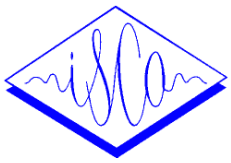

The ESCA Workshop
on Speech Synthesis
Autrans, France
September 25-28, 1990

AUTOMATIC ASSIGNMENT OF LEXICAL STRESS IN ITALIAN

PHILIPPE MARTIN

Experimental Phonetics Laboratory
University of Toronto
300 Huron St., Toronto
Ontario. Canada M5S 2X6

ABSTRACT

The generation of prosody in text-to-speech synthesizers usually requires the producion of an adequate fundamental frequency movement on the stressed syllables of the sentence. These pitch variations are supposed to correspond, among other factors, to some hierarchy which organize stress groups, and which can be equated to the syntactic structure of the sentence. With this hypothesis, the pitch changes located on the stressed syllables function as acoustic cues to pre-analyze (parse) the speech continuum as it is produced by the synthesizer. This paper deals with one part of this process: the automatic positioning of stress syllables from text in Italian, which is not graphically represented (except in final position), and for which no simple stress rule is known.

1. INTRODUCTION

Most of the text-to-speech synthesizers implicitly recognize the following correlates of sentence intonation:

1) a physiological constraint simulated by the declination line;
2) approximate limits of lexical units (word stress);
3) the borders of (large) syntactic groups;
4) the hierarchical organization of these groups in the sentence;
5) marking of words or larger units having a specific semantic/pragmatic role;
6) phonostylistic characteristics of the sentence/speaker.

The automatic generation of the sentence pitch pattern from text can be considered as a reverse ingineering process, attempting to recover information carried by the sentence intonation from information possibly present elsewhere in the text. It is clear that this process cannot be accomplished successfully in all cases from the sentence text alone, and without external information. This is particularly the case for lexical stress in Italian, whose position in the word results from an underlying morphological mechanism.

2. ASSIGNMENT OF LEXICAL STRESS

Lexical Stress in Italian can be located on one of the last four syllables in nouns, adjectives, pronouns and adverbs, and on one of the last six syllables in verbal forms. Since no simple stress rule is known, existing methods for automatic stress assignment are based on statistical correlations observed between stressed syllables and specific syllables structures, considered phonetically, phonologically or morphologically [Delmonte, 1981, Sandri and

Vivalda, 1981]. With the use of extensive lists of exceptions, statistical approaches lead to 90-95% of correct results, depending on the corpus used for tests.

The method described here is based on morphological analysis of the word and on a relatively unknown mechanism for stress assignment in Italian [Garde, 1968]. This mechanism specifies that, if a word can be analyzed into a stem and a flexion, with optional prefix, suffix and clitic pronouns (used in certain verbal forms)

word = (prefix) + stem + (suffix) + flexion + (clitic)

each element has a specific property of *stressability,* i.e. a form can be found where this element will bear the word stress:

a prefix is non stressable;
a stem is stressable on its last or penultimate syllable;
suffix and flexion are either stressable or non-stressable;
clitics are never stressable.

The stress rule simply specifies that the last stressable syllable determines word stress.

For example, the stem *'oper* has its penultimate syllable stressable, which leads to the following stress patterns for its derivatives:

'opera : 'oper + -a (non stressable flexion) "opera"
oper'oso: 'oper + -'os- (stressable suffix) + -o (non stressable) "hard working"
oper'etta: 'oper + -'ett- (stressable suffix) + -a (non stressable) "operetta"
operosit'a: 'oper + -'os- (stressable suffix) + -it'a (stressable) "fact of being hard working"

The same stem can lead to 2 or more derivatives:

turb'ina: turb'in + -a "turbine" *t'urbine*: t'urbin + -e "whirlwind"

Two different stems can lead to identical derivations (homographs), belonging either to different grammatical categories:

comp'ito: c'omp + -'ito (stressable verbal ending) "accomplished"
c'ompito: c'omp + -it- (non stressable suffix) + -o "assignment"

or to the same category:

'auspici: 'auspic + -i (non stressable noun ending), plural of 'auspice
ausp'ici: ausp'ic + -i (non stressable noun ending), plural of ausp'icio "predictions"

Automatic lexical stress assignement is based on the analysis of the word entry in its morphological components, using somewhat extensive lexicon of word stems (about 50,000 entries), prefixes (100 entries), suffixes (250 entries), endings (300), and clitics (40). Some 10,000 irregular forms of verbs, adjectives and nouns were included in this database.

Each entry is describes by grammatical and morphological features allowing appropriate matching of stems, suffixes, flexions and clitics. Classical categories of Noun, Adjective, Verb, Adverb, Pronoun, Preposition, Conjunction, Article and Proper Noun were used, together with morphological features of verbal mode, tense, person, number and gender. Morphological features specific to Italian verbs, nouns and adjectives were also used (Nouns in -a/-co/-io,..., verbs in -are, -iare, -care, ...). These features provide some

possibility of desambiguation for homographs belonging to different syntactic categories, and will allow syntactic labelling for fundamental frequency generation at a later stage of the work.

3. IMPLEMENTATION

To assign the lexical stress from text based on morphological analysis, a large database of stems, prefixes, suffixes, flexions (endings..) and clitics was build. This process was made easier thanks to a set of tools designed to help the operators to analyze more than 60,000 entries.

The program to enter the lexical stems operates from ordinary dictionary entries. Once the operator has entered a word in its canonical form, and specified its grammatical category, an analysis routine attenpts to extract the appropriate stem, together with the morphological sub-category and the most probable stressable syllable. A command then allows then the operator to scan through matching suffixes and endings, so that all possible derivatives of the stem are displayed on screen with the resulting stress pattern. If a form is not correct, assumption on categories and morphological features can then be easely corrected with a single keystroke. A lexicon of prefixes, suffixes, flexions and clitics are on line and can be edited on the fly as well.

4. ANALYSIS

In the analysis phase, each sentence is divided into groups ended by punctuation marks, and each word is analyzed into its possible morphological components. Since in the general case more than one grammatical category and morphological description is obtained for each word in a block, a desambiguation routine operates on groups of 1, 2 and 3 or more successive words, retaining the most probable sequence of categories in a scanning process proceding from left to right. The probabilities of occurence of these sequences are obtained by a statistical analysis of some words, whose categories and feautures were manually labelled. This process constitutes an approximation of a more elaborate error-correcting stochastic grammar, which would allow for a more systematic correction of wrongly labelled elements in the sentence.

5. AN EXAMPLE

The processing of the successive words the sentence *Oggi altero il mio spirito altero* (Today I change my proud spirit") gives the following results:

```
Oggi        'Oggi           Noun Mas   Inv
Oggi        'Oggi           Adv

altero      'altero         Vrb        Sing Ind  1st  Pres
altero      alt'ero         Adj  Mas   Sing

il          'il             Art  Mas   Sing

mio         m'io            Adj  Mas   Sing
mio         m'io            Pro  Mas   Sing

spirito     sp'irito        Noun Mas   Sing
```

```
altero    'altero        Vrb       Sing Ind  1st  Pres
altero    alt'ero        Adj  Mas  Sing
```

In this simple sentence, 4 words receive multiple grammatical land morphological labellings, which leads to 2 different stress assignment for the word *altero*, which is a homograph ('altero "I change" vs. alt'ero "proud"). The analyzer will then sort the combinatorial possiblities of 3 successive categories considered from left to right in the sequence:

```
'Oggi  'altero  il   m'io  sp'irito  'altero

  Adv    Vrb    Art   Adj     Noun      Vrb
               Mas   Mas     Mas
         Sing   Sing  Sing    Sing      Sing
         Ind                             Ind
         1st                             1st
         Pres                            Pres
```

Since the sequence **Adj Noun Vrb** has been found statistically more frequent than **Adj Noun Adj**, the desambiguation process fails to assign a correct category to the second occurence of *altero*, and the wrong version of the homograph is retained. Clearly more a more elaborate strategy must be used even in simple cases.

AKNOWLEDGMENT

Thanks to D. Heap and T. Nadasdi for their comments and suggestions.
This research was supported by a contract CNET No 88 1B 108.

REFERENCES

Delmonte, R. (1981), "L'accento di parola nella prosodia dell'enunciato dell'italiano standard", Studi di Grammatica Italiana, Vol. X, 351-394.

Garde, P. (1968), "L'accent", PUF, Paris.

Martin, Ph. (1979), "L'intonation de la phrase en Italien", Studi di Grammatica Italiana, Presso l'Accademia della Crusca, Firenze, 395- .

Martin, Ph. et Profili, O. (1987), "Accent de mot et structure syntaxique en italien", Information-Communication, Univ. of Toronto, 15-26.

Martin, Ph. (1990), "Positionnement automatique de l'accent lexical en italien en vue de la synthèse", Actes des XVIIIèmes Journées d'Etudes sur la Parole, Montréal, Canada, 149-152.

Quazza, S. and Vivalda, E. (1987) "Contextual Syntactic Analysis for Text-to-Speech Conversion", Proc. ICASSP 87, 389-392.

Sandri, S. and Vivalda, E. (1981), "Automatic Stress Assignment for Italian Text-to-Speech Synthesis", CSELT Rapporti Tecnici, Vol. VIII, No3, June 1981, 213-216.