ISCA Archive
http://www.isca-speech.org/archive

3rd International Conference on
Spoken Language Processing (ICSLP 94)
Yokohama, Japan
September 18-22, 1994

MULTILINGUAL LANGUAGE GENERATION ACROSS MULTIPLE DOMAINS[1]

James Glass, Joseph Polifroni, Stephanie Seneff[2]

Spoken Language Systems Group
Laboratory for Computer Science
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139 USA

ABSTRACT

In this paper we describe a mechanism developed to generate well-formed sentences across multiple domains from a common semantic representation. This language generation component is embedded in conversational systems that permit users to query databases via spoken requests. The generation mechanism serves two distinct but overlapping roles: *paraphrasing* user utterances, and *responding* to user queries. The generation component is completely table-driven, with separate tables controlling the generation for each distinct language. Convenient mechanisms for specifying inflectional endings and for handling movement phenomena have been developed. The system can generate responses/paraphrases for a variety of languages and for several different database-query tasks. We report evaluation results for paraphrasing in an air-travel domain.

INTRODUCTION

Over the past five years our group has become increasingly involved in the development of conversational systems. We believe such systems will play an important role in facilitating information retrieval between humans and computers. In such interactive problem solving situations, the role of language generation is important since it can provide useful and easily comprehensible feedback to the user. This capability will be an even higher priority for displayless systems (e.g., telephone-based) where there is no visual feedback at all and the information must be succinctly captured in the spoken response.

In working in this area, our group has focused on multiple application domains, such as urban navigation using Yellow Pages and geographical map information [1, 2], and air-travel planning using an on-line airline reservation system [3]. In addition, we have worked with several different languages in many of these domains [4, 5, 6]. Aside from English we have thus far worked most extensively with Japanese, Italian, and French, and have ongoing efforts with German, Spanish, and Mandarin. One of the benefits of working in multiple domains and languages is that we are highly motivated to focus on issues of modularity, generalization, and portability.

In this paper, we describe the current state of our language generation component (called GENESIS) which is used both to paraphrase meaning representations and to generate responses to the user. This component is used in all of our domains and across all output languages. We will first give an overview of our conversational system architecture before describing the generation mechanism in more detail. This will be followed by a description of our paraphrasing capability and response generation. We will then report results on some paraphrasing evaluations. Finally, we summarize this work, and report on future plans in the area of language generation.

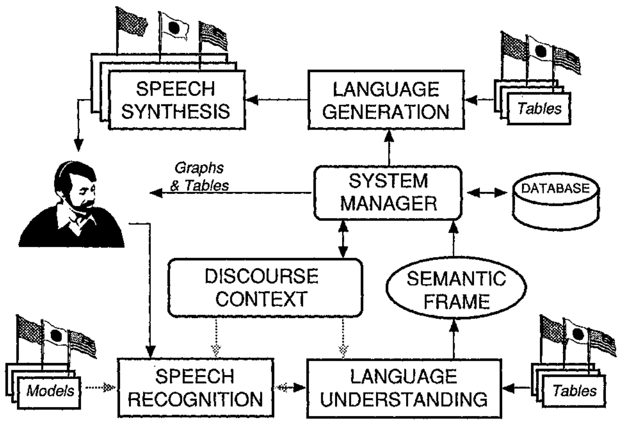

Figure 1: Schematic of prototypical MIT-SLS conversational system.

GENERAL DESCRIPTION

Figure 1 shows a diagram of our conversational system architecture. Speech input is processed by our segment-based speech recognizer [7] and probabilistic natural language component [8], and is converted to a meaning representation called a *semantic frame*. The semantic frame is designed to capture the meaning of the utterance in a way that preserves the critical hierarchical dependencies among the various parts of the utterance.

Our view of an appropriate format for meaning representation has undergone significant evolution since we began to address this problem five years ago. We have tended toward simplification of nomenclature, leading to the current view that all major constituents in a sentence can be characterized as one of: *clause*, *topic*, and *predicate*. The pred-

[1]This research was supported by ARPA under Contract N00014-89-J-1332 monitored through the Office of Naval Research.

[2]Names in alphabetical order.

```
clause: wh_query
        topic: restaurant
               quantifier: def
               name: royal_east
        predicate: serve
                   topic: a_cuisine
                          quantifier: which
                          name: kind
```

Figure 2: Semantic frame for the query: "Can you tell me what kind of food the Royal East serves?"

icate category includes adjectives and prepositional phrases as well as the standard verbal predicates. The relevant syntactic organization of the sentence is encoded through the hierarchy in the semantic frame. An example semantic frame for the utterance "Can you tell me what kind of food the Royal East serves?" is shown in Figure 2. Notice that the semantic frame has not preserved the politeness form in the main clause of the utterance, promoting the noun clause to the top level.

GENERATION MECHANISMS

Language generation is composed of three modules: a lexicon, a set of message templates, and a set of rewrite rules. These modules are language-dependent and external to the GENESIS system itself. In this way, porting the language generation component of an entire conversational system to a new language is confined to developing a new lexicon, messages, and rewrite rules, with the system kernel remaining the same. Since the semantic frame uses English as its specification language, entries in all lexicons, including English, contain words and concepts found in the semantic frame, expressed in English, with corresponding surface realization forms in the target language. The following sections describe each of these modules in greater detail and provide examples of their use in English and French.

Lexicon

The lexicon's main role is to specify the surface form of a semantic frame entry, including the construction of inflectional endings (gender, case, number, etc.). A sample lexicon for English and French is shown in Table 1. As can be seen in the table, each entry in the lexicon contains a *part of speech* tag (e.g., N (Noun), V6 (Regular Verb #6)), a stem, and various derived forms (e.g., the entry for "which" has several realizations in French depending on gender and number). For entries whose morphological variants are regular there are default endings specified under generic *part of speech* entries (e.g., a typical noun (N) in English forms plurals by the addition of an "s"). These defaults can be overridden by an exception specified for the particular entry, as in the English verbs "be" and "do".

Individual entries in the lexicon are also able to specify grammatical attributes that are necessary to control lexical form. In French for example, nouns can specify their gender (e.g., "flight" is masculine), which is required for proper generation of adjectives and quantifiers. In addition, entries can specify default quantifiers (e.g., "royal_east", a proper noun, prefers a definite article). Furthermore, certain auxiliary verbs, such as "do" and "will", can set the verb mode for the main verb (e.g., to "root" mode in English). Finally, the surface form of a particular lexical entry can be controlled by the semantic class of its parent. For instance, numbers can be entered in the semantic frame as simple integers, and realized as cardinal or ordinal (e.g., "second" vs "two") depending upon the semantic class of their parent.

English Lexicon:		
V	V	"Verb" THIRD "es" ROOT "e" ING "ing"...
N	N	"Noun" PL "s"
be	AUX	"be" ROOT "be" THIRD "is" ING "being"...
do	AUX	"do" THIRD "does"... MODE "root"...
indef	Q	"a" PL "any"
which	TRACE	"what"
royal_east	P	"Royal East" Q "def"
serve	V	"serv"
on_street	PREP	"on"
French Lexicon:		
N	N	"Noun" PL "s" F "e" FPL "es"
V6	V	"Verb" ROOT "vir" THIRD "t" FPL "vons"...
be	AUX	"etre" ...FPL "sommes"...
which	TRACE	"quel" F "quelle" MPL "quels" FPL "quelles"
royal_east	P	"Restaurant Royal East" Q "def"
serve	V6	"ser"
on_street	PREP	"dans"

Table 1: Selected entries from the lexicon for the English and French examples given in the text.

existential	(:AUX be) there :TOPIC .
wh_query (English)	:TRACE (:AUX be) (:TOPIC it) :PRED :PREP ?
wh_query (French)	:PREP :TRACE (:PRED be) :TOPIC ?
topic	:QUANTIFIER :NOUN_PHRASE
street	:TOPIC :STREET_TYPE
serve	:PREDICATE :TOPIC
np-on_street	:NOUN_PHRASE :PREDICATE :TOPIC

Table 2: Selected message templates for examples given in the text.

Messages

The catalog of message templates is primarily used to recursively construct phrases describing the topics, predicates, and clauses of a semantic frame. Table 2 shows example message templates for English and French. A message template consists of a message *name* and a sequence of one or more word strings and keywords. There is also a mechanism for optionally specifying a default value in the event the keyword has no value. The set of message templates controls the ordering of constituents, which are instantiated recursively.

Rewrite Rules

The rewrite-rules are intended to capture surface phonotactic constraints and contractions. For example, in French the sequence "de le" is realized as "du". In English we use rewrite rules to generate the proper form of the indefinite articles "a" or "an", or to merge "a other" into "another".

PARAPHRASING

We have found the ability to paraphrase a semantic frame to be very useful for a number of different purposes. It serves as a kind of translation among the various languages supported by the system, with the semantic frame acting as a form of "interlingua". It is also of great use to system developers when porting to a new domain or language, by providing a confirmation that the natural language component successfully parsed the input query *and* generated an appropriate semantic frame. Finally, some aspects of the paraphrasing are used as part of the response generation.

```
clause: existential
        topic: bank
                number: pl
                quantifier: indef
                predicate: on_street
                            topic: street
                                    name: Main
                                    street_type: street
```

Figure 3: Semantic frame for the query: "Are there any banks on Main Street?"

The following sections describe issues involved in paraphrasing the three major structures of the semantic frames: clauses, topics, and predicates. Examples will be drawn from the semantic frames in Figures 2 and 3 using the lexicon and message templates described in Tables 1 and 2. Finally, we discuss the special issue of movement phenomena.

Clauses

Clauses are typically the top-level structure of the semantic frame. Their role is to synthesize the paraphrases from the sub-level structures (i.e., topics and predicates) into a sentence-level message. For example, the input query "Are there any banks on Main Street?" will produce a top-level clause of type [EXISTENTIAL] as shown in Figure 3. From Table 2 we see that the corresponding message has the form "(:AUX be) there :TOPIC" in English. In the event that the auxiliary is not explicitly marked at the clause level, it generally defaults to "be" unless there is a *verbal* predicate at clause level (verifiable from the lexicon), in which case it defaults to "do". Such complicating factors do not apply for our simple example, since "existential" never expects a predicate at clause level. The auxiliary's surface form depends upon number (singular or plural) which is established from the value in the clause's topic.

Topics

Topics typically correspond to noun phrases, and in addition to their generic type, can contain quantifiers, names, and one or more predicate modifiers. Topic paraphrases are synthesized by first creating a core topic noun phrase (which is capable of handling conjunctions), and then recursively adding predicate information. The quantifier is generated once the topic core has been created so that lexical features (e.g., gender, number), contained in the semantic frame and/or looked up in the lexicon during the generation of the core, can be correctly incorporated. In our current example, the core topic noun phrase is "banks" and the quantifier is the plural form of the indefinite article, "any".

The ordering of multiple predicates in the semantic frame is ignored by the generation. Instead, predicates are added in the order of their occurrence in the message catalog, allowing the ordering to be language dependent. In addition, superlatives (e.g., cheapest, earliest) are marked by a separate predicate entry in the topic and can be treated differently from other predicates if desired.

Compound noun phrases present a complicating factor. The parser marks modifiers (predicates) as "global" when it is inferred from the parse tree that they scope over both nouns. The local predicates are first generated to attach to the individual nouns during the construction of the topic core, and any remaining global predicates are subsequently generated.

When necessary, topics can have specific messages which override the generic topic message. For example, the second topic of the example query, [STREET] is synthesized using its corresponding message ":NAME :STREET_TYPE", to produce the paraphrase "Main Street".

Predicates

In general, the surface form for a particular predicate depends upon whether or not it is contained in a main clause. Thus, the system offers separate control of the generation for predicates in clauses as contrasted with predicates modifying noun phrases. For example, the "serve" predicate in Figure 2 must generate the main verb phrase "serve chinese food", whereas this same "serve" predicate within a noun phrase might only have to generate "Chinese" as a premodifier to the noun, as in "Chinese restaurant". This distinction is indicated through separate "np-" prefixed predicate message templates, as shown in Table 2. In our bank example, the predicate [ON_STREET] adds its own paraphrase to the core noun phrase "banks", by consulting the [NP-ON_STREET] message template: ":NOUN_PHRASE :PREDICATE :TOPIC", thus producing "any banks on Main Street". The topic referred to here is the topic internal to the [ON_STREET] predicate.

Topic Movement

One of the most difficult aspects of multilingual paraphrase generation involves movement phenomena, such as occurs in wh-queries, which are prevalent in database query domains. In this section, we will examine two such wh-queries in English and French, to demonstrate how our paraphrase mechanism deals with the different restrictions on movement for these two languages. In the first example, the surface form sentence in English "What kind of food does the Royal East serve?" is realized by moving the quantified noun phrase "what kind of food" forward from its logical syntactic position after the main verb to the very beginning of the sentence. Our analysis component restores this phrase to its deep-level structural position, as shown in Figure 2. A direct generation of the paraphrase would produce "Does the Royal East serve what kind of food?" which would not be a fluent English form.

The corresponding French utterance, "Quel genre de cuisine sert le restaurant Royal East?" exhibits similar movement. However, in French the subject and predicate are inverted, whereas in English the auxiliary "do" marks the question form. French and English exhibit other differences in the formation of wh-queries, as in the English-French pair "What street is MIT on?" and "Dans quelle rue est MIT?" In English, just the wh-quantified noun phrase moves, leaving the preposition behind; in French, the entire prepositional phrase must be moved forward.

These differences between English and French are captured mainly through the distinct message templates for [WH_QUERY] for the two languages, as shown in Table 2. The quantifier "which" flags a noun phrase as a [:TRACE]. Predicates containing such noun phrases are identified as preposition [:PREP] or verb [:PRED] based on their *part of speech* entry in the lexicon, and their positioning is separately controlled in the [WH_QUERY] template. In French, the [:PREP] entry precedes the [:TRACE], and, if there is no verbal predicate, the auxiliary "be" is introduced. In English there is

Correct Paraphrase	Incorrect Paraphrase	Incorrect Frame	Out of Domain
77%	7%	9%	7%

Table 3: Tabulation of English paraphrase results for 100 unseen test sentences within the air-travel domain.

always an auxiliary, and both the prepositional [:PREP] and the verbal [:PRED] predicates are positioned after the topic.

RESPONSE GENERATION

As shown in Figure 1, the language generation component of our conversational systems is controlled by the *System Manager*. In this capacity its role is to provide answers, clarification requests, help, and other computer-initiated feedback, in order to enable a user-friendly dialogue. Responses are constructed via the same procedure used for paraphrasing an utterance, as previously discussed.

Responses are typically constructed from a synthesis of information provided by the user and by the database. The *System Manager* creates a response semantic frame derived from the input frame and modified to reflect the outcome of the database query. For instance, in GALAXY, the quantifier and number of the response frame depend on whether the result is a null set ("There are no <NP>"), a single item ("There is only one <NP>"), or a larger set ("There are twenty five <NP>"). The noun phrase <NP> is generated directly from the main topic of the input semantic frame, and may be a complex noun phrase such as "flights from Boston to Denver serving lunch and arriving between 2:30 and 3:30 p.m." This communication serves a useful role in verifying the system's understanding of the input query.

Complex messages involving *navigational directions* can be controlled by a small set of possible message types, with the particular values for verbs, etc., instantiated from the path-finding algorithm.

