation*) optimally relevant to the audience. Assuming that both German and Japanese have their respective conceptual structures of the academic system rooted in their own cultures, translation candidates that have optimally relevant relationships identified from these two conceptual systems could avoid the gratuitous inferential processing effort on the audience´s part.

The optimization of the relevance between two concepts could be well explained by the cognitive theory, Tversky´s Set-theoretic model of similarity [2]. Thus the next Chapter reviews Tversky´s model and considers how this model could be used in the context of optimizing the relevance of communication.

3 Tversky´s Set-theoretic Model of Similarity

The concept of similarity has a long history within the area of psychological theory. Tversky´s view of similarity [2] is distinguished from the traditional theoretical analysis (c.f. [9]) on two points: 1) while the theoretical analysis of similarity relations has been dominated by the continuous metric space models, [2] argues that *the assessment of similarity between objects may be better described as a comparison of features rather than as the computation of metric distance between points;* and 2) although *similarity has been viewed by both philosophers and psychologists as a prime example of a symmetric relation*, the asymmetric similarity relation has been demonstrated in [2] based on several empirical evidences.

Based on these two points, [2] proposed a classic feature-set model of similarity as follows:

$$sim(a,b) = \frac{f(A \cap B)}{f(A \cap B) + \alpha \cdot f(A - B) + \beta \cdot f(B - A)} \quad (1)$$

Here, A and B are the feature sets of object *a* and object *b*. *f* denotes a measure over the feature sets. (A⋂B) represents the sets of features present in both A and B, (A-B) represents the sets of features present in A but not in B, and (B-A) represents the sets of features present in B but not in A. α and β are free parameters representing an asymmetric relationship between A and B. Since the similarity score in this equation is normalized, the obtained score lies between 0 and 1.

An interesting point is that the application of Tversky´s model requires *a limited list of relevant features and the representation of an object as a collection of features that is viewed as a product of a prior process of extraction and compilation* [2]. In fact, the principle of TO in a way follows rigid rules of categorization. This can systematically extract the collection of features based on the subdividing dimensions. Therefore, the hypothesis is that Tversky´s model is applicable to data-sets extracted from the terminological ontologies. Another important point in the context of the Relevance Theory of Communication is that translation should provide the set of assumptions that are adequately relevant to the audience, and the stimulus (that is translation) produced by the translator is such that it avoids gratuitous inferential processing effort on the audience´s part. Considering that *similarity serves as an organizing principle by which individuals classify objects, form concepts, and make generalizations* [2], the most similar concept to a SL concept, which is identified in the audience´s culture through the feature matching, could be the set of assumptions which are adequately relevant to the audience. Thus, the second hypothesis is that the optimization of the relevance required in an inter-cultural communication can be achieved by aligning the ontological graphs (conceptual hierarchies) and feature specifications which constitute concepts in the two language-dependent terminological ontologies. In order to assess these hypotheses, terminological ontologies are developed from corpora describing real-world concepts in the two remote cultures. The similarity score of the selected concepts are computed by applying Tversky´s model [2] based on the collection of features extracted from these ontologies. This is dealt with in the next chapter.

4 Feature Matching Based on Tversky´s Model

4.1 Corpora

Texts describing the Japanese educational system have been identified from the "Multilingual Living Information[1]" site provided by the Council of Local Authorities for International Relations and from a pamphlet entitled "Higher Education in Japan[2]" published by the Japanese Ministry of Education, Culture, Sports, Science and Technology. For the Danish educational system, documents that are downloaded from the Eurydice web-site[3] published by the Education, Audiovisual and Culture Executive Agency under the EU commission have been used as text corpus. All these documents are officially published in English by reliable authorities of each country. Thus, all English translated terms and expressions in their original languages are considered as official terms. It means that it is feasible to identify terminological expressions in an original language from documents published by the respective authorities. This enables one to eventually identify translation equivalences linking between, in this case, Danish and Japanese. In this study, only the English documents describing language-dependent concepts in the two cultures are used as text corpora.

The Eurydice publishes documents describing the educational systems in a majority of the EU member countries both in English and in their native languages. It means that the same method can, in principle, be applied to other language combinations.

