Sharing Problems and Solutions for Machine Translation of Spoken and Written Interaction

Sherri Condon   Keith Miller

The MITRE Corporation
7515 Colshire Drive
McLean, VA 22102-7508
{scondon, keith}@mitre.org

Abstract

Examples from chat interaction are presented to demonstrate that machine translation of written interaction shares many problems with translation of spoken interaction. The potential for common solutions to the problems is illustrated by describing operations that normalize and tag input before translation. Segmenting utterances into small translation units and processing short turns separately are also motivated using data from chat.

1 Introduction

The informal, dialogic character of oral interaction imposes demands on translation systems that are not encountered in well-formed, monologic texts. These differences make it appear that any similarities between the machine translation of text and speech will be limited to core translation components, as opposed to pre- and post-processing operations that are linked to the medium.

In this paper, we demonstrate that many challenges of translating spoken interaction are also encountered in translating written interaction such as chat or instant messaging. Consequently, it is proposed that solutions developed for these common problems can be shared by researchers engaged in applying machine translation technologies to both types of interaction. Specifically, preprocessing operations can address many of the problems that make dialogic interaction difficult to translate in both spoken and written media.

After surveying the challenges that are shared in machine translation of spoken and written interaction, we identify several areas in which preprocessing solutions have been proposed that could be fruitfully adopted for either spoken or written input. The speech recognition problem of discriminating out of vocabulary words from unrecognized vocabulary words is equivalent to the problem of discriminating novel forms that emerge in chat environments from words that are unrecognized due to nonstandard spellings. We suggest that a solution based on templates like those used in example-based translation could be a useful approach to the problem for both spoken and written input. Similarly, other preprocessing operations that tag input for special processing can be used to facilitate translation of problematic phenomena such as discourse markers and vocatives. Finally, we explore the possibility that the complexity of translating interaction can be reduced by translating smaller packages of input and exploiting participants' strategies for packaging certain discourse functions in smaller turn units.

2 Challenges for translation of spoken and written interaction

In illustrating the problems for machine translation that are shared by both spoken and written interactions, we take for granted that readers are aware of examples that occur in spoken interaction because these are available in the literature and from direct observation of personal experience. Therefore, we focus on providing examples of written interaction to demonstrate that the same kinds of challenges arise in translation of chat and instant messages. Most of the examples we present are taken from logs of chat interactions collected from 10 chat channels in 8 languages during July of 2001. The examples are presented exactly as they appeared in the logs.

2.1 Ellipsis and fragments

The elliptical and fragmentary quality of ordinary spoken dialogue is well-known and is characteristic of chat interaction, too, as in (1).

(1) a. faut voir
    b. voir koi?

The French expression *il faut* "it is necessary" is used without the pleonastic pronoun *il*, and the verb *voir* "to see" is used without a direct object expressing what it is necessary to see. The writer may have intended to use *voir* intransitively, as in the English expression *we'll have to see,* but the interlocutor who responded with (1b) asks "to see what?" and omits both the pronoun and the verb *faut*. (Creative spelling such as the convention that replaces the *qu* in *quoi* with *k* in *koi* is discussed in section 3.3.) Though it is unlikely that preprocessing operations will be able to add information that is missing from fragments and elliptical expressions, these problems interact with preprocessing operations such as segmentation of units for translation (see 2.9).

2.2 High function/low content terms

Spoken interaction is replete with formulaic expressions that serve significant interactional functions, but have little or no internal structure. These include greetings, leave-takings, affirmations, negations, and other interjections, some of which are illustrated in (2).

(2) a. re esselamu aleyküm
    b. in like califonia
    c. jose…lets make 1

The expression *re* is a conventional greeting in chat interaction with the function re-greet, as in *hello again*. In (2a) it is used on a Turkish chat channel preceding a greeting borrowed from Arabic with the literal meaning "peace to you."

