Comparison of Google Translation with Human Translation

Haiying Li, Arthur C. Graesser and Zhiqiang Cai
University of Memphis, Institute for Intelligent Systems, Memphis, USA
{hli5; graesser; zcai}@memphis.edu

Abstract

Google Translate provides a multilingual machine-translation service by automatically translating one written language to another. Google translate is allegedly limited in its accuracy in translation, however. This study investigated the accuracy of Google Chinese-to-English translation from the perspectives of formality and cohesion with two comparisons: Google translation with human expert translation, and Google translation with Chinese source language. The text sample was a collection of 289 spoken and written texts excerpts from the *Selected Works of Mao Zedong* in both Chinese and English versions. Google translate was used to translate the Chinese texts into English. These texts were analyzed by the automated text analysis tools: the Chinese and English LIWC, and the Chinese and English Coh-Metrix. Results of Pearson correlations on formality and cohesion showed Google English translation was highly correlated with both human English translation and the original Chinese texts.

Instruction

The use of the automatic machine translation has increased in recent years with the dramatic increase in communication between countries. Google Translate provides a billion translations a day for 200 million users (Shankland 2013), and apparently offers better performance than other machine translation tools available to the public (Seljan, Brkić and Kučiš 2011). Google translation is a widely used translation tool for inexpensive and instant access to general information about the original texts for moderate quality translation (Anazawa et al. 2013). The previous evaluation of Google translation focused on the levels of words, phrases, sentence length, syntactic structure (Seljan, Brkić and Kučiš 2011), intelligibility and usability (Anazawa et al. 2013), and BLEU (Bilingual Evaluation Understudy) (Papineni et al. 2002). No empirical studies have been conducted on the quality at the discourse level. This study investigated the accuracy of Google translation from the perspectives of discourse level on two metrics: formality and cohesion.

Google Translation

Google is an automatic machine-translation service provided by Google Inc. It translates one written source language to another directly or with English as a medium (Boitet et al. 2009).

Google translation employed the statistical machine translation (Brown et al. 1990) by using both linguistic modeling, statistical decision theory, and matching probabilities (Ney 1995) to determine the most often used translation. The most popular system of Google Translate included the phrase-based model with small text chunks and reordering (Koehn, Och, and Marcu 2003), hierarchical phrase-based models (Chiang 2007), and hierarchical and syntactic models (Zollmann and Venugopal 2006). These three models achieved similar quality of translation, but hierarchical and syntactic models showed more benefits in Chinese-to-English translation (Zollmann and Venugopal 2006).

The criteria of the translation quality encompassed adequacy, fidelity, and fluency of the translation (Hovy 1999). The classic viewpoint of measuring translation performance is that "the closer a machine translation is to a professional human translation, the better it is" (Papineni et al. 2002). The components of the judgment included "translation closeness metric" like word error rate metrics, and "a corpus of good quality human reference translations" (Papineni et al. 2002). However, these criteria were all restricted to the word accuracy and the clause or sentence level. Few automatic evaluation added semantic metric between automatic and reference translations by comparing shallow semantic roles and discourse representations, such as *Asiya* (Giménez and Gonzàlez 2010).

Machine translation fails in the accuracy in grammar, complex syntactic, semantic and pragmatic structures. This results in nonsensical errors in grammar and meaning processing. Some languages were translated more accurately than others, such as French to English (Shen 2010) and Italian to English (Pecorao 2012). As for Chinese to English translation, the quality of Google English translation was better if the original Chinese texts to be translated were short and simple sentences (Shen 2010). However, no empirical studies have been conducted systematically to compare the Chinese-English Google translation at a multi-text, discourse level.

Human Translation

Human translation is influenced by the characteristics of source-target language transfer, cultural context and individual translators' translation ability (Bassnett and Lefevere 1992; Wong and Shen 1999). During the process of translation, the ways of handling decoding and recoding, problems of equivalence (Gentzler 2001), loss and gain, and untranslatability (Bassnett 2002) will yield varied translated versions by different translators due to the different interpretation of both source and target languages. With machine translation, however, the impact of individual translation capacity will be avoided. Nonetheless, the quality of machine translation is still a concern.

To evaluate the quality of machine translation, it is a necessity to compare the machine translation with the human translation and the source language at a deeper and more comprehensive textual level, including the levels of the words, syntax, semantics, pragmatics and discourse. With this multilevel comparison, we may have an overall view on the quality of machine translation, as compared with human translation and the source language.