4.2 Ontology Construction

The terms and their definitions describing the educational systems in each country are manually identified from the respective English documents. Based on these terms and their definitions, terminological ontologies representing the educational system in each

1 http://www.clair.or.jp/tagengorev/en/j/index.html

2 http://www.mext.go.jp/english/koutou/detail/1287370.htm

3 http://eacea.ec.europa.eu/education/eurydice/index_en.php

of the two countries are developed using a Computer Aided Ontology Structuring prototype (CAOS) that is based on the TO principles defined in [1]. As described in Chapter 1, the uniqueness of TO is feature specifications and subdivision criteria [6]. A feature specification is presented as attribute-value pair - for example in Figure 1, [ENTRANCE REQUIREMENT: high school graduate]. Thus, a representation of a whole concept is a feature structure, i.e. a set of feature specifications corresponding to the unique set of characteristics that constitutes that particular concept [1]. In Figure 1, each box that represents a particular concept is divided into three layers: 1) top layer, lexical representation (term), 2) middle layer, dimension specifications, and 3) bottom layer, feature structure (set of feature specifications).

The use of feature specifications is subject to principles and constraints described in detail in [1]. Most importantly, a concept automatically inherits all feature specifications of its superordinate concepts. Secondly, polyhierarchy is allowed so that one concept may be related to two or more superordinate concepts. On the other hand, subdivision criteria that have been used for many years in terminology work are strictly implemented in TO by introducing dimensions and dimension specifications [1][6]. This enables the CAOS prototype to perform consistency checking which helps in constructing ontologies [1]. A dimension of a concept is an attribute occurring in a non-inherited feature specification of one or more of its subordinate concepts [1][6]. Values of the dimension allow a distinction among sub-concepts of the concept in question. In Figure1, the concept “academic degree” has the dimension [LENGTH OF EDUCATION] whose values are [2-3 years | minimum 4 years]. These dimension values distinguish the sub-concepts: “junior college degree” and “university degree”. This clarification makes it much easier to identify subdivision criteria and differentiating characteristics [6]. The same feature attribute can only occur on sister concepts and a given value can only appear on one of these sister concepts. In this way a concept must be distinguished from each of its nearest superordinate concepts as well as from each of its sister concepts by at least one feature specification [1][6].

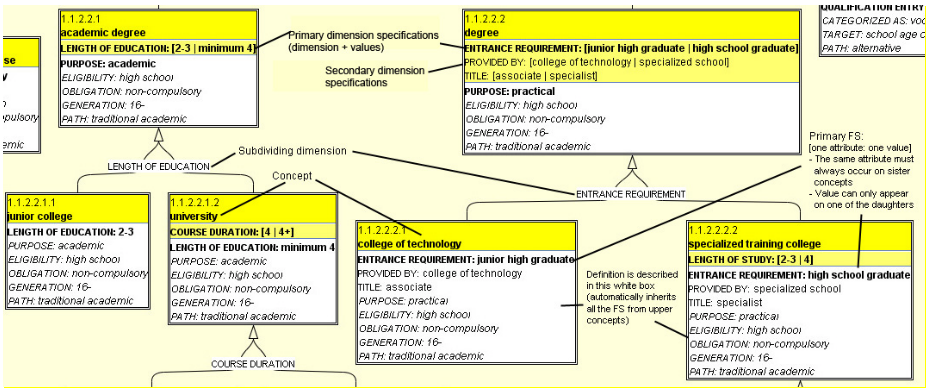

Fig. 1 Example of the Terminological Ontology.

By using the CAOS prototype that performs the consistency checking of the TO principles, the two educational ontologies are developed based on the terms and definitions manually extracted from the corpora.

4.3 Feature Matching Based on Tversky´s Contrast Model

The basic techniques generally used in the ontology matching are string-based (lexical) matching, graph-based (structural) matching, and feature-based matching [10]. Accordingly, string-based matching is manually performed as the first step. The Japanese and Danish educational system ontologies, respectively, contain 42 and 65 concepts consisting of terms (lexical representations) and their feature specifications. Among these, only two terms are completely matched. It indicates that graph-based matching based on the lexically matched nodes is not sufficient in this case. Therefore, feature-based matching is manually implemented in a top-down manner as the second step. The third highest dimension in the Japanese ontology and the highest dimension in the Danish ontology are "generation". Hence, the two ontologies are categorized into the following three blocks based on feature values of this generation dimension: "0-6 years old"; "6-15 years old"; and "16 years old and above."

Table 1. List of terms and feature sets.

