Towards Translating Spoken Language Pragmatics in an Analogical Framework

Keiko Horiguchi
Department of Language Engineering, UMIST
& D-21 Laboratory, Sony Corporation
6-7-35 Kitashinagawa
Shinagawa-ku, Tokyo 141
Japan
keiko@pdp.crl.sony.co.jp

Abstract

This paper argues that stylistically and pragmatically high-quality spoken language translation requires the transfer of pragmatic information at an abstract level of "utterance strategies". A new categorization of spoken language phenomena into essentially non-meaningful "speech errors", and purposeful "natural speech properties" is introduced, and the manner in which natural speech properties convey pragmatic information is described. Finally, an extension of the analogical speech translation approach is proposed that accounts for such higher-level pragmatic information.

1 Introduction

Traditional grammar, at its origin highly prescriptive, was aimed at written sentences, and completely ignored all characteristics of spoken language. This distinction was codified in Chomsky's competence-performance distinction (Chomsky, 1965). Chomsky singled out the abstract notion of grammatical competence, the rules or constraints characterize grammatical sentences, as the proper subject for the study of language. All other characteristics of language were relegated to the category of language performance, essentially meaningless by-products of the system that happens to implement language production in humans. Traditional approaches to computational linguistics have also focused on the rules of grammar.

1.1 Spoken Language

Spoken language, however, has many characteristics that are different from written language. When an utterance is produced on-line, the speaker does not have a lot of time to think and plan the entire utterance. For this reason, spoken utterances tend to be relatively short, have less complex structure, and contain more fixed or semi-fixed expressions than written sentences. At the same time, the on-line nature of spoken language also gives rise to so-called "disfluencies". Furthermore, written communication usually aims at an accurate transfer of information, in an interactive mode of verbal communication, each utterance often carries larger portion of pragmatic information such as a variety of illocutionary and perlocutionary forces.

1.2 Pragmatic Information

Accurate handling of pragmatic information in speech translation is gaining importance as speech recognition technology improves. As can be observed in communications between a native speaker and an intermediate/advanced second-language learner, pragmatic inappropriateness in otherwise perfectly grammatical utterances causes more communicational damage than purely syntactic mistakes, since the listener tends to interpret it as intentional or malicious, instead of viewing it as a result of the speaker's linguistic incompetence.

2 Previous Approaches: Recognizing Speech Act Types

For the reasons outlined above, it is important to handle pragmatic information in spoken language translation. The most studied area in pragmatics has been the illocutionary force of utterances. This type of information has been shown to be useful for reducing ambiguities and improving the accuracy of speech recognition and translation in many systems (Woszczyna and Waibel, 1994), (Nagata, 1992), (Qu et al., 1996).

2.1 Rule-based Approaches

One of the traditional approaches to this area of pragmatics is to recognize speech act types compositionally using syntactic and semantic rules plus

a few pragmatic principles, such as felicity conditions for each speech act type. Spoken language expressions, however, tend to deviate from conventional grammars, and a system consisting of layers of rule-based modules is often too brittle to handle naturally-occurring spoken input. Furthermore, there are a number of fully- or semi-lexicalized morpheme sequences that carry specific illocutionary forces but that are not totally predictable from its forms. These sequences have an institutionalized function in the particular community, and are best accounted holistically rather than analytically (Pawley and Syder, 1983).

2.2 Pattern Matching

Many spoken language systems have thus been using robust pattern-matching techniques to overcome these problems. They use detailed, task-specific templates and semantic grammars, which can recognize various fixed phrases to mark speech act types while skipping over disfluencies in the input. This method has been shown to be successful in many dialogue systems (Jackson et al., 1991), (Ward, 1991).

3 Other Pragmatic Information

When people engage in face-to-face dialogues, the focus is usually on establishing and maintaining a good relationship among the interlocutors, rather than mere transfer of information. Each spoken utterance thus usually carries a large portion of what (Traugott, 1982) calls the *expressive* component, which expresses the speaker's attitude toward the proposition, toward the interlocutor, and toward the speech situation.

