

# Some Issues in Intelligent Processing of Linguistic Information

Hiroya Fujisaki, Non-member

## ABSTRACT

In order to achieve a higher and more intelligent level of processing linguistic information than is currently accomplished by conventional technologies, it is believed necessary to obtain deeper understanding of some of the fundamental aspects of language, as well as of the way humans deal with these aspects. The present paper first takes up ambiguity, vagueness, and variability, and presents some definitions and formulations, mainly on the basis of the author's own studies. It then discusses human processes of language comprehension and language acquisition, and proposes to model these processes to be incorporated into machine processing. Finally, new frontiers of research are briefly mentioned.

## 1. INTRODUCTION

Recent progress in technologies for processing both spoken and written languages has been quite remarkable, and has certainly facilitated exchange of linguistic information between humans and machines. Nobody will deny, however, that these technologies are still far from being satisfactory, in the sense that they do not provide an interface that is smooth and robust enough to let the humans feel that they are dealing with an intelligent partner. In other words, humans will be more comfortable if machines are capable of processing linguistic information more intelligently than is done by current technologies. The goal of such 'intelligent' processing would be a machine performance comparable to that of a human being with an average intelligence, though we see no reason why machines should not be more intelligent than humans are. As our society becomes more and more dependent on information processing machines, the human language, especially in its spoken form, will surely be of primary importance in man-machine interface, since language is the quintessential tool for human thought and communication. Thus a 'human-like' intelligence will be a reasonable goal for machines in the age of man-machine symbiosis. Such an ability of intelligent processing can only be achieved by looking into the very nature of the language with its possibilities and limitations as a means

for information transmission, and of the human processes of language use. One of the main reasons for the limited success of our conventional technologies is, in the author's opinion, that they have been developed without paying due attention to these points.

It may be appropriate here to state the author's view on the role of language in human communication. Language is a medium for expressing, transmitting, and storing information and knowledge, which constitute the ultimate substance of a message, be it in the form of a written text or a spoken discourse. Thus the goal for language processing is the interconversion between linguistic representations and the underlying knowledge representations. In this respect, studies on knowledge representation in the field of artificial intelligence and knowledge engineering are certainly important, but we would perhaps need more realistic representations based on psychological investigations. Although in most situations the sender of a message intends to incur a unique representation of knowledge on the part of the receiver, the reconstruction usually relies very heavily on the prior knowledge, either conveyed by previously sent messages or being shared by the sender and the receiver even without direct exchange of information. Without full support from such prior knowledge, a message can very well be ambiguous or may even be ununderstandable. Systematic studies on *ambiguity* and of the means of disambiguation therefore constitute one of the important areas of research in intelligent processing of language.

Whereas the information to be conveyed by language may be continuous in nature, language can express and transmit only discrete form of information. Thus there arises a certain kind of indeterminacy, which may be called *vagueness*, between the information being intended by the sender and the information retrieved by the receiver. Quantitative analysis and formulation of such vagueness is another area of research.

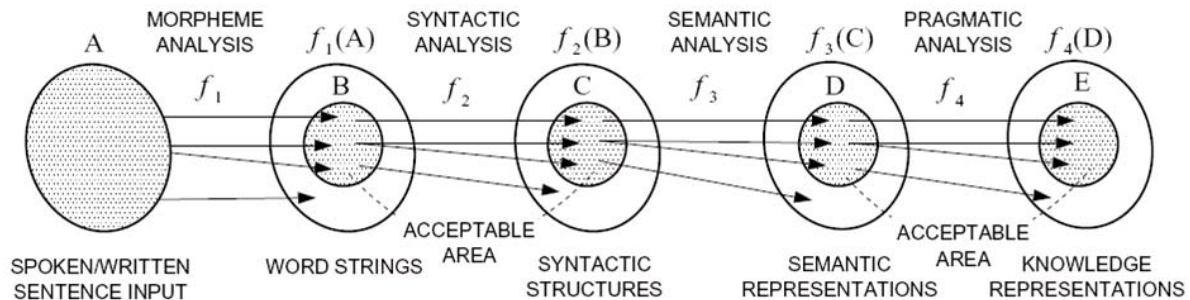
Furthermore, it is generally taken for granted that a language is a well-defined, definite system of codes, and the way people in the same language community use the language is identical. However, both the code system and the way it is used vary from one individual to another. Intelligent processing of language has to cope with this sort of *variability*.

The present paper is meant to present the author's view on these points, drawing heavily on the author's own works. It will first discuss the three aspects of

---

Manuscript received on September 10, 2006.

The author is Professor Emeritus, University of Tokyo, Bunkyo-ku, Tokyo, 113-8656 Japan; E-mail : fujisaki@alum.mit.edu



**Fig.1:** Successive stages of message interpretation, indicating non-unique mapping between representations at different stages.

human language use that have to be taken into account in accomplishing intelligent processing: ambiguity, vagueness, and variability. It will then discuss the characteristics of human processes that the author considers to be crucial for the realization of intelligent processing.