Japanese concepts	
a:high school	A= {formal education, 16 years +, non-compulsory, lower secondary graduate}
b:general course	B= {formal education, 16 years +, non-compulsory, lower secondary graduate, general}
c:specialized training course	C= {formal education, 16 years +, non-compulsory, lower secondary graduate, specialized}
d:technical course	D= {formal education, 16 years +, non-compulsory, lower secondary graduate, specialized, technical}
e:business course	E= {formal education, 16 years +, non-compulsory, lower secondary graduate, specialized, business}
f: higher education	F= {formal education, 16 years +, non-compulsory, secondary graduate}
Danish concepts	
g:upper secondary education	G= {16 years +, lower secondary graduate}
h:general upper secondary education	H= {16 years +, lower secondary graduate, access to higher education}
i:gymnasium	I= {funded by state, 16 years +, lower secondary graduate, access to higher education}
j:business college	J= {self governing, 16 years +, lower secondary graduate, access to higher education}
k:HHX program	K= {self governing, 16 years +, lower secondary graduate, access to higher education, business}
l:HTX program	I= {self governing, 16 years +, lower secondary graduate, access to higher education, technical}
m:vocational or technical education	M={16 years +, lower secondary graduate, access to labor market}
n: tertiary education	N={secondary graduate, project and research}

In this study, the focus is on the block having feature value "16 years old and above." From this block, some of the sub-concepts and their feature values listed under the Japanese term, "non-compulsory education", and the Danish terms, "upper secondary education" and "tertiary education" are manually selected in Table 1 (due to the paper space, redundant data – e.g. concepts having similar constitution of feature sets – has intentionally been omitted). In order to apply Tversky´s model, synonymous feature expressions identified from the country specific corpora are approximately standardized by hand in Table 1. One thing to notice in Table 1 is that if a concept is categorized into several sub-concepts based on a dimension, an extra feature specification is added to each of them according to the principles of TO. Hence it is possible to observe the hierarchical structure from the feature values listed in Table 1.

Now the question is how to assign the asymmetric parameters in accordance to the translation direction. In [2], the direction of asymmetry is determined by the relative salience of the stimuli, in other words, the variant is more similar to the prototype than vice versa. Thus, *if sim(a,b) is interpreted as the degree to which a is similar to b, then a is the subject of the comparison and b is the referent. Hence the features of the subject are weighted more heavily than the features of the referent*. When considering a translation scenario, translators´ task is to identify a concept in audiences´ conceptual structure that is optimally relevant to the concept in the SL. It means that the stimulus selected by a translator should to the maximum extent be similar to a concept in the SL concept. Therefore, the features of a stimulus should be weighted more heavily than the ones of an SL concept in accordance to [2]. Hence, the asymmetric parameters are manually set as α=0.7 and β=0.3 in this empirical study. The result is shown in Table 2 and 3. In Table 2, the Danish concepts (*g*-*n*) are set as subject of the comparison and the Japanese (*a*-*f*) as referent. Opposite to this, the Japanese concepts (*a*-*f*) are set as subject of the comparison and the Danish (*g*-*n*) as referent in Table 3.

Table 2. Tversky´s similarity score: *a*-*f* (JP) as referent (SL), *g*-*n* (DK) as stimulus (TL).

SL	*g*	*h*	*i*	*j*	*k*	*l*	*m*	*n*
a	**.77**	.61	.50	.50	.43	.43	.61	.00
b	**.69**	.56	.47	.47	.4	.4	.56	.00
c	**.69**	.56	.47	.47	.4	.4	.56	.00
d	**.63**	.51	.43	.43	.38	.57	.51	.00
e	**.63**	.51	.43	.43	.57	.38	.51	.00
f	**.38**	.30	.25	.25	.21	.21	.30	**.38**

Table 3. Tversky´s similarity score: *g*-*n* (DK) as referent (SL), *a*-*f* (JP) as stimulus (TL).

SL	*a*	*b*	*c*	*d*	*e*	*f*
g	**.58**	.49	.49	.41	.41	.29
h	**.54**	.45	.45	.39	.39	.27
i	**.50**	.43	.43	.37	.37	.25
j	**.50**	.43	.43	.37	.37	.25
k	.47	.4	.4	.35	**.53**	.23
l	.47	.4	.4	**.53**	.35	.23
m	**.54**	.45	.45	.39	.39	.27
n	.00	.00	.00	.00	.00	**.29**