The example in (2b) demonstrates that chat interaction includes expressions such as *like* that are usually associated exclusively with speech. *Like* and discourse markers such as *well, now, so* and *anyway* occur frequently in chat interaction. Wiebe et al. (1995) identify discourse markers as a major area of difficulty for translation of spoken interaction. Many discourse markers are homophonous and/or homographic with lexical items that have very different meanings, and the need to disambiguate polysemous words has received much attention in the language processing and machine translation literature.

The use of vocative proper names illustrated in (2c) is frequent in chat interaction, where participants' nicknames are used to direct messages to specific interlocutors. In a small sample of 76 messages from our chat logs, 31% included vocative uses of participants' nicknames. In speech and in chat interaction like (2), where punctuation is unpredictable (see 3.2), capitalization cannot be relied on to identify proper names. The complexity of translating proper names has also received considerable attention in the machine translation research community, and translation of proper names has been proposed as an evaluation measure for machine translation (Papineni et al., 2002; Vanni and Miller, 2002).

2.3 Vagueness

Though vagueness is a problem in all language use, Wiebe et al. (1995) identify it as a major problem for translation of spoken interaction, citing metaphorical expressions such as *de alli,* literally, "from there," translated as *after that*. More pervasive are the deixis and vagueness that result from the shared online context that participants in real time interaction can rely on, compared to communication environments in which relevant context must be encoded in the text itself. Researchers have demonstrated the increased explicitness and structural complexity of asynchronous interaction, in which delays between the transmission of messages preclude immediate feedback, compared to synchronous interaction, in which it is expected that messages will be responded to immediately (Chafe, 1982; Condon and Cech, forthcoming; Sotillo, 2000). Similarly, Popowich et al. (2000) report that a high degree of semantic vagueness is a problem for translating closed captions because the visual context provided by the television screen supplies missing details.

2.4 Anaphora

Another consequence of the synchronous communication environments in written and spoken interaction is the high frequency of pronouns and deictic forms. Wiebe et al. (1995) report that 64% of utterances in a corpus of spoken Spanish contained pronominals compared to 48% of sentences in a written corpus. Similarly, numerous studies have demonstrated the high frequency of personal

pronouns, especially first person pronouns, in chat interaction compared to other types of written texts (Ferrara, Brunner and Whittemore, 1991; Flanagan 1996; Yates, 1996). Pronouns are particularly problematic for translation when the target language makes distinctions such as gender that are not present in the source language. To determine the appropriate inflection, the antecedent of the pronoun must be identified, and resolution of pronoun antecedents is another thorny problem that has attracted much attention from researchers in machine translation.

2.5 Juncture

Along with the liberties that participants in chat interaction take with spelling and punctuation conventions (see 3.1,2), they also deliberately (and undoubtedly sometimes accidentally) omit spaces between words, as in (3).

(3) a. selamunaleykum (= selamun aleykum)
    b. aleykumselam (= aleykum selam)

The Turkish "peace to you" greeting in (3a) is a variant of (2a) and is usually represented as two words, though the merged forms in (3a) and in the conventional reply (3b) occur several times in a sample of our corpus. Consequently, one of the basic challenges for speech recognition, identification of word boundaries, is also a problem in chat interaction.

2.6 Colloquial terms, idioms and slang

Wiebe et al. (1995) use the *term conventional constructions* to refer to idiosyncratic collocations and tense/aspect usage. Colloquial or idiomatic usage complicates translation of both spoken and chat interaction, though it is less frequent in formal writing. (4) provides some examples from chat.

(4) a. have a ball, y'all
    b. do u sleep there n stuff

Like discourse markers, expressions such as *have a ball* in (4a) and *[and] stuff* in (4b), have both compositional and idiomatic meanings, which causes ambiguity that must be resolved.

2.7 Code-switching

Code-switching is common in multilingual speech communities, and the participants in communication environments like chat tend to be multilingual. The Turkish to English switch in (5a) illustrates.

(5) a. anlamami istedigin seyi anlamadim sorry
    b. salam mon frere

In (5b) from the #paris chat channel, Arabic *salam* "peace" is used as a greeting. Not only are these switches problematic for translation engines designed to map a single source language into a single target language, but also translation into a single language eliminates the sociolinguistic and pragmatic effects of code-switching.

2.8 Language play

Another consequence of the informal contexts in which speech and chat interaction occur is the playful use of language for entertainment, and in online environments like chat, where fun often is the primary attraction, humor and language play are valued (Danet et al., 1995). In addition to play with identity and typographic symbols, which have become conventional in chat interaction, novel games like (6) emerge.

(6) a. wew
    b. wiw
    c. wow

(6) is part of a sequence on a Cebuano language channel in which the game seems to be to produce consonant frames with different vowels. It ocurred after another game in which participants inserted a vocative use of *baby* (as in *hey baby*) into almost every message, often accompanied by additional codeswitching into English, and finally prompting the protest, "you guys have been on that baby thing for ages."

2.9 Segmentation of translation units

Just as spoken interaction does not include clear delimiters for word boundaries, it also lacks conventional means of marking larger units of discourse that would be analogous to sentence punctuation in written texts. Similarly, though chat interaction is written, punctuation is often inconsistent or absent. For example, vocative nicknames used to address messages to specific participants may not be separated from the remainder of the message by any punctuation or they may be separated by commas, colons, ellipses, parentheses, brackets, and emoticons. The same range of possibilities occurs for punctuation between sentences, which is frequently absent. Consequently, it is difficult to segment input into consistent units that translation components can anticipate.

3 Analogous Challenges in Spoken and Written Interaction

Another set of problems that arise in translation of written interaction are not found in spoken interaction because they involve the typographic symbols that render language in written form. However, most of the problems have analogies in spoken interaction, just as lack of punctuation in writing causes the same juncture and segmentation problems encountered in speech. Most of the challenges in Section 2 represent problems for translation from the source language to the target language, whereas the challenges in this section primarily complicate the problem of recognizing the source message.

3.1 Unintentional misspellings and typographical errors

Nonstandard spellings occur so frequently in chat interaction that it is difficult to find examples that do not contain them, as (2b) illustrates above. Online interaction also contains many deliberate misspellings that are discussed in the next section. In addition to misspellings like (2b) and (7a), we classify as typographic errors the many instances like (7b) in which participants fail to punctuate contractions (though these may be deliberate).

(7) a. hi evenybody
    b. bon jai mon vrai nick crisse
       "good, I have my true nickname crisse"

Unlike the English contraction *I've,* the French contraction *j'ai* "I have" is not optional: it is always spelled *j'ai* and neither *je ai* nor *jai* exist in the French language. This kind of misspelling is analogous to mispronunciations and speech errors like slips of the tongue in speech, though clearly these anomalous forms are much more frequent in chat interaction than in speech.

Another type of problem is the failure to use diacritic symbols associated with letters in some orthographic systems. For example, in French the letter "a" without an accent represents the 3rd person singular present tense form of the verb "have," while the form à is the preposition "to." Both of these forms are pronounced the same, but in other cases the diacritic signifies a change in both pronunciation and meaning. For example, *marche* is the 3rd person singular present tense form of the French verb *marcher* "to work, go" and is pronounced like English *marsh* with one syllable, but *marché* "market" is a noun pronounced with a second syllable [e]. Consequently, the failure to follow orthographic conventions, creates homographs that present the same identification and ambiguity problems as homophones do in speech.

3.2 Creative spelling, rebus, and abbreviations

Online interaction is famous for the creative and playful conventions that have emerged for frequently used expressions, and though most of these originated in English, it is now possible to observe French *mdr* (mort de rire "dying of laughter") with English *lol* (laughing out loud) or *amha* (à mon humble avis "in my humble opinion") like English *imho* (in my humble opinion) and even Portuguese *vc* (voce "you"). Like the nonstandard spellings that are unintentional, these deliberate departures from convention are so frequent that we have already seen several instances of rebus forms in (2c) and (4b) and the replacement of Romance *qu* by *k* in (1). Other examples include English *pls* "please" and *ur* "your," Turkish *slm* (selam "peace") and the French forms in (8).

(8) a. ah wi snooppy ? "ah yes, snooppy?"
    b. et c pa toi   "and it is not you"

In (8a) *oui* "yes" is spelled *wi*, which reflects the pronunciation, and in (8b) the rebus form *c* represents *c'est* "it is," both pronounced [se], while *pas* is also spelled as pronounced, without the silent *s*. In the creative and rebus spellings, the nonstandard forms typically reflect the pronunciation of the word and the pronunciation is often a reduced one, as in *hiya* "hi you" and *cyah* "see you." Consequently, these forms can be viewed as analogous to the variation in speech that is caused by use of reduced forms. Alternatively, these representations might be viewed as analogous to the out of vocabulary words that plague current speech recognizers.

3.3 Register and dialect differences

Chat interaction is subject to the same kinds of register and dialect variation that occurs in speech. For example (2a) employs a form of the standard Turkish greeting *esselamu aleyküm* that is closer to the original Arabic because it uses the Arabic definite article *es-*, though the umlaut is not Arabic. In contrast the variant in (3a),

selamun aleykum, employs the Turkish suffix on *selam*, but omits the umlaut. (8) illustrates other variants that occurred in a sample of chat from the #ankara channel.

(9) a. selamun aleyküm
    b. selamýn aleykum
    c. Selammmmmmmmm
    d. selamlar
    e. selam all

Another example is the variable use of *ne* in French constructions with negative polarity. Though formal French uses *ne* before the verb and *pas* after the verb for sentence negation, most varieties omit the *ne* in everyday contexts, as observed in (8b), where the absence of both *ne* and the *s* on pas creates serious problems for any translation engine that expects negation to appear as *ne* and *pas* in French. This variation combines with a creative spelling based on reduction to produce examples like (10).

(10) shuis pa interessé "I am not interested"

The standard form of (10) is *je ne suis pas interessé*, but in casual speech, *ne* is dropped, the vowel in *je* is omitted and the two adjacent fricatives merge to produce the sound that is typically spelled *sh* in English (though it is usually spelled *ch* in French).

3.4 Emotives and repeated letters

Two challenges for speech recognition are non-lexical sounds such as laughter or grunts and the distortions of pronunciation that are caused by emphasis, fatigue, or emotions such as anger and boredom. These complications have analogies in written interaction when participants attempt to render the same sounds orthographically, producing forms like those in (9c) and (10).

(10) a. merhabaaaaaaaaaaaaaaaaaaaaaaaaa
    b. ewww
    c. eeeeeeeeeeeeeeeeeeeeeeeeeeee
    d. hehehe

In (10a) the final syllable of the Turkish greeting *merhaba* is lengthened in the same way that it would be in an enthusiastic and expansive greeting, and (10b) effectively communicates a typical expression of disgust. Laughter is rendered in a variety of ways including *ha ha*, *heh heh,* and (10d). The variability of spellings in these cases resembles the variability of nonverbal sounds in speech.

3.5 Emoticons

Another way that chat participants express emotion is by using emoticons and messages that consist entirely of punctuation, as in (11).

(11) a. hey Pipes` >:) how u doing?
    b. o)))*******
    c. !!!!
    d. ????

Like the emotives and repeated letters described in 3.4, these can be viewed as analogous to the non-lexical sounds that occur in speech. However, they are probably more easily identified because they are drawn from a very limited set of symbols.

4 Sharing solutions to shared problems

Because machine translations of spoken and written interaction share so many challenges, it is likely that solutions to the problems might also be shared in ways that will allow research on the newer phenomenon of written interaction to benefit from the years of experience with spoken interaction. Conversely, approaches to written interaction, not biased by previous efforts, can provide fresh perspectives on familiar problems. We present some examples in which there appears to be strong potential for this kind of mutual benefit, drawing on our efforts to improve the performance of TrIM, MITRE's Translingual Instant Messenger prototype. TrIM is an instant messaging environment in which participants are able to interact by reading and typing in their own preferred languages. The system translates each user's messages into the language of the other participants and displays both the source language and target language versions.

TrIM's translation services are provided by the CyberTrans system, which provides a common interface for various commercial text translation systems and several types of text documents (e.g. e-mail, web, FrameMaker). It incorporates text normalization tools that can improve the quality of the input text and thus the resultant translation. Specifically, preprocessing systems provide special handling for punctuation and normalize spelling, such as adding diacritics that have been omitted.

4.1 Spelling and recognition problems

Closer consideration of the problems created by nonstandard spellings reveals the strong similarities between the complexity of speech recognition and recognition of written words in "noisy" communication environments such as chat. In both cases, there is a need to discriminate between words that are not recognized because they are not in the system and words that are in the system, but are not recognized for other reasons, such as variation in phonetic form or a problem in the recognition process. Two properties of chat interaction make this problem as serious for identifying written words as it is for spoken input. First, though a much larger vocabulary can be maintained in digital memory than in the models of speech recognition systems, the creativity and innovation that is valued in chat environments provides a constant source of new vocabulary: it is guaranteed that there will always be out of vocabulary (OOV) words. Second, the high frequency of intentional and unintentional departures from standard spelling matches the variability of speech and makes it essential that the system be able to normalize spellings so that messages are not obscured by large numbers of unidentified words.

A variety of methods have been proposed in the speech recognition literature for detecting OOV words. Fetter (1998) reviews four approaches to the problem and observes that they can be classified in two broad groups: *explicit* acoustic and language models of OOV words "compete against models of in-vocabulary words during a word-based search…*Implicit* models use information derived from other recognition parameters to compute the likelihood of OOV-words" (Fetter, 1998: 104). In the latter group, he classifies approaches that use confidence measures, online garbage modeling in keyword spotting, and the use of an additional phoneme recognizer running in parallel to a word recognizer. These approaches might be adapted to the problem of discriminating misspelled and OOV words in chat interaction, just as approaches to spelling correction might provide alternative solutions to the analogous problem in speech recognition. For example, models of OOV words might compete with models of in-vocabulary recognition errors using Brill and Moore's (2000) error model for noisy channel spelling correction that takes into account the probabilities of errors occurring in specific positions in the word. By modeling recognition errors, the model captures the stochastic properties of both the language and the individual recognition system.

Because our goal is not only recognizing, but also translating messages, we are especially interested in solutions that will facilitate the translation system and process. Consequently, solutions based on modeling the contexts of OOV word use and the contexts of nonstandard spellings seem most promising. For example, it would be worth exploring whether the templates used in example-based translation could be used to model these contexts.

4.2 Preprocessing for special cases

Seligman (2000) observes that current spoken language translation systems use very different methods for recognizing phones, words, and syntactic structures, and he envisions systems in which these processes are integrated, proposing alternatives that range from architectures which support a common central data structure to grammars whose terminal symbols are phones. The latter approach appears to be too narrow because it precludes the possibility of employing preprocessing operations that structure input to facilitate translation.

The success of TrIm and CyberTrans suggests that preprocessing operations offer useful approaches to the challenges we have identified. For example, a preprocessing system in CyberTrans identifies words which are likely to be missing diacritic symbols and inserts them before the input is sent to the translation engines. As a result, the chat message in (12a) is correctly translated as (12b) rather than (12c) or (12d), which are the results from two systems that did not benefit from preprocessing.

(12) a. et la ca va mieux
     b. and there that is better
     c. and Ca is better
     d. and the ca goes better

The French form *la* is the feminine definite article, whereas the form *là* is the demonstrative deictic "there." CyberTrans recognized that the form should be *là* and that *ca* should be *ça*.

Another example concerns the problems of forms such as discourse markers, vocatives, and greetings. (13) shows that when the discourse marker *well* is separated by a comma, as in

(13a), it is correctly translated as a discourse marker in (13b), whereas without the comma in (13c), the translation uses *puits*, the word that would be used if referring to a water well.

(13) a. Well, I'm feeling great!
   b. Bien, je me sens grand!
   c. well aren't we happy today?
   d. puits ne sommes-nous pas heureux aujourd'hui?

Therefore, the comma served as a signal to the system to translate the discourse marker differently, and clearly, a preprocessing operation that identifies and tags items like discourse markers can facilitate their translation.

4.3 Segmentation of translation units

Though clause or sentence units are not clearly marked in speech, the grammars on which analysis and translation rely typically operate with the clause or sentence as the primary unit of analysis. Consequently, the issue of segmenting speech into sentence-like units has received considerable attention in the speech translation community, especially the possibility that prosodic information can be used to identify appropriate boundaries (Kompe, 1997; Shriberg et al., 2000). Other efforts identify lexical items with high probabilities of occurring at sentence boundaries and incorporate probabilistic models, taking into account acoustic features (Lavie et al. 1996) or part of speech information (Zechner, 2001).

In contrast, some researchers have proposed that identifying sentence boundaries is less important than finding appropriate packaging sizes for translation. For example, Seligman (2000) reports that pause units divided a corpus into units that were smaller than sentence units, but could be parsed and yielded understandable translations when translated by hand. Popowich et al. (2000) deliberately segment closed caption input into small packages for translation, claiming that it provides a basis for handling the vocatives, false starts, tag questions, and other non-canonical structures encountered in captions. They also claim that the procedure reduces the scope of translation failures and constrains the generation problem. Since much interaction is already fragmented, processing that relies on smaller units rather than sentences seems worth investigating.

A related possibility is that much of the problematic input is already packaged in small units of short turns or messages. Yu et al. (2000) report that 54% of turns with a duration of 0.7 seconds or less in their data consisted of either *yeah* or *mmhmm* and about 70% of the turns contained the same 40 words or phrasal expressions. They took advantage of these facts by building a language model tailored specifically for short turns.

In a pilot study, we examined a small sample of chat data in order to determine whether short messages were more likely to contain the problematic features described in sections 2 and 3. Of 76 chat messages, 46 or 61% were 3 units or less, where units were delineated by space and punctuation (without separation of contractions) and emoticons counted as a unit. The frequency of items such as greetings, vocatives and acronyms was counted for each message size. Of 8 greetings in the corpus, 7 occurred in messages of 3 or fewer words, and all 9 greeting-plus-vocative structures occurred in messages of 3 or fewer words. Messages with 3 or fewer words also contained 11 of the 14 emotives in the corpus (including emoticons), 8 of the 10 messages with repeated letters, 3 of the 3 acronyms, 2 of the 3 discourse markers, and both of the interjections in the corpus. These results support the claim that much of the problematic usage in chat interaction is limited to short turns that can be identified and processed separately.

5 Conclusions

This paper demonstrates that many of the problems which complicate translation of spoken interaction are shared by written interaction such as chat and instant messaging. It is proposed that solutions developed to solve similar problems in the two communication environments can be profitably shared, and several examples are presented where mutually beneficial approaches might be developed. Specifically, we noted that the problem of discriminating unrecognized OOV words from in-vocabulary words in spoken interaction is analogous to the problem of discriminating unrecognized OOV words from misspelled words in written interaction. We suggested that some of the same methods used in spelling correction might be adapted to speech recognition, especially language models that incorporate probabilities of errors in specific

positions in the word. We also observed the potential of preprocessing operations that structure input for translation systems to allow special treatment of problematic language, including the possibility that much complexity can be avoided by processing and translating smaller units separately. We look forward to exploring these possibilities in future work.

References

Brill, Eric and Moore, Robert C. 2000. An improved error model for noisy channel spelling correction. *Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics*.