## 2. AMBIGUITY AND DISAMBIGUATION

### 2.1 Sources of Ambiguity

Ambiguity of a linguistic expression (word, phrase, sentence, or even paragraph) is defined as the state of having more than one meaning. Since, however, it is quite common that the meaning of an expression can be definite only when the context and the background information are given, the judgment on ambiguity is possible only after the context and the assumed background knowledge are specified. Thus an expression may be ambiguous without any context, but may be unambiguous within a certain context. A systematic study has been conducted by the author and his coworker on the possible sources of ambiguity in Japanese and the possibility of disambiguation (Fujisaki and Hoshiai, 1984).

Assuming that the process of interpretation of a message written in Japanese consists of four successive analysis stages, i.e., morpheme analysis, syntactic analysis, semantic analysis and pragmatic analysis, each stage of processing can be considered as a mapping  $f$  of its input domain onto its output domain, as shown in Fig. 1. For example, the first stage of processing, i.e., morpheme analysis, maps the set  $A$  of all possible character strings onto  $f_1(A)$ , which is divided into two sets: one is the set of possible word strings (denoted by  $B$ ), the other is the set of strings containing non-words of Japanese (denoted by  $f_1(A) - B$ ). The arrows indicate the following three possible cases of mapping:

- 1) the input string is mapped uniquely onto a possible word string
- 2) the input string is mapped onto two (or more) possible word strings (i.e., the result is ambiguous)
- 3) the input string is mapped onto a string that contains non-words of Japanese.

The same figure also applies to spoken language inputs, if we assume phonetic symbols as input. Thus the figure shows that various types of ambiguity become explicit as the result of analysis/processing, and also that the possibility of disambiguation exists at a later stage of processing. Figure 2 indicates possible sources of ambiguity at each stage of analysis, and Table 1 lists the possible stages of processing at which these ambiguities may be dissolved (disambiguated).

It is to be noted that some of the sources of ambiguity are language specific. For example, the word-boundary ambiguity is present in written Japanese but not in written English.

### 2.2 Size of Context Necessary for Unique Interpretation

Granting that certain sources of ambiguity can be eliminated by the use of context and/or background knowledge, we would like to know which of the two is more important, and what is the necessary size of context. As a partial answer to this question, a small-scale study was conducted on the size of context necessary to derive intended interpretations for the Japanese demonstrative pronoun "sore" (meaning 'it') and the corresponding demonstrative adjective "sono" (meaning 'its') used in 31 editorials of the newspaper "The Asahi Shimbun" (Fujisaki et al., 1986). The results, summarized in Table 2, will provide a rough idea on the size of necessary context, though the result may certainly vary depending on the style as well as on the specific language.

These results suggest that about 40% of the demonstratives can be correctly interpreted within the context of a sentence, while 50 to 60% of them refer to the context of the immediately preceding sentences, and less than 10% of them require a still larger context, mostly sentences in the same paragraph.

### 2.3 Role of Ellipses and Their Restoration

Assuming that the ultimate goal for linguistic communication is to establish a desired knowledge representation on the part of the receiver, the sender only needs to transmit information just sufficient for the

**Table 1:** Various sources of ambiguity in Japanese sentences and the levels of processing at which their disambiguation may be possible.

LEVELS AT WHICH DISAMBIGUATION MAY BE POSSIBLE  CLASSIFICATION OF AMBIGUITY	MORPHEME ANALYSIS				SYNTACTIC ANALYSIS	SEMANTIC ANALYSIS	PRAGMATIC ANALYSIS
	SPOKEN LANGUAGE		WRITTEN LANGUAGE				
	SEGMENTAL FEATURES	PROSODIC FEATURES	KANA	KANA KANJI			
WORD BOUNDARY AMBIGUITY	○	○	○	○	○	○	○
HOMONYMY		○		○	○	○	○
HOMOGRAPHISM	○	○	○		○	○	○
POLYSEMY					○	○	○
CONSTITUTIONAL HOMONYMY						○	○
REFERENTIAL AMBIGUITY						○	○
AMBIGUITY DUE TO ELLIPSIS						○	○
AMBIGUITY IN THE ROLE OF SURFACE STRUCTURE CASES							○
SCOPE AMBIGUITY OF QUANTIFIERS							○
INTENSION/EXTENSION AMBIGUITY							○

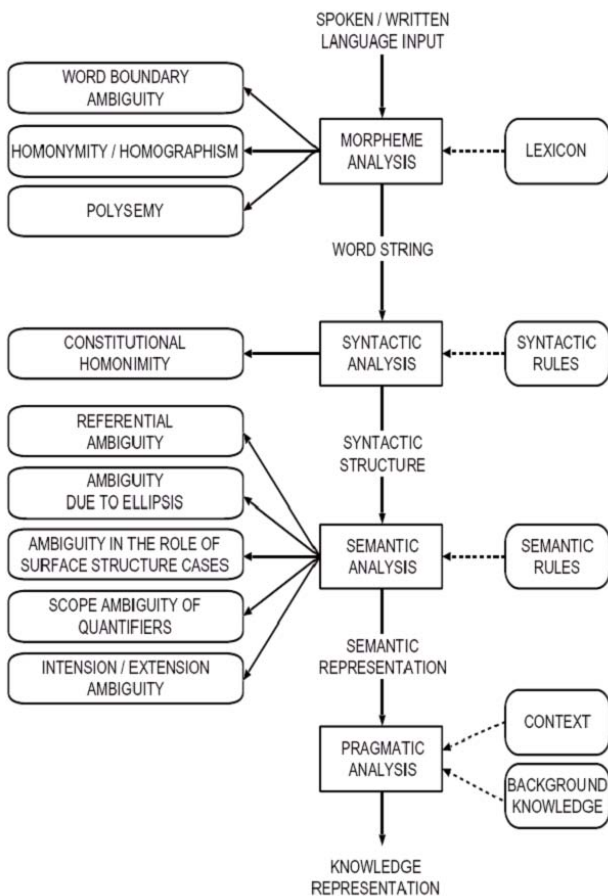
**Table 2:** Size of context necessary to derive intended interpretation for demonstrative pronouns and adjectives of Japanese.