Formality and Cohesion

Formality is proposed to "be the most important dimension of variation between styles or registers" (Heylighen and Dewaele 2002). Formality was defined as "The type of speech used in situations when the speaker is very careful about pronunciation and choice of word and sentence structure" (Richards, Platt, and Platt 1997), or as "a linguistic system based on logic and/or mathematics that is distinguished by its clarity, explicitness, and simple verifiability" (Bussmann 1996). Even though these definitions failed to provide specific linguistic features that could predict formality, they claimed that formality was able to be represented by linguistic features.

Heylighen and Dewaele (2002) claimed that the formal language has the features of detachment, accuracy, rigidity, cognitive load, and dense information. Conversely, the informal language has the features of flexibility, directness, implicitness, involvement and less information. They proposed a measure of formality with the frequencies of part of speech (POS) at the coarse lexical level of formality. They asserted nouns, adjectives, articles and prepositions are more frequently used in formal style, but pronouns, adverbs, verbs and interjections are more frequently used in informal style. The formula is as follows:

$$F - score = [(noun + adjective + preposition + article - pronoun - verb - adverb - interjection + 100)/2]$$

However, this measure was limited to the word level and alphabetic language so it is difficult to generalize to the symbolic Chinese language. Therefore, we propose a new method to measure formality at the multiple discourse levels, including cohesion, narrativity, space and time, and embodiment. These features were provided by both Chinese and English automated discourse analysis tools, namely Coh-Metrix (Graesser et al. 2004; McNamara et al. 2014) and Linguistic Inquiry and Word Count (LIWC) (Pennebaker, Booth, and Francis 2007).

Coh-Metrix is a computational tool that analyzes texts at multi-textual levels related to conceptual knowledge, cohesion, lexical difficulty, syntactic complexity, and simple incidence scores (Graesser et al. 2004; McNamara et al. 2014). Coh-Metrix automatically measures the diverse types of cohesion through the quantitative connection between text elements and constituents with individual linguistic indices as well as five principal components (McNamara et al. 2014). These five major dimensions include (1) Narrativity, closely affiliated with stories and everyday oral conversation, (2) Deep Cohesion, text with more cohesive connectives such as causal, intentional, and temporal connectives, (3) Referential Cohesion, words and ideas overlapping across sentences and the entire text, (4) Syntactic Simplicity, few words and simple, familiar syntactic structures, and (5) Word Concreteness, easily arousing mental images and richer semantic specifications.

LIWC is a text analysis software program with a text processing module and an internal default dictionary. The LIWC tool counts the percentage of words in a document that maps a specific word in linguistic or psychological categories (Pennebaker et al. 2007). The Chinese LIWC dictionary was developed by National Taiwan University of Science and Technology based on the LIWC 2007 English dictionary. In addition, some word categories unique to the Chinese language were added in the Chinese LIWC dictionary (Huang et al. 2012). The empirical studies showed reliable correlations between English LIWC categories and Chinese LIWC categories (Li et al. 2012).

Latent Semantic Analysis (LSA) (Landauer et al. 2007) is a computational model in natural language processing to extract semantic representation from large corpus. We use LSA to study semantic overlap between sentences and paragraphs. The semantic overlap is computed by the cosine between vectors representing two sentences of paragraphs. In this paper, we only consider semantic overlap between adjacent sentences. Our Chinese LSA space is generated from large reference Chinese corpus with various genres including economy, language arts (e.g., classic fictions and modern fictions), social studies (e.g., history, philosophy and politics), science and military (Li et al. 2012).

Content Word Overlap (CWO) measures cohesion with the overlap of content words between adjacent sentences in the documents (McNamara et al. 2014). In our computation, a word is considered as a content word if it does not occur in a function word list. The function word list we use is based on *Dictionary of Function Words in Modern Chinese Language* (Wang 1998). The content word overlap score between two adjacent sentences is binary (either "0" or "1"). The content word overlap score of a document is the average of the scores of all the adjacent sentence pairs.

This study focuses on two research questions at the multiple textual levels: (1) whether Google translation is similar to professional human translation in terms of formality and semantic cohesion; and (2) whether Google translation is as similar to the Chinese source language as is human translation to the Chinese language in terms of formality and cohesion.