Note1: Bold font is the highest score in each row

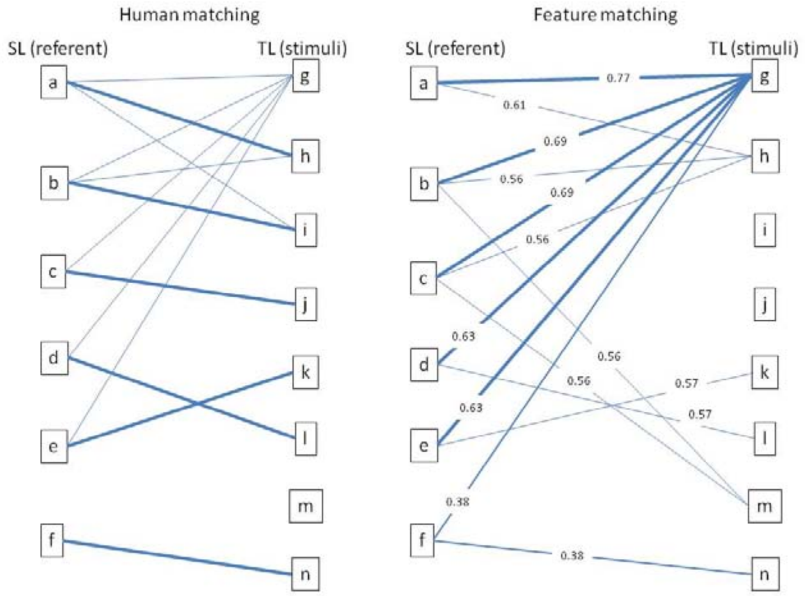

Fig. 2 Comparison with human matching: Danish terms as stimuli

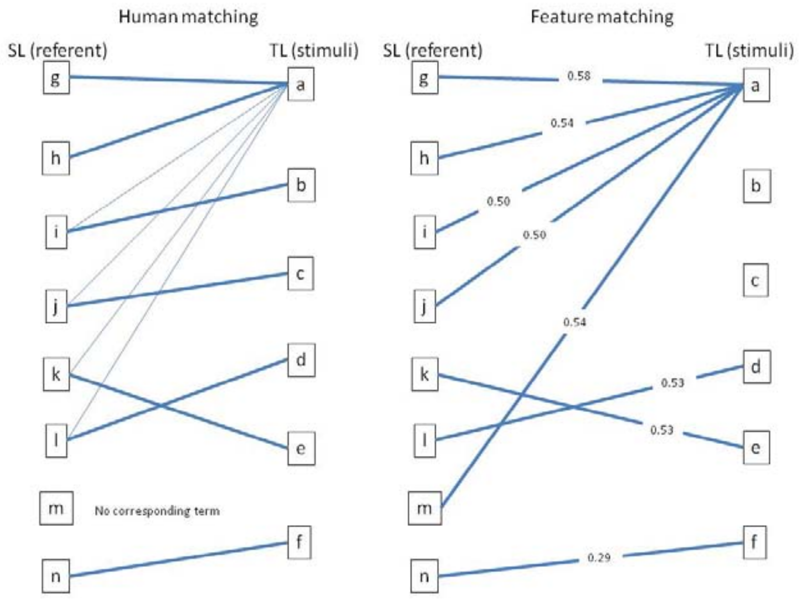

Fig. 3 Comparison with human matching: Japanese terms as stimuli

In Table 2 and 3, the scores with bold fonts are the highest scores in each row. From these tables, it can be interpreted that a concept with the highest score has the most optimal relevance to an SL concept.

The first notable point is that the application of the asymmetric parameters resulted in the asymmetric bidirectional relationships in most of the links between Danish and Japanese concepts. To be more precise, it is less optimal to use the Japanese term as stimulus, when conveying the original meaning of the Danish concept (e.g. the asymmetric score for term *c* as stimulus and *g* as referent: 0.49). On the other hand, it is more optimal to use the Danish term as stimulus, when conveying the original meaning of the Japanese concept (the asymmetric score for term *g* as stimulus and *c* as referent: 0.69).

Another point is that the majority of the identified optimal stimuli in Table 2 and 3 were the most general terms located at the highest hierarchy in the data-sets. From this viewpoint, it is difficult to assess whether the application of Tversky´s model to the terminological ontologies is successful. Hence the feature matching results are compared with the human matching results in Figure 2 and 3. In the human matching charts (the left side of the figures), the bold line indicates the ideal stimulus (optimal translation candidate) and the slim line indicates the acceptable stimuli (reasonably acceptable translation candidates) for each SL term. In the feature matching charts (the right side of the figures), the bold line indicates the optimal stimulus having the highest score for each SL term. The slim line indicates the stimuli that are not the highest score for a SL term, but having scores over 0.55.