Necessary Context	(a) Same Sentence	(b) Immediately Preceding Sentence	(c) Same Paragraph Except (a)&(b)	(d) Other
Demonstrative Pronoun "sore" ('it')	40%	59%	1%	0%
Demonstrative Adjective "sono" ('its')	37%	50%	8%	5%

purpose. Thus it is quite common that the semantic information derived from the received message is not sufficient by itself, but the receiver can make inference on the missing information and derive the complete knowledge representation. In many situations where the rules of inference are well understood both by the sender and the receiver, it is also quite common to apply systematic rules of simplification (ellipsis) in producing the messages, to avoid redundancy. Figure 3 illustrates the processes of such simplification and restoration by inference involved in the production

and comprehension of linguistic messages.

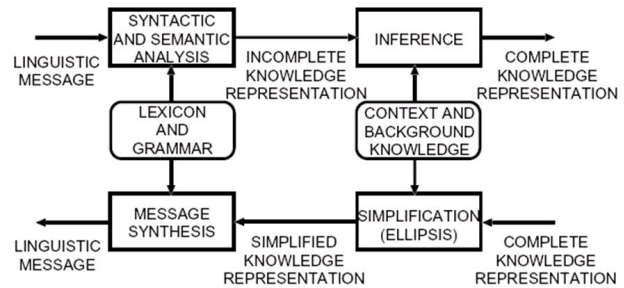
In a study on the interconversion between linguistic messages and knowledge representations for weather forecasts, we have found that such ellipses occur in 56% of all the weather forecast sentences investigated, but the missing information can always be retrieved by empirical rules of inference (Fujisaki and Kameda, 1984). In fact, the weather forecast messages generated without these systematic omissions do sound overly redundant and unnatural.



**Fig.2:** Sources of ambiguity in Japanese arising at various levels of processing for deriving knowledge representations.

### 3. VAGUENESS OF LINGUISTIC EXPRESSIONS

While the ambiguity discussed above refers to the indeterminacy in the choice of a multiplicity of distinctively different interpretations, there exists another kind of indeterminacy in linguistic expressions, viz. vagueness. Vagueness of meaning arises where a discrete linguistic expression is assigned to a continuum of properties or events, as, for example, in the case of color naming. Although such vagueness of correspondence between linguistic expressions and the underlying properties is conventionally treated as fuzziness, very little work seems to have been actually performed on the quantification of the relationship between language and objects. A deeper insight into the nature of vagueness indicates that the indeterminacy of information transmission in this case has two different aspects, one is the indeterminacy of coding on the part of the sender, the other is the indeterminacy of decoding on the part of the receiver. In an earlier work (Fujisaki and Katagiri, 1981), we have demonstrated a method of quantitative measurement of the sender's coding characteristics and the re-



**Fig.3:** Processes of interconversion between linguistic messages and knowledge representations.

ceiver's decoding characteristics. The coding characteristics can be expressed by a set of categorizing curves, i.e., a set of probability distribution functions for the category boundaries. The decoding characteristics can be expressed by a set of probability density functions for the responses which are incurred by each of the linguistic expressions. Figure 4 illustrates examples of such coding and decoding characteristics for one informant of Japanese in using age terms such as "yō-nen" (childhood), "shō-nen" (boyhood), "sē-nen" (youth), "sōnen" (manhood), "rō-nen" (old age).

It was demonstrated that the method can be used to quantify inter-individual differences as well as context effects in the use of these categorizing terms.

The formulation of a functional relationship between a physical variable and a linguistic expression was treated by Zadeh, though quite qualitatively, as an example in his proposal for the concept of "fuzziness" (Zadeh, 1965). The membership function is defined, not as the probability, but as the "grade of membership" that a stimulus is considered to belong to a "fuzzy" set. The present author considers, however, that the membership function proposed by Zadeh is exactly equivalent to the probability distributions in the coding characteristics obtained by observing the sender's behavior. Despite the conceptual distinction made by Zadeh between probability and membership function, the experimental procedure of determining the membership function will be exactly the same as that of determining the coding characteristics described in the present study. It should also be noted that the concept of fuzziness corresponds only to one kind of indeterminacy in language use, i.e., the indeterminacy in regards to coding. There exists another kind of indeterminacy, i.e., indeterminacy in regards to decoding. The formulation of indeterminacy in language use cannot be complete unless both of these factors are taken into account.