EVALUATION

It is not very clear how to design an evaluation metric for this generation system, particularly for the response component, where judgments are necessarily highly subjective [1]. However, we decided it would be possible to evaluate the *paraphrasing* capability (English to English) for a small unseen set of 100 utterances in the air-travel domain, by using subjective judgment to decide whether a paraphrase was a correct rendition of a particular semantic frame. The situation was complicated by the possibility that the parser might not have produced a correct semantic frame from the original sentence. The system being tested had been developed based on a set of some 300 to 400 training queries.

Table 3 summarizes the results of this experiment. The parser produced an inappropriate semantic frame for 16% of the queries, approximately equally divided among "answerable" and "unanswerable" queries. Of the remainder, nearly 92% produced a correct English paraphrase. Only seven queries had a correct frame but an incorrect paraphrase. Minor bugs accounted for two of the errors. An additional error was easily corrected by rule editing, and another had to do with the inability to specify vocabulary items under multiple syntactic categories. The remaining three errors were due to minor ordering problems or idiosyncracies.

We also maintain a generation development set of several hundred utterances in each of the domains in which we work. These utterances are chosen by hand, from training data, to represent the various types of queries possible, and are augmented with the correct paraphrase. These provide a convenient mechanism to check the integrity of the system following any modifications.

SUMMARY AND FUTURE WORK

We have described a generation system that has proven to be quite useful for the limited task of paraphrase/response in multilingual and multi-domain database query systems. However, our system is by no means complete. We have slowly increased the complexity of the system to accomodate new phenomena, but we believe that spoken language is inherently simpler in construct than written language, especially when restricted to database query. For instance, we rarely encounter a sentence not in the present tense, and therefore tense has been neglected in our generation mechanisms. The lexicon should be enhanced to allow for multiple parts of speech for the same word (e.g., "connecting" as an adjective/verb), a phenomenon which occurs only rarely in our domains. Our handling of movement phenomena has been tested out only for English and French, and, even then, only for wh-query movement.

We are just now beginning to develop a *displayless* version of our PEGASUS system. In this case, much more information must be conveyed via synthetic speech, so it will be crucial to have powerful and easily controllable generation techniques.

REFERENCES

[1] V. Zue, J. Glass, D. Goodine, H. Leung, M. Phillips, J. Polifroni, and S. Seneff, "The VOYAGER Speech Understanding System: Preliminary Development and Evaluation," *Proc. ICASSP-90*, Albuquerque, NM, pp. 73–76, April, 1990.

[2] V. Zue, E. Brill, J. Glass, D. Goddeau, C. Pao, M. Phillips, J. Polifroni, and S. Seneff (1994b), "A Spoken Language Interface in the GALAXY Travel Domain," *these proceedings*, 1994.

[3] V. Zue, J. Glass, D. Goddeau, D. Goodine, C. Pao, M. Phillips, J. Polifroni, and S. Seneff, "PEGASUS: A Spoken Dialogue Interface for One-line Air Travel Planning," *Proc. International Symposium of Spoken Dialogue*, pp. 157–160, November, 1993.

[4] G. Flammia, J. Glass, M. Phillips, J. Polifroni, S. Seneff, and V. Zue "Porting the Bilingual VOYAGER System to Italian," *these proceedings*, September, 1994.

[5] J. Glass, D. Goodine, M. Phillips, S. Sakai, S. Seneff, and V. Zue, "A Bilingual VOYAGER System," *Proc. Third European Conference on Speech Communication and Technology*, pp. 2063–2066, September, 1993.

[6] H. Bonneau-Maynard, J.-L. Gauvain, L. Lamel, J. Polifroni, and S. Seneff, "A French Version of the MIT ATIS System: Portability Issues," *Proc. Third European Conference on Speech Communication and Technology*, pp. 2059–2062, September, 1993.

[7] V. Zue, J. Glass, D. Goodine, M. Phillips, and S. Seneff, "The SUMMIT Speech Recognition System: Phonological Modelling and Lexical Access," *Proc. ICASSP-90*, Albuquerque, NM, pp. 49–52, April, 1990.

[8] S. Seneff, "TINA: A Natural Language System for Spoken Language Applications," *Computational Linguistics*, Vol. 18, No. 1, pp. 61–86, 1992.