5 Discussion

As described in Chapter 4, among all the concepts between the two ontologies, only two terms were completely matched in the string-based matching. This indicates that the English educational terminology used in respective knowledge sharing communities is immensely dissimilar, even though the educational concepts existing in the two countries are relatively similar. From this observation, it can be elaborated that it is very complicated to link concepts in two remote languages. This is because language resources having direct links between two remote languages are usually very limited, and therefore a pivot translation via English is often required both for dictionary-based human translations and for statistically-based machine translations. This also emphasizes the necessity for carefully analyzing how the meanings of a concept in one culture can be conveyed to a person in another culture through English as lingua franca.

From this viewpoint, the empirical analysis in Chapter 4 showed modest progress. The results listed in Table 2 and 3 as well as Figure 2 and 3 indicate that Tversky´s contrast model [2] is to a certain extent applicable to data-sets extracted from terminological ontologies. The application of the asymmetric parameters showed an interesting indication that it is less optimal to use the Japanese term, e.g. “c: high school-specialized course” as stimulus for a Japanese audience, when conveying the original meaning of the Danish concept e.g. “g: upper secondary education” (the asymmetric score is 0.49). On the other hand, it is more optimal to use the Danish term e.g. “g: upper secondary education” as stimulus for a Danish audience, when conveying the original meaning of the Japanese concept e.g. “c: high school-specialized course” (the asymmetric score is 0.69). Even though concepts in two cultures are mapped to each other, it is not necessarily true that a translational equivalence holds in a bidirectional way, if the two concepts are not 100% identical. This result can be explained as follows: when considering equation (1) of Tversky´s model, it is obvious that if features for the parameter α (features that are in the feature set A but not in set B: (A-B)) increases, the similarity score severely decreases. The reason for the aforementioned result (the use of Japanese term as stimulus is less optimal) is that, when we categorize the two ontologies into three blocks, we select the “generation” dimension that is the third highest dimension in the Japanese educational system. Therefore, the feature values of the first- and second highest dimensions have been inherited to all the Japanese concepts listed in Table 1. This indicates that the asymmetric parameters of Tversky´s model, to a certain degree, reflect the hierarchical structure hidden behind the feature structures of the terminological ontologies.

A final point is that, in most cases, the identified optimal stimuli based on the similarity score in Table 2 and 3 is the most general term located at the highest hierarchy in the data-sets. According to the principle of TO, when a concept is subdivided into several sub-concepts based on a dimension, an extra feature is added to each sub-concept. Hence it is often the case that concepts having more features are more specific sub-concepts. It means that the lower a concept is located in the ontology, the more features the concept inherits from superordinate concepts. If dimensions and their values at the lower part of the two ontologies are not consistent, all the inherited features are simply acting as noise in the data-sets. The positive interpretation could be that Tversky´s model is applicable to identify corresponding pairs with less noise, in other words, pairs that optimally share common features with less noise. The negative interpretation could be that Tversky´s model has limitations in identifying corresponding pairs at the optimally specific level. Considering communication in the real world, it is not incorrect to say that the relevance required in the communication is achieved in this way, since people can usually achieve a mutual understanding much easier at a reasonably general level than at a very specific level. However, as Figure 2 and Figure 3 illustrate, the optimal translation candidates selected by human are the optimally specific concepts. Hence one of the challenges is to identify the reasonably specific terms from noisy data-sets. In order to achieve this, additional investigations (e.g. implementing the feature matching for the all concepts in the two ontologies) are required.

Another future challenge is to further investigate and compare this empirical study with data-sets obtained from terminological practices as well as from the translation practices in the real world. The data-sets used in this study are English documents published by the EU commission for the Danish educational system. The EU commission has used English terminology that is standardized based on the International Standard Classification of Education (ISCED) defined by the United Nations Educational, Scientific and Cultural Organization. Therefore, the documents describing the educational system for the majority of EU member countries are based on the standardized classification and English terminology. Hence, it may be much easier to align the educational system ontologies constructed from documents published by the EU member countries. On the other hand, the Japanese ontology does not conform to the same standardized classification and terminology. By applying Tversky´s contrast model to the different data-set combinations, it may be possible to investigate the behavior of data in different scenarios. Extending from this viewpoint, the further development of the CAOS prototype for automating the knowledge extraction, ontology construction and update described in [11] could efficiently be synchronized in order to link language-specific concepts existing in different countries. To be more concrete, it might be effective to automatically extract knowledge from the domain-specific corpora based on pre-defined feature dimensions derived from a standardized classification and the terminology e.g. the ISCED classification.