### 4. VARIABILITY OF SPEECH SOUNDS OF A LANGUAGE

The acoustic characteristics of a spoken language vary due to various factors. Some factors introduce stochastic and unpredictable variations, while others

cause systematic and predictable differences. The domain of variability may be within a speaker, among speakers of the same social or dialectal group, or among groups of speakers. Table 3 shows a rough classification of these factors based on their characteristics.

The so-called common Japanese, i.e., the language which is now commonly used in Japan, emerged toward the end of the 19th century on the basis of the Tokyo dialect. Although there existed and still exist a number of dialects in Japan, the common Japanese was spread rather rapidly over the entire country through educational and legislative endeavors, as well as by the rapid and widespread use of radio and television broadcasts. Thus at present essentially nobody has difficulty in understanding speech of the common Japanese. When it comes to speaking, however, there exist considerable regional differences which is to be differentiated from dialectal differences. Since we aimed at investigating the variability of the common Japanese, the speech material was collected in Tokyo and eight other major cities of Japan with a large population, high cultural activities, and rather close connections with Tokyo (Fujisaki et al., 1985). These cities are distributed over the entire country as shown in Fig. 5. In order to avoid variability due to age and gender, the informants were restricted to male adults, between 20 and 65 years of age, who were born and grew up in the respective regions. Their parents were also natives of the respective regions in most cases. The number of informants was eight to ten for each region.

The speech material consisted utterances of five Japanese vowels uttered in isolation. The informants read a randomized list of these vowels and recorded six to eight tokens of each vowel. Some disyllabic and polysyllabic words were also recorded and analyzed to investigate variability due to context, but here we will examine only the variability of isolated vowels. The recorded sounds were digitized at 10 kHz with an accuracy of 12 bits per sample for further computer processing. The procedure for formant extraction is based on Analysis-by-Synthesis of the spectral envelope (Fujisaki et al., 1970).

Figure 6 illustrates the extent of intra-speaker variability of F1 and F2 of a vowel in the case of a male adult informant from Tokyo. Each point indicates an utterance, and each ellipse indicates the vowel area within which an utterance of this particular speaker is expected to fall at a probability of 90% under the assumption of two-dimensional normal distribution. As an index for the magnitude of intra-speaker variations, we adopt here the diameter of a circle in the log  $F1$  - log  $F2$  plane whose area is identical to that of the ellipse. Table 4 shows the mean values of the index for each vowel averaged over 10 informants from Tokyo. The extent of intra-speaker variability is seen to be greatest for the vowel /u/, and smallest for the



**Fig. 5:** Distribution of the nine cities selected for the present study.

vowel /e/.

Figure 7(a) illustrates the extent of interspeaker variability for 10 male adult informants from Tokyo. Each point indicates

**Table 4:** Index of intra-speaker variations of each vowel averaged over ten Tokyo speakers.

/a/	/i/	/u/	/e/	/o/
0.0267	0.0229	0.0289	0.0194	0.0258

the average of six utterances of one informant, and each ellipse indicates the vowel area within which the averaged formants of a vowel by an arbitrary speaker from Tokyo are expected to fall at 90% probability. Considering that the dialect spoken by the Tokyo informants constitutes the reference for the common Japanese, these ellipses may be regarded as indicating the ranges of tolerance for the vowels of the common Japanese. The same index of variability as used for intra-speaker variations is adopted here to express the magnitude of inter-speaker variations of mean formant frequencies of each vowel within a region.

Figure 7(b) illustrates the extent of interspeaker variability for 30 informants from Tokyo including male and female adults to 4-year old children. Interspeaker variability in this case is so great that the vowel regions tend to overlap in the F1-F2 plane, but are shown to be separated in the three-dimensional space of F1-F2-F3.

Figure 8 shows similar F1-F2 diagrams for each of the eight other regions (cities). While Sendai and Nagaoka display vowel areas similar to those of Tokyo, the remaining six cities show vowel areas somewhat

