MYRTILLE II : AN UNIFIED APPROACH TO PSEUDO-NATURAL LANGUAGE UNDERSTANDING

Jean-Paul HATON

CRIN - UNIVERSITE DE NANCY I (France)

Introduction

Our group has been involved for the past ten years in various research areas of man-machine communication : drawing and image analysis and interpretation, speech recognition and understanding.
After having developed several systems for the recognition of isolated words and of artificial language sentences, we started in 1976 the MYRTILLE II project for pseudo-natural language understanding. In this paper we present the basic ideas involved in the project and we briefly describe the various processing levels of the system.
MYRTILLE II was designed in order to meet the following requirements (which are to be compared with the specifications given in the ARPA Project (Ref. 1) :
- understanding of an "expert language", i.e. a subset of spoken french with few limitations on syntax and semantically restricted by the universe of the application (presently : inquiry to a meteorological data-base, this application corresponds to a 400-word vocabulary)
- experimental, modular system : MYRTILLE II will be used for testing several algorithms and strategies on various applications,
- importance of dialog in order to provide a non-trivial communication with the user.

The system is made up of three major components :
- an acoustic-phonetic decoder which converts the input sentence into a string or a lattice of phonemes,
- a linguistic recognizer which accepts the phoneme string as input and recognizes the sentence, i.e. it yields its syntactic-semantic representation under the form of a tree structure (surface structure),
- a pragmatic interpreter which derives from the surface structure the semantic interpretation (deep structure) of the sentence. This representation is readily usable for answering the user's query (e.g. access to a data-base, etc...). The interpreter is made up of a semantic analyzer using a case-grammar model (Ref. 2) modified for spoken input and of a dialog procedure. The dialog generates questions necessary for solving ambiguities which may appear during the recognition process ; it also processes the user's answers to these questions.

The pragmatic level is presently being implemented and it has not yet been tested. Therefore we will now concentrate on the two other processing levels and examine some results.

The Acoustic-Phonetic Decoder (Ref. 3)

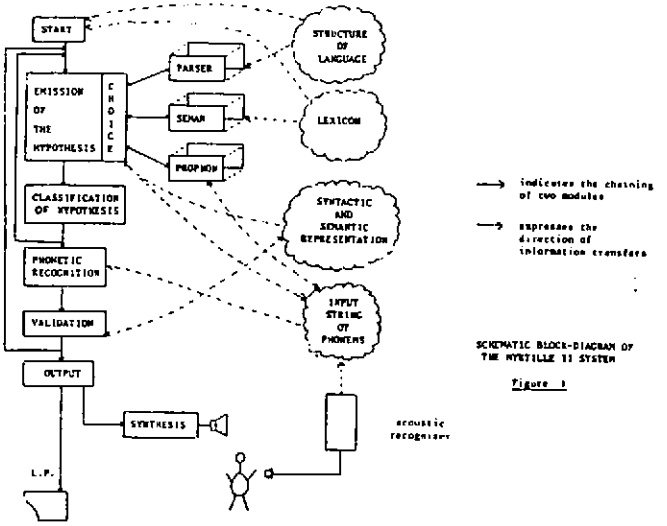
In order to keep all the available information the system directly works on the digitized waveform.
The speech signal is first low-pass filtered and then 10 bit quantized at a 10 kHz rate. The phoneme recognition process is two fold :
- segmentation of the speech wave into phoneme-like units by computation of a segmentation function on successive, overlapping 30 millisecond-length windows.

MYRTILLE II : AN UNIFIED APPROACH TO PSEUDO-NATURAL LANGUAGE UNDERSTANDING

This function incorporates the following parameters : zero-crossing rate, energy, length of the signal ; the segmentation process works differently on voiced and unvoiced portions of speech.
- Identification of the segments by evaluating several features on these segments (formant frequencies computed by LPC, energy in frequency bands computed by FFT, duration, etc...). Rather than directly evaluate all parameters for each segment we only compute the subset of parameters necessary to identify a given segment. The recognition system therefore appears like a hierarchical tree structure of processors. Each processor, whenever it is activated, yields a set of weighted answers which can be either other(s) processor(s) or phoneme classes. This structure turns out to be very convenient for removing or adding parameters.
The system has been tested for one speaker, after training and several results will be shown. We are presently working on some improvements of the recognition and on the important problem of speaker adaptation.

The Linguistic Recognizer (Ref. 4)

As with several other systems, MYRTILLE II operates on the principle of hypothesis-and-test. The hypotheses can be emitted by several knowledge sources, i.e. syntax, semantics and phonetics-phonology. Instead of data driven system, for instance HEARSAY II (Ref. 5), MYRTILLE II is controlled by a supervisor which decides at each step of the recognition process which knowledge source to activate. Figure 1 shows the overall organization of the system.