Method

Corpora

The corpora included both original Chinese and English translated documents: the original Chinese corpus (OC), the Google translation (GT), and the human translation (HT). OC was collected from *Selected Works of Mao Zedong*, which included representative articles Mao written during the different periods of the Chinese revolution. The reasons for the choice of Mao's articles are twofold. First, these articles were all written by Mao himself, so the different individual writing styles would be avoided. Second, these articles have been translated into English by a special committee who are all experts at both the English language and the Chinese language. Therefore, the impact of the individual translators on the translation would be controlled. Meanwhile, the accuracy and fluency of this human translation is considered impeccable.

This expert translation was used as the human English translation corpus (HT). Then we used Google Translate to translate the Chinese documents to English, which is the GT corpus. Each corpus consisted of 289 documents, totally 867.

Procedure and Design

Three steps were involved before the data analysis. First, the automated text analysis tools were used to analyze all the documents. The Chinese corpus was performed by the Chinese LIWC (CLIWC; 71 indices), LSA and CWO. The two English corpora were processed with the English LIWC (ELIWC; 64 indices) and the English Coh-Metrix (ECoh-Metrix; 54 indices), which includes the LSA and CWO. Thus, five data sets were included in this study: Chinese corpus generating CLIWC data set, both GT and HT corpora generating ELIWCs and ECoh-Metrixes.

Second, the Component Model (Li et al. 2013b) was used to compute the principal component scores for these five data sets. The Component Model employed the large English and Chinese reference corpora to perform a principal components analysis (PCA; see Table 1) to get the means, standard deviations and the coefficients of each variable in this analysis. The Component Model is computed with the formula below:

$$y = \sum_1^n \left(\frac{x-\mu}{s} \gamma \right)$$

In this formula, y is a component score we want for a particular corpus (PC), x is the value of each index in a document of PC, μ is the mean of the corresponding index from reference corpus (RC), s is the standard deviation of the corresponding index from RC, γ is the coefficient of the corresponding variable from RC. 1 to n means the number of indices in each component based on PCA in reference corpora. ∑ means the sum of all the values of the indices in each component. Thus, we get the component scores for PC according to this model.

Table 1. Principal components in the English and Chinese reference corpora and the formulas of the composite formality score

ID	CLIWC	ELIWC	ECoh-Metrix
1	Narrativity	Narrativity	Narrativity
2	Cognitive Complexity	Cognitive Complexity	Deep Cohesion
3	SpaceTime	Conjunctions	Referential Cohesion
4	Positive Emotion	Positive Emotion	Syntactic Simplicity
5	Negative Emotion	Negative Emotion	Word Concreteness
6	Embodiment	Embodiment	—————
7	Personal Concerns	—————	—————
Formality	(-1-2-3)/3	(-1-2+3)/3	(-1+2+3-4-5)/5

We performed three Component Models: CLIWC on the Chinese data set, along with both ELIWC and ECoh-Metrix on English data sets. Thus, each component score of either Chinese or English corpus was obtained and standardized from these three models in terms of Chinese and English reference corpora.

The composite formality scores of these five data sets were computed based on the component scores. The formulas of formality scores are listed in Table 1.

The correlational design was used to examine the relationship between GT, HT, and OC in terms of formality and cohesion (LSA and CWO).

Results and Discussion

The relationships between Google translation, human translation and the original Chinese documents were examined first at the overall textual level by means of formality and then at the cohesion level. The Pearson correlations were performed on the composite formality scores first and then on LSA and CWO.

Correlations with Formality

The overall formality scores were analyzed by the Pearson correlations with the Chinese and English LIWC data sets to examine the associations among the original Chinese, Google translation and human translation (See Table 2). The Chinese formality score (*Mean* = -.46, *Standard Deviation* = .40) had a significant and positive correlation with both human translation (*Mean* = .46, *Standard Deviation* = .37) and Google translation (*Mean* = .26, *Standard Deviation* = .38). These findings suggest that from the perspective of formality computed by LIWC, both Google translation and human translation are associated with the original Chinese language. The Chinese and English LIWC dictionaries had a very high inter-rater reliability (Huang et al. 2012), and the components used in the computation of formality score had the very high correlations between the Chinese component and its corresponding English component (Li et al. 2012). Therefore, findings imply that Google translation is similar to human translation from the perspective of formality.