Chafe, Wallace. 1982. Integration and involvement in speaking, writing, and oral literature. In *Spoken and Written Language: Exploring Orality and Literacy,* Deborah Tannen (Ed.), Norwoord, NJ: Ablex, pp. 35-53.

Condon, Sherri and Cech, Claude. (Forthcoming) Discourse management in three modalities. In *Computer-Mediated Conversation,* Susan Herring (Ed.), Hampton Press.

Danet, Brenda, Ruedenberg-Wright, Lucia, and Rosenbaum-Tamari, Yehudit. 1995. Hmmm... Where's that smoke coming from?" Writing, Play and Performance on Internet Relay. *Journal of Computer-Mediated Communication,* 1 (2).

Ferrara, Kathleen, Brunner, Hans, and Whittemore, Greg. 1991. Interactive written discourse as an emergent register. *Written Communication*, 8 (1), 8-34.

Fetter, Pablo. 1998. Detection and transcription of OOV words. Verbmobil Technical Report 231.

Flanagan, Mary. 1996. Two years online: Experiences, challenges and trends. *Expanding MT Horizons: Proceedings of the Second Conference of the Association for Machine translation in the Americas*, 2-5 October, pp. 192-197.

Kompe, Ralf. 1997. *Prosody in Speech Understanding Systems*. Berlin: Springer.

Lavie, Alon, Gates, Donna, Coccaro, Noah and Levin, Lori. 1996. Input segmentation of spontaneous speech in Janus: A speech-to-speech translation system. *Proceedings of the ECAI 96*, Budapest, Hungary.

Papineni, K., Roukos, S., Ward, T., Henderson, J., and Reeder, Florence. 2002. Corpus-Based comprehensive and diagnostic MT evaluation : Initial Arabic, Chinese, French, and Spanish results. *Proceedings of the Human Language Technology Conference*. San Diego, California.

Popowich, Fred, McFetridge, Paul, Turcato, Davide, and Toole, Janine. 2000. Machine translation of closed captions. *Machine Translation*, 15, 311-341.

Shriberg, Elizabeth, Stolcke, Andreas, Hakkani-Tur, Dilek, and Tur, Gokhan. 2000. Prosody-based automatic segmentation of speech into sentences and topics. *Speech Communication* 32(1-2).

Seligman, Mark. 2000. Nine issues in speech translation. *Machine Translation*, 15, 149-185.

Sotillo, Susana M. 2000. Discourse functions and syntactic complexity in synchronous and asynchronous communication. *Language Learning & Technology,* 4 (1), pp. 82-119.

Vanni, Michelle. and Miller, Keith. 2002. Scaling the ISLE framework: Use of existing corpus resources for validation of MT evaluation metrics across languages. In Proceedings of LREC 2002. Las Plamas, Canary Islands, Spain.

Wiebe, Janice, Farwell, David, Villa, Daniel, Chen, J-L, Sinclaire, R., Sandgren, Thorsten, Stein, G., Zarazua, David, and Ohara, Tom. 1995. ARTWORK: Discourse Processing in Machine Translation of Dialog. Final Report Year 1. New Mexico State University: Computing Research Lab. http://crl.NMSU.Edu/Research/ Projects/artwork/index.html

Yates, Simeon. 1996. Oral and written linguistic aspects of computer conferencing: A corpus based study. In *Computer-Mediated Communication: Linguistic, Social and Cross-Cultural Perspectives.,* Susan Herring (Ed.), Philadelphia: John Benjamins, pp. 29-46.

Yu, Hua, Tomokiyo, Takashi, Wang, Zhirong, and Waibel, Alex. 2000. New developments in automatic meeting transcription. International Conference on Speech and Language Processing, Beijing, China.

Zechner, Klaus. 2001. Automatic Generation of Concise Summaries of Spoken Dialogues in Unrestricted Domains. *Proceedings of the 24th ACM-SIGIR International Conference on Research and Development in Information Retrieval*, New Orleans, Louisiana.