Finally, Tversky´s contrast model has been extended by several researchers in different disciplines. Especially in the area of cognitive science, Tenebaum and Griffiths [12] proposed a framework that subsumes Tversky´s model of similarity by recasting Shepard´s universal law of generalization [10] in a more general Bayesian framework. Frank et al. [13] further extended this framework in [12] in order to model informative communication based on [7]. Hence, it is an obvious future challenge to apply these extended cognitive models to the aforementioned different combinations of data-sets.

6 Conclusion

In this paper, the applicability of Tversky´s contrast model derived from the cognitive theory [2] to data-sets extracted based on the Terminological Ontology method is investigated. The study indicates that the application of [2] to [1] could, to a certain extent, enable one to analyze not only the degree of relevance between concepts in two cultures, but also the degree of asymmetric relationship between the concepts. By extending [2] to e.g. [12][13], it may be feasible to investigate further how meanings of a concept in one culture can be effectively conveyed to a another culture through English as lingua franca. However, further investigations using data-sets obtained from terminological practices as well as from translation practices are needed in order to clarify the limitations pointed out in this study.

Acknowledgments. I would like to express my gratitude to my supervisors, Hanne Erdman Thomsen and Daniel Hardt for their valuable guidance and support on my Ph.D. project. I would also like to thank my colleagues in the psycholinguistic reading group at Copenhagen Business School, who provided me with a selection of highly relevant papers and constant inspiration in the area of intercultural communication and translation theories.

References

1. Madsen, B.N., Thomsen, H.E., Vikner, C.: Principles of a System for Terminological Concept Modelling. In: Proc. The 4th International Conference on Language Resources and Evaluation, ELRA, 15-19 (2004)
2. Tversky, A.: Features of Similarity. In: Psychological Review 84, 327-352 (1977)
3. Declerck, T., Krieger, H.U., Thomas, S.M., Buitelaar, P., O´Riain, S., Wunner, T., Maguet, G., McCrae, J., Spohr, D., Montiel-Ponsoda, E.: Ontology-based Multilingual Access to Financial Reports for Sharing Business Knowledge Across Europe. In: Rooz, J., Ivanyos, J. Eds.: Internal Financial Control Assessment Applying Multilingual Ontology Framework, Budapest: HVG Press, 67-76 (2010)
4. Vossen, P., Agirre, E., Calzolari, N., Fellbaum, C., Hsieh, S., Huang, C., Isahara, H., Kanzaki, K., Marchetti, A., Monachini, M., Neri, F., Raffaelli, R., Rigau, G., Tescon, M.: KYOTO: A System for Mining, Structuring and Distributing Knowledge Across Languages and Cultures. In: Proc. The 6th International Conference on Language Resources and Evaluation, Morocco, 1462-1469 (2008)
5. ISO 704: Terminology Work – Principles and Methods. Genève: ISO (2000)
6. Madsen, B.N., Thomsen, H.E., Vikner, C.: Comparison of Principles Applying to Domain Specific Versus General Ontologies. In: Oltramari, A., Paggio, P., Gangemi, A., Pazienza, M.T., Calzolari, N., Pedersen, B.S., Simov, K. (eds.): OntoLex 2004: Ontologies and Lexical Ressources in Distributed Environments. ELRA, 90-95 (2004)
7. Sperber D., Wilson, D.: Relevance: Communication and Cognition. Blackwell, Oxford (1986)
8. Guut, E.A.: A Theoretical Account of Translation – without a translation theory. In: Target: International Journal of Translation Studies, 2/2, 135-164 (1990)
9. Shepard, R.N.: Toward a Universal Law of Generalization for Psychological Science. In: Science, New Series, Vol. 237, No. 4820, 1317-1323 (1987)
10. Euzenat J., Shvaiko, P.: Ontology Mapping. Springer-Verlag, Berlin Heidelberg (2007)
11. Madsen B.N., Thomsen H.E., Halskov J., Lassen T.: Automatic Ontology Construction for a National Term Bank. In: Proc. Terminology and Knowledge Engineering Conference 2010, Dublin, 502-533 (2010)
12. Tenenbaum, J.B., Griffiths, T.L.: Generalization, Similarity, and Bayesian Inference. In: Behavioral and Brain Sciences 24, 629-640 (2001)
13. Frank, M. C., Goodman, N. D., Lai, P., Tenenbaum, J. B.: Informative Communication in Word Production and Word Learning. In: Proc. The 31st Annual Meeting of the Cognitive Science Society (2009)