Furthermore, the correlation between LIWC formality and Google translation and human translation was extremely high (r = .80), but for Coh-Metrix formality (*Mean* = .29, *Standard Deviation* = .27 for human translation; *Mean* = .29, *Standard Deviation* = .29 for Google translation), the correlation was also high (r = .68), but not as high as LIWC formality. These findings suggest that Google translation is more similar to human translation with LIWC formality than Coh-Metrix formality. It is likely that LIWC formality was measured by the word categories in terms of the linguistic and psychological word categories, whereas Coh-Metrix formality was measured at the multiple discourse levels such as referential and deep cohesion, syntactic structure, narrativity, and word concreteness. This phenomenon of LIWC formality of Google translation more similar to the source language perhaps supports the mechanism of Google translation, phrase-to-phrase translation.

Table 2. Pearson correlations of formality scores

	1	2	3	4	5
C-LIWC	1				
HT-LIWC	.24**	1			
GT-LIWC	.23**	.80**	1		
HT-Coh-Metrix	-.06	.08	.06	1	
GT-Coh-Metrix	-.03	.25**	.21**	.68**	1

Note. ** p <.01 level. C=Chinese data set; HT=English human translation data set; GT = English Google translation data set (Same as below).

Correlations with Cohesion

LSA and CWO are important indices to examine the cohesion of the text. Since the sentence length (SL) impacts the LSA and CWO value, the SL is also included in the correlational analysis. The results indicated that the LSA of Google translation (*Mean* = .24, *Standard Deviation* = .09) had a high positive correlation with that of the Chinese version (*Mean* = .27, *Standard Deviation* = .09) compared with the human translation (*Mean* = .23, *Standard Deviation* = .09). Furthermore, the two English translations were highly correlated in LSA.

Table 3. Pearson correlations of LSA, CWO and Sentence Length (SL)

	LSA		CWO		SL	
	C	HT	C	HT	C	HT
HT	.59**		.42**		.77**	
GT	.76**	.70**	.57**	.68**	.97**	.80**

Note. ** p <.01 level. HT means human translation and GT means Google translation.

Similarly, the CWO of Google translation (*Mean* = .12, *Standard Deviation* = .06) had a higher positive correlation with that of the Chinese (*Mean* = .67, *Standard Deviation* = .18), compared with human translation (*Mean* = .11, *Standard Deviation* = .06). Like LSA, the two English translations were highly correlated in CWO.

These findings suggest that Google translation is more similar to the Chinese language than human translation at the level of the referential cohesion and conceptual cohesion. It is perhaps that Google translation uses the same words or expressions without considering the flexibility of choosing the alternative expressions. It is also perhaps that Google translation mechanically uses the period of the Chinese as an end of a sentence without considering any flexibility of reorganizing the sentences based on the complete semantic meanings in the Chinese language. Therefore, the SL is an important index to examine whether Google translation has any flexibility with the Chinese at the level of changing the number of sentences like human translation.

Similarly, the SL of the Google translation (*Mean* = 25, *Standard Deviation* = 9) was positively correlated with the SL of the Chinese (*Mean* = 25, *Standard Deviation* = 8). This correlation was enormously high, almost up to 1. The SL of the human translation (*Mean* = 26, *Standard Deviation* = 7) was also highly correlated, but not as high as the Google translation.

Table 4. Examples from the Chinese, Google translation and human translation with underlined words and expressions being the same in both translations

ID	Version	Texts
1	Chinese	这篇文章的观点是正确的。合作社必须强调做好政治工作。政治工作的基本任务是向农民群众不断地灌输社会主义思想，批评资本主义倾向。
	EGT	The view of the <u>article</u> is correct. The <u>cooperatives</u> <u>must</u> be stressed that the good <u>political work</u>. <u>The basic</u> task of the political work is to continue to instill <u>socialist ideology</u>, criticized the tendency of capitalism to the peasant masses.
	EHT	The viewpoint of this <u>article</u> is correct. <u>Co-operatives</u> <u>must</u> stress doing <u>political work</u> well. <u>The basic</u> requirement of political work is constantly to imbue the peasant masses with a <u>socialist ideology</u> and to criticize capitalist tendencies.
2	Chinese	在打倒地主阶级和官僚资产阶级以后，中国内部的主要矛盾即是工人阶级与民族资产阶级的矛盾，故不应再将民族资产阶级称为中间阶级。
	EGT	Down with <u>the landlord class and the bureaucrat-capitalist class,</u> the principal <u>contradiction</u> in China's internal is the <u>contradiction</u> <u>between the working class and the national bourgeoisie</u>, national bourgeoisie, it <u>should no longer</u> be called the middle <u>class</u>.
	EHT	With the overthrow of <u>the landlord class and the bureaucrat-capitalist class</u>, the <u>contradiction between the working class and the national bourgeoisie</u> has become the principal <u>contradiction</u> in China; therefore the national bourgeoisie <u>should no longer</u> be defined as an intermediate <u>class</u>.