3.1 Ignoring Expressive Information

When spoken language understanding is performed in a goal-oriented dialogue system, it is usually acceptable to strip off any "extraneous" information in order to map the speaker's intention onto an unambiguous system command. This is not possible, however, in a spoken language translation system that acts as a human-human verbal communication aid, where the expressive information encoded in utterances plays a far bigger role. For example, if we consider a conversation between two persons who meet for the first time at a party and extract only propositionally meaningful chunks and translate them, the result will resemble an interrogation rather than a pleasant conversation.

3.2 Translating Pragmatic Information

In our work, we take the view that important pragmatic information is actually encoded in many of the characteristics of spoken language that have been viewed as defective or ill-formed. We believe that many of such characteristics carry specific communicative functions that must be preserved in order to obtain translations with high stylistic and pragmatic accuracy. We propose that the spoken language phenomena that have been labeled as "disfluencies" or "ill-formedness" be divided into two categories: those that serve a communicative function, and those that are non-communicative by-products of the speech production process.

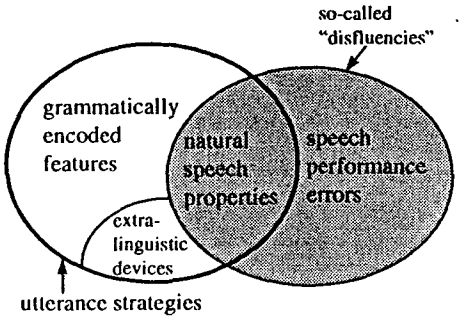

Figure 1: Communication vs. Performance

4 Communication and Performance

We refer to the spoken language phenomena that are non-communicative by-products of the speech production process as "speech performance errors", and to the phenomena that serve a communicative function as "natural speech properties".

4.1 Speech Performance Errors

Speech performance errors are obvious errors not intended by the speaker and (for our purposes) not bearing any information. This includes errors in pronunciation, word selection, and structure selection. When speech performance errors are corrected by the speaker within the utterance, they result in *slip-of-the-tongue repairs.*

4.2 Natural Speech Properties

In contrast to speech performance errors, natural speech properties are produced intentionally by the speaker, and usually carry specific pragmatic, communicative functions. For example, inverted word order and repetitions usually emphasize certain parts of the utterance. Incomplete sentences are often used to soften so-called "face threatening acts" (Brown and Levinson, 1987), speech acts that might have negative effects on the listener, such as rejections or requests.

4.3 Speech Repairs

Some types of repairs, in which a phrase is paraphrased or repeated with more information, also ful-

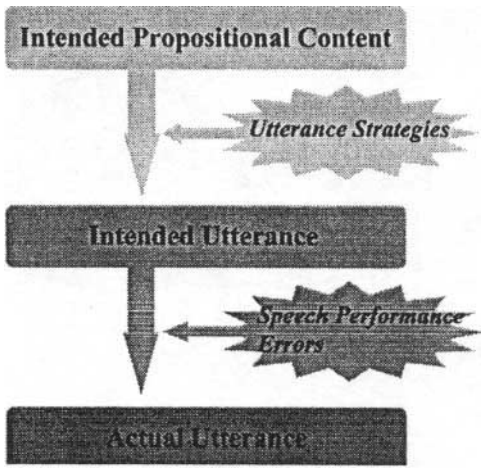

Figure 2: Two-level Distortion Model of Spoken Language

fill communicative functions. We call such repairs *elaborating repairs*, in contrast to *slip-of-the-tongue repairs* which result from correcting speech performance errors. Elaborating repairs like the example below may signal the listener about the status of the speaker's internal processing, or reduce the face-threatening effect of the utterance.

> I gave it to you on Monday, yeah, probably on Monday the 27th.