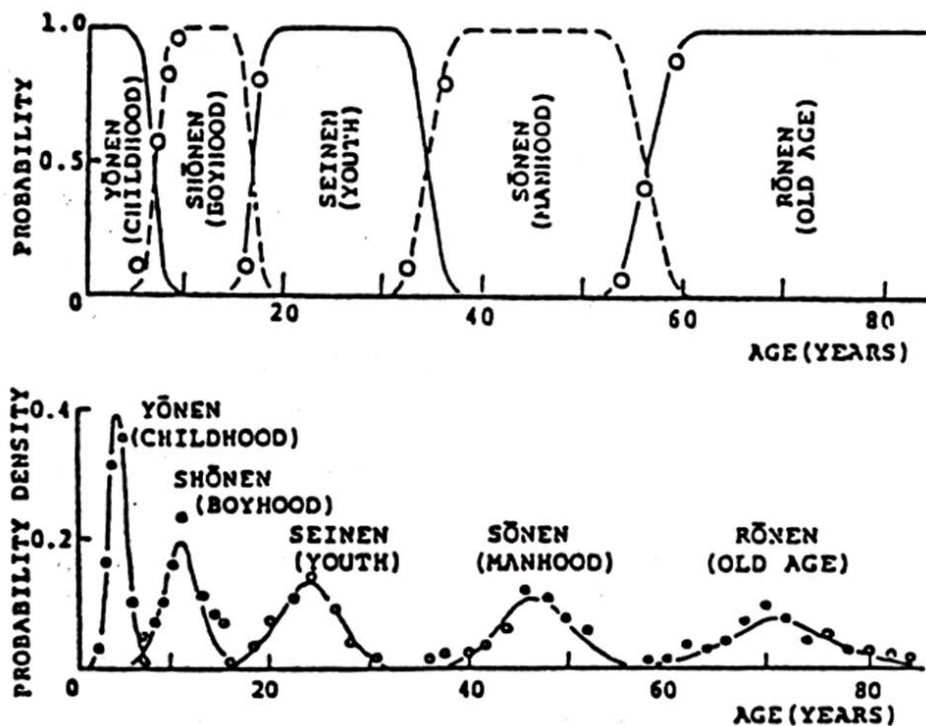


Fig.4: An example of coding and decoding characteristics in the use of 'age' terms of Japanese.

Table 3: Factors contributing to variability of speech sounds of the common Japanese.

Variability	Within a Speaker	Among Speakers	Among Speaker Groups
Stochastic & Unpredictable	Random Fluctuations	Individual Differences	
Systematic & Predictable	Contextual & Situational Differences	Age & Gender Differences	Regional & Dialectal Differences

different from those of Tokyo. The difference is seen to be most prominent for the vowel /u/.

Figure 9 shows the mean formant frequencies of vowels spoken in each of the nine regions. Each symbol represents the average F1-F2 values of a vowel spoken within a region, and the lines connect the five vowels of Tokyo. Regional differences are most conspicuous in the vowel /u/, but can also be observed, to a lesser extent, in vowels /a/ and /o/.

If we take into account that inter- and intraspeaker variations will be superposed on the regional differences, it can easily be seen that an error-free, speaker-independent recognition will be impossible even in the case of sustained vowels of the common Japanese. In fact, even human listeners are not free from recognition errors when these vowel samples are randomly presented. However, human listeners are capable of

quickly adapting to the vowel system of a particular speaker, when vowel samples from one speaker are presented in succession. These findings suggest that an intelligent system for automatic speech recognition should not aim at perfect speaker independence, but at efficient speaker adaptation based on a general knowledge of the sound system of the spoken language in question and a procedure to utilize the information contained in the immediately preceding context.

### 5. VARIABILITY IN SPONTANEOUS UTTERANCES

While the messages in a written text or in read speech follow the rules of grammar rather strictly and are thus well-formed, the messages commonly used in spontaneous speech (such as those occurring in con-

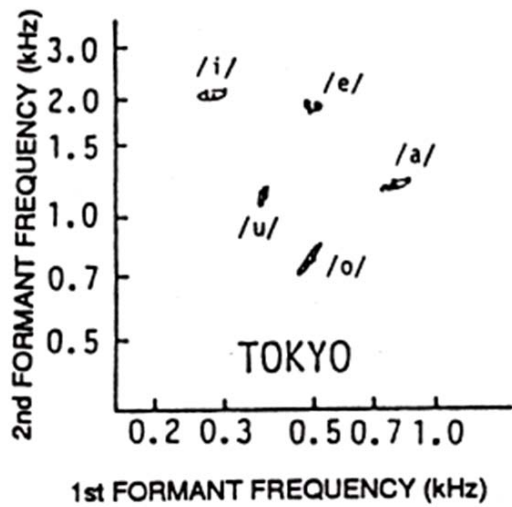


Fig.6: An example of intra-speaker variations of vowel formant frequencies. Each point corresponds to an utterance, and each ellipse indicates the area within which an utterance sample is expected to fail at a probability of 90%.

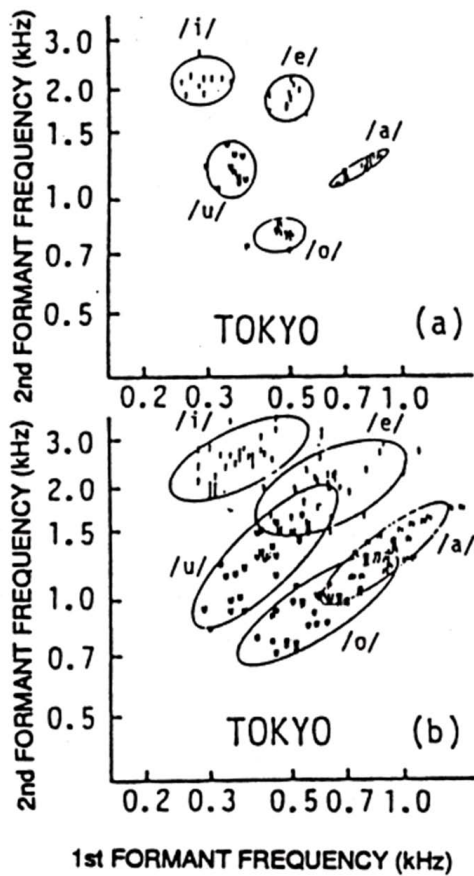


Fig.7: Inter-speaker variations of vowels of the common Japanese of Tokyo. (a) 10 male adult speakers. (b) 30 speakers from 4-year old children to adults, male and female.