SCHEMATIC BLOCK-DIAGRAM OF THE MYRTILLE II SYSTEM

Figure 1

The data-structures in MYRTILLE II have been chosen for the sake of efficiency, especially at the lexical and syntactic-semantic levels.
The lexicon stores the phonetic transcriptions of the words together with usual phonemic alterations. Moreover, for each word are added references to phonological rules with the usual phenomena : plurals, liaisons, etc...

MYRTILLE II : AN UNIFIED APPROACH TO PSEUDO-NATURAL LANGUAGE UNDERSTANDING

The lexicon also includes the syntactic and semantic definitions of the words under the form of a hierarchical tree representation : 4 levels are defined which correspond to syntactic-semantic classes, with semantic relations between the classes. This representation makes it easier to retrieve the different adjectives which may qualify a noun, or the nouns which may be subject of such or such verb, etc... A similar solution was adopted in the SRI speech understanding system (Ref. 5).
The syntactic-semantic data essentially correspond to the description of the grammar of the language. It is quite impossible to describe them in a context-free formalism in the case of a natural language. We have chosen a network representation that we called "procedural network" which is very efficient for emitting hypotheses. In this representation, a branch corresponds to a word or to a sub network whereas procedures are linked to each node in order to take into account semantic and contextual information.
A preliminary version of MYRTILLE II is working in time-sharing on a IRIS 80 computer (Ref. 6). This version uses a best-first strategy and is able to recognize such french sentences, pronounced by one speaker after training :
Does it rain in Nancy ?
I would to know the speed of the wind in Brittany.
Will the sun shine tomorrow in the region of Paris ?
Etc...

References

1. A. NEWELL et al "Speech Understanding Systems : A Final Report "North-Holland, 1975.
2. J. FILLMORE in "Universals in Linguistic Theory" E. Bach, R.T. Harms, editors, Rinehard and Winston, 1967.
3. J.P. HATON, C. SANCHEZ "An Experimental System for Acoustic-Phonetic Decoding of Continuous Speech" IEEE ICASSP, Washington, 1979.
4. J.M. PIERREL, J.P. HATON "Data Structure and Organization of the MYRTILLE II System" 4th Int. joint Conf. Pattern Recognition, Kyoto, Japan, Nov. 1978.
5. D. WALKER "The SRI Speech Understanding System" IEEE Trans. ASSP, 23, pp. 397-416, 1975.
6. J.M. PIERREL "Etude et mise en oeuvre de contraintes linguistiques en compréhension automatique du discours continu". Doctorat d'Etat Thesis, University of Nancy 1, March 1981.

Appendix

The following is a recognition exemple of the french sentence :
"Est-ce que les risques de petites plaques matinales de verglas persisteront demain dans la région de Nancy ?".
The surrounded numbers indicate the numbers of word hypotheses emitted at each step.
Finally we give below the tree structure determined by the system.

MYRTILLE II : AN UNIFIED APPROACH TO PSEUDO-NATURAL LANGUAGE UNDERSTANDING

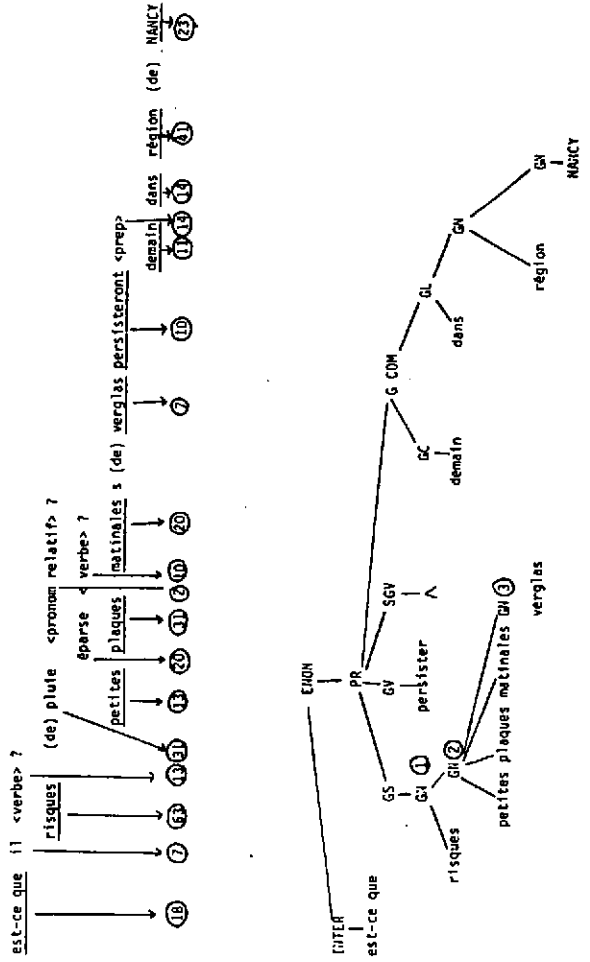


Figure 2

Recognition of the french sentence :
"Est-ce que les risques de petites plaques matinales de verglas persisteront demain dans la région de Nancy ?".