Unlike a slip-of-the-tongue repair, where the speaker would have deleted the original phrase had there be time and means, in an elaborating repair, deleting the phrase (*on Monday* in the example above) would result in a different effect on the dialogue.

4.4 Utterance Strategies

We refer to all the devices that serves pragmatic or communicative functions as "utterance strategies". They concern the speaker's intention of "how to say" an utterance, as opposed to "what to say" (propositional content) of the utterance. For this reason, their semantics are non-truth-conditional. Utterance strategies range from grammatically-encoded information to extra-linguistic devices such as facial expressions and body language, to natural speech properties (Figure 1).

4.5 Distortion Model

Based on this discussion, we arrive at the two-level model of spoken language phenomena shown in Figure 2. The speaker's intention of "what to say" (intended propositional content) is combined with the speaker's intention of "how to say it" (pragmatic utterance strategies) to form the "intended utterance", which contains natural speech properties. When the speaker actually produces the utterance, speech performance errors might occur, resulting in the "actual utterance" that is to be interpreted by the listener. We believe that spoken language translation systems need to be able to translate what is conveyed through natural speech properties in order to fully convey the speaker's intentions in verbal communication.

5 Communicating Pragmatic Information

This section discusses the different types of pragmatic information that play a role in spoken dialogues.

5.1 Discourse Structure

One type of pragmatic information relates to signaling discourse structure. This concerns how the propositional content of the current utterance is related to what the conversational participants already know, and to the structure of the discourse. It indicates the *theme* and the *rheme* of the utterance, and places new or contrasting information into *focus*. It may signal a new topic, or the return of an old topic, the speaker's attempt to hold the conversational floor. This type of pragmatic information has been noted to be very important for automatically synthesizing utterance with appropriate intonation (Prevost, 1996), and for generating sentences with appropriate word order in free word order languages such as Turkish (Hoffman, 1996).

Natural speech properties that carry the information related to discourse include word order variations (inversions, right or left-dislocations), and filled pauses and hedges used as a floor-holding device to signal the listener not to take her turn. Some interjections and hedges can be used to help the listener prepare herself for the subsequent information, or aid the listener's processing and comprehension of the current utterance.

5.2 Interpersonal Intention: Politeness

The majority of utterance strategies express the speaker's *interpersonal intentions*, the main aspect being politeness. There are two different types of politeness expressed in spoken language, "discernment" and "volition" (Ide, 1989). Discernment refers to the speaker's recognition of her relationship with the addressee and the situation. This is mainly expressed through the speaker's choice of conversational topic, lexical items, and syntactic structures. For example, the same conversational participants may use different linguistic forms depending on the speech situation, such as a discussion during a formal meeting versus an informal hallway-chat after the meeting. The choice of formal and informal predicate forms in Japanese and the choice of distant and

familiar second person pronouns in French and German are examples of lexically encoded discernment markers.

The volitional aspect of politeness is usually expressed through projection of "face". There are two distinct aspects of face, "positive face" and "negative face" in the theory of (Brown and Levinson, 1987), which are rephrased as involvement and independence, respectively, in (Scollon and Scollon, 1995). Positive face or involvement concerns one's desire to be liked by others, to be involved with others, and to be part of the same group. Negative face or independence, on the other hand, concerns one's desire to maintain privacy and independence, and to avoid the imposition or dominance of others.

5.3 Positive Face

Strategies to project positive face, which are called "positive politeness strategies"(Brown and Levinson, 1987) or "solidarity politeness" (Scollon and Scollon, 1995), can be carried out by the use of intensifiers accompanying positively affecting speech acts such as thanking and complimenting. The use of the first-person plural pronoun *we* in English is also an example of solidarity politeness. Most of the other linguistic items used in solidarity politeness strategies, however, do not bear propositional content. For example, the speaker may try to appeal to mutual beliefs or affective common ground by using the English interjection *you know* and the Japanese sentence-final particle *ne* (Cook, 1988). The speaker may also try to invite the addressee's involvement by using hearer-oriented question tags such as *right*, *all right*, *okay*, *would you* or *will you*, or by using devices to attract the addressee's attention such as *look*, *listen*, *hey* and informal or affectionate address terms.

Telegraphic utterances that omit obvious information can be interpreted as a strategy to emphasize common knowledge among the interlocutors. Involvement strategies "assert the speaker's right to advance his or her own position on the grounds that the listener will be equally interested in that position and in advancing his or her own position (Scollon and Scollon, 1995):85". The speaker may achieve this by displaying an assertive, "non-challengeable attitude" (Kawanishi, 1994) with the Japanese *janai* form, or by aligning herself and the listener on the same side by using distant demonstrative *ano (that)* (Cook, 1993).

5.4 Negative Face

Strategies to project negative face, which are called "negative politeness" or "deference politeness", are mainly carried out by the use of "toning down" devices accompanying negatively affecting speech acts such as criticizing, giving advice, requesting, or refusing an offer or request. Expressions of the speaker's hesitation or tentativeness, such as hedges (*well*, *I don't know*, *I think*, *I am wondering if...*), use of the interrogative form, or the past tense *I was wondering if...* or the subjunctive mood *it would be better* are examples of such devices to soften the force of the utterance, and to make it easier for the addressee to refuse. Sometimes even questions to ask permissions to ask a question are used to give the addressees ways to answer negatively without directly refusing the request, as shown in the following example given (Yule, 1996), pp.64-65:

> I know you're busy, but might I ask you if–em–if you happen to have an extra pen that I could, you know–eh–maybe borrow?

There are also content "downtoners" such as *little*, *a bit*, *just*, ... *and so on*, and the use of colloquial expressions (such as *to give a hand* instead of *to help*), which trivialize the action mentioned in the utterance. The speaker may also try to create distance between the addressee by avoiding reference to both the speaker and the addressee, as in an agentless passive sentence *I would like a reservation to be made.*

5.5 Expressing Attitude

Another type of utterance strategy expresses the speaker's attitude towards the propositional content in the utterance. This information can be conveyed through various forms of evidential markers and devices to express the speaker's certainty/uncertainty, or the speaker's perspective.

6 Handling Pragmatic Information in Speech Translation

In the context of spoken language translation, the crucial characteristic of pragmatic utterance strategies is that the surface forms in which they are realized are often different across languages.

6.1 Example: English vs. Japanese Politeness

For example, softening the effect of an imperative force by questioning the addressee's ability to perform the action (*Can you do X for me?*) or asserting the speaker's desire (*I would like you to do X for me*) can be found across many languages. However, strategies to further reduce the imposing effect of these request forms are usually not directly transferable across languages. In English, a more polite way of phrasing the request *Can you do X for me?* would be the use of the subjunctive mood, *Could you do X for me?*, but no corresponding form exists in Japanese. Instead, the Japanese speaker may use the negative form *X shite itadakemasen ka?*. If this Japanese form is translated literally into English, the result would be *Can't you do X for me?*, which has a quite different pragmatic meaning, and definitely does not convey the same degree of politeness as the Japanese expression.

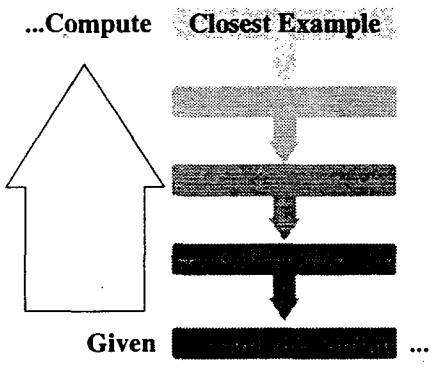

Figure 3: Computing an Interpretation

6.2 Abstract Pragmatic Transfer

To what extent word order can be altered, and how easily known information can be elided, are also largely dependent on the syntax of each language. One strategy that is realized as fronting may be marked as intonation in another language. It is not sufficient, therefore, to recognize the surface form of each pragmatic strategy and directly transfer it to the same surface form in another language. Spoken language systems are thus required to transfer pragmatic utterance strategies at a more abstract level and to be able to recognize and generate appropriate surface forms in each language, in order to achieve high-quality translation. In our approach, we treat pragmatic strategies as additional information that are superimposed upon basic propositional content, try to recognize and extract them, and transfer them to the appropriate target language expressions.

7 An Analogical Framework for Translating Pragmatics

This section gives a brief overview of our approach to translating spoken language.

7.1 The Role of Lexicalization

Our approach to translating spoken utterances resonates well with the insights of (Pawley and Syder, 1983) about native speakers' competence and knowledge. According to Pawley and Syder, a native speaker has a number of fully- or semi-lexicalized morpheme sequences in her long-term memory, in addition to a set of productive syntactic rules. When people engage in a conversation, there are a number of cognitively intensive tasks that they have to perform other than encording and decoding internal structure of each utterance, such as planning a larger unit of discourse, planning and interpreting perlocutionary effects, and paying attention to the surroundings. The use of pre-established expressions helps both the speaker and the addressee, since such expressions can be easily and quickly retrieved from their long-term memory, and little encoding and decoding work is required. As Pawley and Syder note, these memorized sequences have varying degrees of lexicalization. While some are completely fixed expressions, most others are "stems" that can be inflected, expanded or transformed to some extent.

7.2 Analogical Translation

This model of a native speaker's linguistic competence fits very well with the analogical framework of translation (Nagao, 1984), (Jones, 1996). In the analogical framework, the translation system is equipped with a large database of pre-translated example pairs, in which the best example that matches the input expression is selected and used for generating an appropriate target language expression. For translating spoken language, an analogical system should have various sentence stems and patterns along with their corresponding translation in its example database. In this framework, the task of an the spoken language translation system can be seen as follows: given the speech recognizer output, the system must recover the closest example available in the example database (Figure 3).

7.3 A Model of Speech Production

There are a number of factors that need to be considered in trying to select the most appropriate example in the database for the given input. Based on four distinct factors that we have identified, we propose a model of spoken language production that we call the "cascaded noisy channel model" (Figure 4). In this model, the speaker first selects an example E that is closest to the core of the message that she intends to express. Then, the speaker modifies patterns by replacing subconstituents, by expanding it with modifiers, and by transforming it into different syntactic constructions (for example, transforming it from the declarative mood 4to the interrogative mood, or from the active voice to the passive voice). This process yields the "intended propositional content". Next, depending on the speech situation and discourse context, the speaker applies certain pragmatic utterance strategies. This results in the "intended utterance", which is characterized by natural speech properties such as ellipsis, inverted word order, or interjections.

When the speaker actually vocalizes the utterance, speech performance errors may occur. The result of this is the "actual utterance" that is presented to the listener. The speech recognition program converts the speech signal to a string of word hypotheses, possibly introducing additional errors and distortions, which results in the "recognizer output".

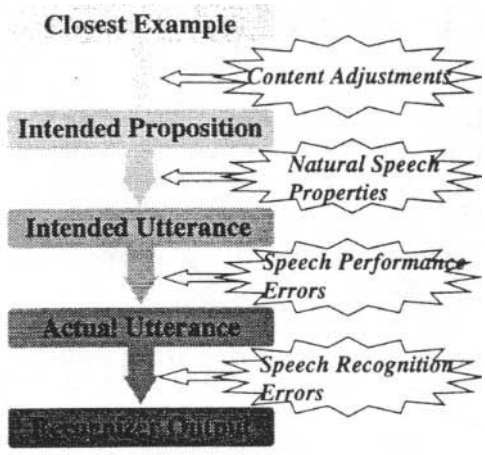

Figure 4: Cascaded Noisy Channel Model of Spoken Language

Thus, the speech recognizer output, which represents the input to the translation engine, has traversed four distinct channels or distortion processes, each of which is associated with different causes and effects on the message. Previous research has shown that speech recognizer errors can be modeled, and corrected, in such a framework (Ringger and Allen, 1996). In our work, we extend this model to cover a sequence of separate sources of distortions.

7.4 A Hybrid Analogical Method for Speech Translation

We have incorporated into the analogical translation method a shallow syntactic analysis module that identifies clause and phrase boundaries and that converts some variations into lexical and syntactic features. Both input and example expressions are matched after shallow syntactic analysis. Analogical matching and transfer is applied recursively to the input syntactic tree. By applying the recursive analogical transfer process from larger linguistic constituents to subconstituents, the system can handle various degree of lexicalization in the input language in an efficient manner.

In our work, the distortion processes are modeled using a number of distortion operators that operate on the shallow syntactic tree of the utterance. Given a number of independence assumptions, the most probable example can be computed efficiently with a dynamic programming algorithm. (See (Horiguchi and Franz, 1997) for more details.)

8 An Example

This section shows an example of the manner in which an expression containing a pragmatic "politeness" component is translated from Japanese to English.

8.1 Japanese Input

In the following example, speaker A is explaining an incident in which she was asked a difficult favor, and speaker B is responding, expressing her understanding of A's difficult position.

```
(1)A: sorede tyotto   kangaesasete
      so      a-little think-CAUSE-PASS
      ''so I said 'let me think for a while'''

      hosii-tte  itta     no
      want-QUOTE say-PAST PART

(2)B: nanka muzukasii-yo-ne         souiu no
      HEDGE be-difficult-PART-PART such  thing

      kotowaru no-tte
      reject   thing-TOP
```

The propositional content of speaker B's response is "To reject something like that is difficult," but the utterance also contains a number of natural speech properties that add certain pragmatic elements of meaning.

8.2 Pragmatic Operators

The "intended propositional content" of the above utterance's can be paraphrased as follows:

```
(3) Souiu no-wo     kotowaru no-ga     muzukashii.
    such  thing-OBJ reject   thing-SBJ be-difficult
```

Our flexible matching process is able to map an inverted construction like example input (2) onto its normalized form (3). Then, the following pragmatic operators are found to have been applied to the "intended propositional content":

- inserting *nanka*
  pragmatic strategy: soften the current assertion
  pragmatic effect: deference politeness
- inserting *-yo*
  pragmatic strategy: express the attitude that the speaker's assertion is non-challengeable
  pragmatic effect: solidarity politeness
- inserting *-ne*
  pragmatic strategy: indicate affective common ground
  pragmatic effect: solidarity politeness
- deleting object marker *-wo*
  pragmatic strategy: emphasize shared knowledge
  pragmatic effect: solidarity politeness
- subject-predicate inversion
  pragmatic strategy: point to previously established or implied referent
  pragmatic effect: discourse coherence, solidarity politeness

The last operator, subject-predicate inversion, is usually employed to describe how the subsequent information connects to the previous discourse by preposing the constituent that is implicitly or explicitly related to something in the previous discourse. In the example above, it is used to point to the situation that A has been talking about as something already established or agreed upon to be difficult, and thus can be interpreted as a solidarity politeness operator which reinforces the common ground between the interlocutors.

8.3 Translating the Utterance

This subsection discusses why it is necessary to analyze which pragmatic operators were applied to the input, and to generate the corresponding pragmatic operators in the output, in order to obtain stylistically and pragmatically high-quality translations.

8.4 Pure Analogical Translation

If we employ a pure example-based translation method, most of the pragmatic information cannot be reflected, since it is not feasible for the example database to contain all possible pragmatically marked permutations of the examples. Therefore, in the best case, the following literal translation of sentence (3) might be obtained:

(4) To reject something like that is difficult.

Since sentence (4) is pragmatically neutral, the pragmatic information from the original sentence has been lost.

8.5 Direct Mapping

A direct one-to-one mapping of each pragmatic strategy operator to the target language is not possible, since many of these operators are not directly translatable to other languages. For example, while many languages have hedges similar to *nanka*, and many languages include means to invert subject and predicate, only few languages include deletable case-markers such as *wo* or sentential particles such as *yo* and *ne*. Thus, if we attempt a direct mapping of the pragmatic operators, we might obtain a translation similar to the following:

(5) Sort of difficult, to reject something like that is

This translation is quite awkward, and does not fully reflect the pragmatic meaning of the original sentence.

8.6 Translating Pragmatic Strategies

By analyzing each operator for its pragmatic effect, we can obtain a translation that preserves the speaker's pragmatic intentions:

Well, it's sort of difficult, isn't it, to reject something like that.

In this translation, the deference politeness strategy is transferred to the hedge words *well* and *sort of*, the solidarity politeness strategies are transferred to the tag question *isn't it*, and the subject-predicate inversion is transferred into the extraposition construction.

8.7 Conclusions and Further Work

Our work is motivated by the goal of pragmatically high-quality translation of spoken utterances of the type that may be found in human-to-human spoken dialogues. In order to accurately render the full range of meaning conveyed by such utterances, it is not sufficient to limit attention to syntactic and semantic aspects of spoken expressions.

Based on a number of independent motivations, we have adopted a hybrid analogical approach to the problem of translating spoken language. Briefly, our approach is motivated by the shortcomings that we perceive in other approaches, such as syntactic or semantic-grammar based, interlingua-based, purely analogical, or purely statistical methods. For more detailed arguments, please refer to (Horiguchi and Franz, 1997).

In this paper, we have described our view of spoken language pragmatics, and we have described how pragmatic information can be translated within the hybrid analogical approach. In future work, we will perform corpus analysis for additional pragmatic operators, and extend the prototype implementation of our analogical speech translation system to cover these phenomena.

References

Bateman, John. 1988. Aspects of clause: Politeness in Japanese: An extended inquiry semantics treatment. In *Proceedings of the 26th Annual Meeting of the Association for Computational Linguistics.*

Bates, Madeleine, Robert J. Bobrow, and Ralph M. Weischedel. 1993. Critical challenges for natural language processing. In Madeleine Bates and Ralph M. Weischedel, editors, *Challenges in Natural Language Processing.* Cambridge University Press, Cambridge, pages 3–36.

Bobrow, R., Robert Ingria, and David Stallard. 1990. Syntactic and semantic knowledge in the DELPHI unification grammar. In *Proceedings of the Speech and Natural Language Workshop*, pages 230–236, June.

Brown, Penelope and Stephen Levinson. 1987. *Politeness: Some universals in language usage.* Cambridge University Press, Cambridge, U.K.

Chomsky, Noam. 1965. *Syntactic Structures.* The MIT Press, Cambridge, Massachusetts.

Cook, Haruko Minegishi. 1988. *Sentential particle in Japanese conversations: A study of indexicality.* Ph.D. thesis, University of Southern California.

Cook, Haruko Minegishi. 1993. Functions of the filler *ano* in Japanese. In Soonja Choi, editor, *Japanese/Korean Linguistics Volume 3*, pages 19–38. CSLI, Stanford University, CA.

Epstein, M., K. Papieni, S. Roukos, T. Ward, and S. Della Pietra. 1996. Statistical natural language understanding using hidden clumpings. In *ICASSP-96*, pages 176–179, Atlanta, GA.

Hoffman, Beryl. 1996. Translating into free word order languages. In *Coling-96*, Copenhagen, Denmark.

Horiguchi, Keiko and Alexander Franz. 1997. A formal basis for spoken language translation by analogy. In *Spoken Language Workshop at ACL/EACL-97*, Madrid, Spain.