The findings of the almost perfect correlation between Google translation and the original Chinese version suggest that Google Translate automatically translates the sentence by sentence with the punctuation in the source language without much flexibility. It neither considers the difference in syntactic structure between the source language and the target language.

However, the human translation is able to separate the complete semantic meanings into individual independent sentences considering the different characteristics between the Chinese and English languages. For example, in Chinese, a sentence is allowed to include multiple subordinate clauses and/or coordinate clauses without conjunctions, but in English, the conjunctions should be used to connect clauses. Meanwhile, the comma plays an important role in Chinese because it frequently occurs at the end of a clause and separates clauses (Lin 2000). Two short excerpts provide this as an illustration (See Table 4).

These two examples illustrate that Google translation almost chooses the same words or expressions as human translation, and perhaps even uses the same sentence structure. For example, the three-sentence structures are all the same except some word choices in the first example. In the second example, both use “with” preposition structure to initiate the sentence. However, the second example shows human translation flexibly breaks up sentences and complements with the connectives such as “therefore.” These examples further support the claim that Google translation helps in general information understanding instead of the grammatical accuracy.

Conclusion

This study evaluated Google English translation with the comparison of human English translation and the original Chinese. The results indicated that both translations had a small, but significant correlation with the Chinese in formality, but the translations had a high correlation in LIWC and Coh-Metrix composite formality score. In terms of cohesion, both translations had a high correlation with each other in LSA and CWO, but Google translation had higher correlations with the Chinese than human translation. Considering the sentence length, both translations were correlated highly, but Google translation had a higher correlation (almost to 1) with the Chinese.

These findings imply that Google translation is associated with the original Chinese similar to human translation from the perspectives of formality and cohesion. Since formality was computed with the comprehensive multiple linguistic and psychological metrics and cohesion was measure with LSA and CWO, it is possible to make a conclusion that Google translation is close to human translation at the semantic and pragmatic levels. However, at the syntactic level or the grammatical level, it needs improving. In other words, Google translation yields a decipherable and readable translation even if grammatical errors occur. Google translation provides a means for people who need a quick translation to acquire information. Thus, computers provide a fairly good performance at translating individual words and phrases, as well as more global cohesion, but not at translating complex sentences.

The further study should enlarge the corpus, such as including documents in the different genres and by different authors. In addition, error analyses should be conducted to find out the specific inaccurate translations so that the translation systems could be improved.

References

Anazawa, R., Ishikawa, H., Park, M. J., and Kiuchi, T. 2013. Online Machine Translation Use with Nursing Literature: Evaluation Method and Usability. *Computers Informatics Nursing* 31(2): 59-65.

Bassnett, S. 2002. *Translation Studies*. London: Routledge.

Bassnett, S., and Lefevere, A. 1992. *Translation, History and Culture*. London: Cassell.

Boitet, C., Hervé B., Mark S., and Valérie B. 2009. Evolution of MT with the Web. In *Proceedings of the Conference Machine Translation 25 Years On*, 1-13. Cranfield: Bedfordshire.

Brown, P. F., Cocke, J., Pietra, S. A. D., Pietra, V. J. D., Jelinek, F., Lafferty, J. D., Mercer, R. L., and Rossin, P. 1990. A Statistical Approach to Machine Translation. *Computational Linguistics* 16(2): 76-85.

Bussmann, H. 1996. *Routledge Dictionary of Language and Linguistics*. London: Routledge.

Chiang, D. 2007. Hierarchical Phrase-based Translation. *Computational Linguistics* 33(2): 201-228.

Gentzler, E. 2001. *Contemporary Translation Theories*. Clevedon: Multilingual Matters.

Giménez, J., and Màrquez, L. 2010. Asiya: an Open Toolkit for Automatic Machine Translation (meta-) evaluation. *The Prague Bulletin of Mathematical Linguistics* 94(1): 77-86.

Graesser, A.C., McNamara, D.S., Louwerse, M. M., and Cai, Z. 2004. Coh-metrix: Analysis of Text on Cohesion and Language. *Behavior Research Methods, Instruments, & Computers* 36(2): 193-202.