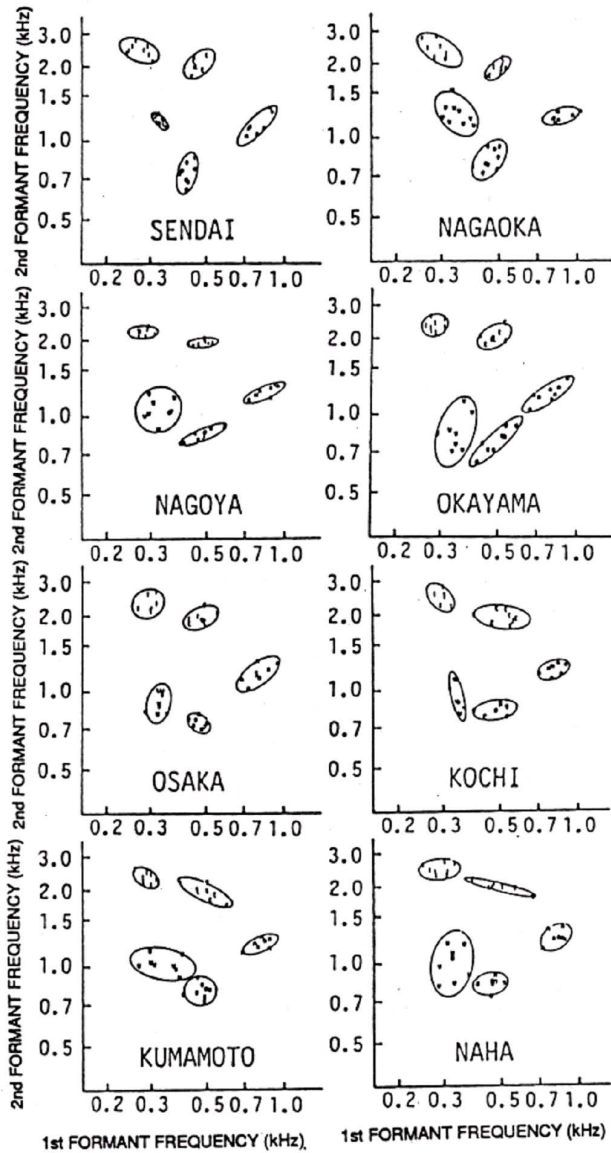


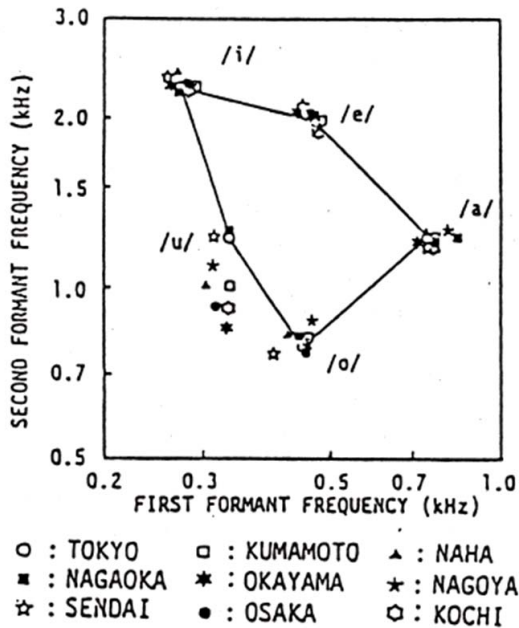
Fig.8: Inter-speaker variations of vowels of the common Japanese observed in eight major cities other than Tokyo.

versations) are quite different, often being fragmentary and illformed from the point of view of grammar of the written language. Their characteristics can be broadly classified into two categories:

- 1) Frequent occurrence of ellipsis and anaphora, because of large amount of shared knowledge between the speaker and the listener.
- 2) Frequent occurrence of errors and repetitions as well as of hesitations, filler sounds, re-starts, etc., because of the need for promptness of communication.

These so-called ill-formed utterances, however, do not cause serious difficulties in a human-to-human dialogue. This is mainly due to the following two principles observed by the speaker and the listener:

- 1) Principle of cooperation - the speaker and the listener cooperate to ensure smooth and prompt ex-



**Fig.9:** Regional differences in vowel formant frequencies of the common Japanese. Each point indicates the mean of eight to ten informants from one region.

change of information.

2) Principle of reliability - the speaker uses various methods for drawing attention, confirmation, indication of speaking turns, repair of errors, etc. to ensure reliability of communication.

These principles have to be incorporated also in a reliable man-machine dialogue system.

## 6. LEARNING FROM HUMAN PROCESSES

### 6.1 Human Processes of Language Comprehension

While it is true that natural language expressions are full of sources of ambiguity, and their disambiguation presents a formidable task to machine processing, a human receiver (listener/reader) does not necessarily notice all the sources of ambiguity present, or rather does not notice most of the sources of ambiguity, and quickly derives the interpretation intended by the sender.

This almost unconscious disambiguation is based mainly on the cognitive strategies and procedures that are quite different from the search procedures for possible interpretations adopted in machine processing. An investigation has been conducted to find out some important features of the cognitive strategies and procedures in reading written messages, and then incorporate some of the findings into a model, which is then included in an experimental system for parsing Japanese texts (Ikemori and Fujisaki, 1984). Figure 10 shows a general model for the human processes in the comprehension of written language mes-

sages, while Table 5 lists some of the findings of our investigation.

The use of such a model in written language processing has two definite advantages:

1) In analysis of linguistic messages, it can disregard most of the sources of ambiguity which a human receiver will not notice, and thus expedite machine processing. In other words, the system will point out only those ambiguities which will present difficulties for a human receiver.

2) In generating linguistic messages, it can provide a check on the ambiguity (for a human receiver) of the message to be generated. If the message is judged to be ambiguous by the model, it is modified in such a way as to avoid ambiguity. In this way, the system can ensure that its output messages are unambiguous for human receivers.

**Table 5:** Major findings of a psychological study on the human processes of comprehending written messages.

#### 1. Size of Processing Units

A string of approximately 15 characters can be accepted visually but is segmented into smaller sized units generally consisting of a content word and a few function words.

#### 2. Determination of Lexical Meanings

Contextual information is used to assign an appropriate meaning for a processing unit. The assignment is based on prediction.

#### 3. Parsing Procedures

Every time a verb is found, a check is performed on the semantic consistency between the verb and the preceding case elements. The check is carried out in the order of proximity to the verb in question.

#### 4. Use of Contextual Information

Utilization of contextual information follows the general principle of "from old to new" information.

An experimental system incorporating these features has been constructed and demonstrated (Fujisaki et al., 1985).

A similar investigation has also been conducted on the cognitive strategies and procedures in hearing spoken messages. Figure 11 shows a model constructed on the basis of the experimental results (Fujisaki et al., 1990).

### 6.2 Human Processes of Language Acquisition

Almost all the existing schemes for machine processing of spoken/written languages are trained on a



pre-determined samples/corpus, or are operated by rules derived therefrom. On the contrary, humans acquire skills of language use only gradually through exposure to a large amount of data over a long period. In fact, they acquire new expressions during the whole span of their active lives. This is absolutely necessary since a language is an open system in which new expressions are created whenever necessary. An intelligent system for machine processing should also be capable of acquiring language, i.e., of inferring the phonetic, syntactic and semantic properties of a new word (i.e., *unknown* word) which is not registered in the system's lexicon, without being taught by human teachers.

An extensive study has been conducted on the classification of the types of new words found in newspaper articles (Fujisaki et al., 1989). A pilot study has also been conducted on the actual process of inferring the meaning of unknown words by human subjects, and the results have been incorporated in a system with the ability of acquiring certain types of unknown words (Kameda, 1995).

## 7. NEW FRONTIERS OF RESEARCH

### 7.1 Beyond Linguistic Information

It is true that speech is a means to convey linguistic information, i.e., lexical, syntactic, semantic, and pragmatic contents of a message, it also conveys other kinds of information. In my terminology, these can be classified into two broad categories: para-linguistic and non-linguistic. The information concerning the speaker's intentions (e.g., exhortation, question, suspicion, etc.), attitude (e.g., politeness, friendliness, etc.), and styles (e.g., fast/slow, formal/informal, etc.), which are usually under the conscious control of the speaker, can be considered to fall into the former category (para-linguistic), while the information concerning the speaker's physical states (e.g., age, gender, health, idiosyncrasies, etc.) and emotional states (e.g., joy, sorrow, anger, fear, etc.), which are usually not under the conscious control of the speaker, can be considered to fall into the latter category (non-linguistic), though conscious simulation is possible, as is done by actors. Establishing a framework for the representation of these kinds of information as well as finding their articulatory, acoustic, and perceptual correlates presents new, unexplored areas of investigation which are important not only for their own sake, but also for certain practical applications such as synthesis/recognition of emotional speech as well as interpreting telephony capable of conveying subtle nuances.

### 7.2 From Mind to Mind - The Ultimate Goal of Language Science and Language Technology

I have already mentioned that language is a means to convey information, but one may naturally ask: Where does the information come from, and where does it go to? Ultimately it comes from the mind of the speaker/writer (i.e., the sender) and goes into the mind of the listener/reader (i.e., the receiver). Thus language is merely a medium of communication between two minds. In this sense, language models that are being widely used in conventional speech recognition systems only serve as crude approximations to the speaker's mind as the source of information. The ultimate system for spoken/written language comprehension should have a model of the sender's mind which involves models of the sender's self, of the receiver, of the rest of the world as seen by the sender, as well as of the processes of selecting a relevant piece of information, and of constructing the message. It should also have a model of the receiver's mind which involves models of the receiver's self, of the sender, and the rest of the world as seen by the receiver, as well as of the processes of constructing the expected message from all these sources of knowledge. Likewise, the ultimate system for spoken/written language generation from concept should also have models of the sender's mind and of the receiver's mind.

## 8. SUMMARY

This paper has discussed several issues that have to be dealt with in order to achieve an 'intelligent' processing of linguistic information, based mainly on the author's own works.

Ambiguity of a language expression has been defined as the state of having more than one meaning within a given context and with given background knowledge, and has been shown to have at least 10 different sources in the case of the Japanese language, some being language specific.

Vagueness has been defined as the result of decoding a linguistic expression into a continuous attribute, such as a color name into a color continuum. It has been made clear that the coding process and the decoding process are two essentially different psychological processes, and that both can be formulated and described in quantitative terms.

Variability of linguistic expressions has been discussed in two contexts, one in connection with speaker variability of Japanese vowels, the other in connection with linguistic forms of spontaneous speech.

In order to cope with these characteristics of language in an intelligent way, it is considered necessary to understand how humans deal with them. This paper discussed two aspects of human processes: language comprehension and language acquisition, with

an aim to model them and to incorporate into machine processing.

Finally, new frontiers of research have been briefly discussed.

## References

- [1] Fujisaki, H. (1995). Future of speech science and spoken language technology. *European Studies in Phonetics and Speech Communication* (Bloothoof, G., Hazan, V., Huber, D. and Llis-terri, J., eds.), OTS Publications, Utrecht, the Netherlands, pp. 15-18.
- [2] Fujisaki, H. (2004). Problems and prospects of spoken language processing. *Proceedings of the International Symposium on Speech Technology and Processing Systems and Oriental COCOSDA-2004*, New Delhi, India, vol. 2, pp. 137-140.
- [3] Fujisaki, H. and Hoshiai, T. (1984). Ambiguity in linguistic expressions. *Knowledge Engineering* (Tanaka, K., ed.), Asakura Publishing Co., Tokyo, Japan.
- [4] Fujisaki, H. and Kameda, H. (1984). Mutual conversion between linguistic expressions and knowledge representations in weather forecasts. *Transactions of Institute of Electronics and Communication Engineers of Japan*, vol. J67-D, no. 8, pp. 924-931.
- [5] Fujisaki, H. and Katagiri, Y. (1981). Formulation of the information transmission process by words and its quantitative evaluation. *Transactions of Institute of Electronics and Communication Engineers of Japan*, vol. J64-D, no. 15, pp. 394-402.
- [6] Fujisaki, H., Kamba, T. and Ikemori, M. (1985). A system for sentence correction within paragraph context (2) - reduction of ambiguity and difficulty of comprehension. *Record of the 31st National Convention of the Information Processing Society of Japan*, pp. 1667-1668.
- [7] Fujisaki, H., Kameda H. and Morita, T. (1989). Construction of a natural language understanding system based on a model of human processes of natural language processing. *Issues in Advanced Processing of Linguistic Information* (Nagao, M., ed.), pp. 657-684.
- [8] Fujisaki, H., Morikawa, H. and Udagawa, H. (1985). Analysis of variability in vowels of the common Japanese. *Proceedings of 1985 IEEE International Conference on Acoustics, Speech, and Signal Processing*, supplement, pp. 1-4.
- [9] Fujisaki, H., Nakamura, N. and Yoshimune, K. (1970). Analysis, normalization and recognition of sustained Japanese vowels. *Journal of the Acoustical Society of Japan*, vol. 26, pp. 152-154.
- [10] Fujisaki, H., Hirose, K., Ohno, S. and Mine-matsu, N. (1990). Influence of context and knowledge in the perception of continuous speech. *Proceedings of 1990 International Conference on Spoken Language Processing*, Kobe, Japan, vol. 1, pp. 417-420.
- [11] Fujisaki, H., Kameda, H., Kamba, T. and Akashi, T. (1986). Analysis of contextual information necessary for natural language understanding. *Record of the 32nd National Convention of the Information Processing Society of Japan*, pp. 1565-1566.
- [12] Ikemori, M. and Fujisaki, H. (1984). A study on the process of language understanding based on psychological experiments. *Record of the 48th National Convention of the Psychological Society of Japan*, p. 419.
- [13] Kameda, H. (1995). Automatic acquisition of unknown words by analogy based on examples. *Proceedings of the 1995 Engineering Sciences Society of the Institute of Electronics, Information, and Communication Engineers*, pp. 225-226.
- [14] Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, vol. 8, pp. 338-353.



**Hiroya Fujisaki** was born on October 18, 1930. He received the BS, MS and Ph.D. from University of Tokyo with the Fulbright Scholar at MIT from 1958 to 1961. He was a guest researcher at KTH on 1960, joined the Faculty of Engineering, University of Tokyo in 1962 and has been a full Professor at EE Department from 1973 to 1991 and serving twice as Chairman, also Professor of Speech Science at Graduate School of Medicine in 1974 to 1977. Since 1991 he has been a Professor Emeritus at University of Tokyo and a Professor at Tokyo University of Science. His research interests include Speech Communication and Spoken Language Processing, Natural Language Processing, Human and Artificial Intelligence.