Ide, Sachiko. 1989. Formal forms and discernment: Two neglected aspects of universals of linguistic politeness. *Multilingua*, 8(2/3):223–248.

Jackson, Eric, Douglas Appelt, John Bear, Robert Moore, and Ann Podlozny. 1991. A template matcher for robust NL interpretation. In *Proceedings of the Speech and Natural Language Workshop*, pages 190–194, February.

Jones, Daniel. 1996. *Analogical Natural Language Processing.* UCL Press, London.

Kawanishi, Yumiko. 1994. An ananlysis of nonchallengeable modals: Korean -canha(yo) and Japanese -janai. In Noriko Akatsuka, editor, *Japanese/Korean Linguistics, Volume 4*, pages 95–112. CSLI, Stanford.

Maruyama, Naoko. 1996. Hanashikotoba no shoso (bt). In *wQN'[gA*, pages 41–58, March.

Mayfield, L., M. Gavalda, W. Ward, and A. Waibel. 1995. Concept-based speech translation. In *ICASSP-95*, pages 97–100, Detroit, MI.

Nagao, Makoto. 1984. A framework of a Machine Translation between Japanese and English by analogy principle. In A. Elithorn and R. Banerji, editors, *Artificial and Human Intelligence*, pages 173–180. North-Holland.

Nagata, Masaaki. 1992. Using pragmatics to rule out recognition errors in cooperative task-oriented dialogues. In *Proceedings of International Conference on Spoken Language Processing (ICSLP-92)*, pages 647–650.

Nakatani, Christine and Julia Hirschberg. 1993. A speech-first model for repair detection and correction. In *Proceedings of the 31st Annual Meeting of the Association for Computational Linguistics*, pages 46–3, Columbus, Ohio.

O'Shaughnessy, Douglas. 1994. Correcting complex false starts in spontaneous speech. In *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, volume I, pages 349–352, April.

Pawley, Andrew and Frances Hodgetts Syder. 1983. Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In Jack C. Richards and Richard W. Schmidt, editors, *Language and Communication*, pages 191–227. Longman.

Prevost, Scott. 1996. An information structural approach to spoken language generation. In *Proceedings of the 34th Annual Meeting of the Association for Computational Linguistics*, pages 46–53, Santa Cruz, CA.

Qu, Yan, Barbara Di Eugenio, Alon Lavie, Lori Levin, and Carolyn P. Rose. 1996. Minimizing cumulative error in discourse context. In *Proceedings of the ECAI*, Budapest.

Ringger, Eric K. and James F. Allen. 1996. A fertility channel model for post-correction of continuous speech recognition. In *Proceedings of International Conference on Spoken Language Processing (ICSLP-96)*, pages 897–900, Philadelphia, PA.

Scollon, Ron and Suzanne Wong Scollon. 1995. *Intercultural Communication: A Discourse Approach.* Blackwell, Oxford UK/Cambridge USA.

Traugott, E. C. 1982. From propositional to textual and expressive meanings: Some semantic-pragmatic aspects of grammaticalization. In W. P. and Y. Malkiel, editors, *Perspectives on historical linguistics.* John Benjamin, Amsterdam/Philadelphia, pages 245–71.

Ward, Wayne. 1991. Understanding spontaneous speech: The PHOENIX system. In *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, pages 365–367, May.

Woszczyna, Monica and Alex Waibel. 1994. Inferring linguistic structure in spoken language. In *Proceedings of International Conference on Spoken Language Processing (ICSLP-94)*, pages 847–850, Yokohama, Japan.

Yamashita, Yoichi, Keiichi Tajima, Yasuo Nomura, and Riichiro Mizoguchi. 1994. Dialog context dependencies of utterances generated from concept representation. In *Proceedings of International Conference on Spoken Language Processing (ICSLP-94)*, pages 971–974, Yokohama, Japan.

Yule, George. 1996. *Pragmatics.* Oxford University Press, Oxford, UK.