Heylighen, F., and Dewaele, J. M. 2002. Variation in the Contextuality of Language: An Empirical Measure. *Foundations of Science* 7(3): 293-340.

Hovy, E. 1999. Toward Finely Differentiated Evaluation Metrics for Machine Translation. In *Proceedings of the EAGLES Workshop on Standards and Evaluation*, 127-133. Pisa, Geneva.

Huang, J., Chung, C. K., Hui, N., Lin, Y., Xie, Y., Lam, Q., Cheng, W., Bond, M., and Pennebaker, J. W. 2012. 中文版語文探索與字詞計算字典之建立 [The Development of the Chinese Linguistic Inquiry and Word Count Dictionary]. *Chinese Journal of Psychology* 54(2): 185-201.

Koehn, P., Och, F. J., and Marcu, D. 2003. Statistical Phrase-based Translation. In Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology, Vol. 1, 48-54. Edmonton, Canada: Association for Computational Linguistics.

Landauer, T., McNamara, D., Dennis, S., and Kintsch, W. 2007. *Handbook of Latent Semantic Analysis*. Mahwah, N.J.: Erlbaum.

Li, H., Cai, Z., Graesser, A. C., and Duan, Y. 2012. A Comparative Study on English and Chinese Word Uses with LIWC. In G. M. Youngblood, and P. M. McCarthy (Eds.), *Proceedings of the Twenty-Fifth International Florida Artificial Intelligence Research Society Conference*, 238-243., Palo Alto, California: AAAI Press.

Li, H., Graesser, A. C., and Cai, Z. 2013b. Component Model in Discourse Analysis. In S. K. D'Mello, R. A. Calvo, and A. Olney (Eds.), *Proceedings of the 6th International Conference on Educational Data Mining*, 326-327. Memphis, TN: International Educational Data Mining Society.

Lin, S. 2000. 标点符号的学习和应用 [Study and Application of Punctuation]. Beijing: Beijing People's Publisher.

McNamara, D., Graesser, A.C., McCarthy, P., and Cai, Z. 2014. *Automated Evaluation of Text and Discourse with Coh-Metrix*. Cambridge, MA: Cambridge University Press.

Ney, H. 1995. On the Probabilistic Interpretation of Neural Network Classifiers and Discriminative Training Criteria. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17(2): 107-119.

Papineni, K., Roukos, S., Ward, T., and Zhu, W. J. 2002. BLEU: A method for Automatic Evaluation of Machine Translation. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, 311-318. Philadelphia, PA.

Pennebaker, J. W., Booth, R. J., and Francis, M. E. 2007. *Linguistic Inquiry and Word Count: LIWC* [Computer software]. TX: LIWC.net, Austin.

Richards, J., Platt, J., and Platt, H. 1997. *Dictionary of Language Teaching and Applied Linguistics*. London: Longman.

Seljan, S., Brkić, M., & Kučiš, V. 2011. Evaluation of Free Online Machine Translations for Croatian-English and English-Croatian Language Pairs. In *Proceedings of the 3rd International Conference on the Future of Information Sciences: INFuture2011-Information Sciences and e-Society*, 331-345. Zagreb, Croatia.

Shankland, S. May 18 2013. Google Translate Now Serves 200 Million People Daily. *CNET*. Retrieved from http://news.cnet.com/8301-1023_3-57585143-93/google-translate-now-serves-200-million-people-daily/

Shen, E. June 2010. Comparison of Online Machine Translation Tools. *Translation and Localization*. Retrieved from http://www.tcworld.info/e-magazine/translation-and-localization/article/comparison-of-online-machine-translation-tools/

Wang, Z. 1998. *现代汉语虚词词 [Dictionary of Function Words in Modern Chinese Language]*. Shanghai: Shanghai Lexicographical Publishing House.

Wong, D., and Shen, D. 1999. Factors Influencing the Process of Translating. *Meta: Journal des Traducteurs [Meta: Translators' Journal]* 44(1): 78-100.

Zollmann, A., and Venugopal, A. 2006. Syntax Augmented Machine Translation via Chart Parsing. In Proceedings of NAACL 2006 Workshop on Statistical Machine Translation, 4-9. New York, NY.

Acknowledgments

This work was supported by the National Science Foundation (BCS 0904909) for the Minerva project: Languages across Culture. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF.