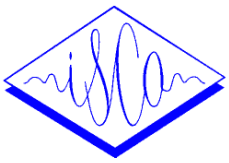

2nd European Conference on
Speech Communication and Technology
EUROSPEECH '91
Genova, Italy, September 24-26, 1991

A CODED DICTIONARY FOR STRESS ASSIGNMENT RULES IN ITALIAN

Marcello Balestri

CSELT – Centro Studi e Laboratori Telecomunicazioni S.p.A
Via G. Reiss Romoli, 274 – 10148 TORINO (Italy) – Tel. +39 11 2169457

ABSTRACT

This paper describes a method for the automatic location of the lexical stress in an Italian text. The method is based on the observation that in Italian many words ending with the same letters show the same stress position; for example, in all words ending in "*–grafia*" like "*fotografia*" (photograph) the stress falls on the penultimate vowel.

We can formulate a rule encompassing the majority of words showing the same stress pattern; words which do not correspond to the pattern being dealt with by an appropriate exception mechanism.

INTRODUCTION

A text–to–speech synthesis system is designed to transform any text written in the normal orthography of a language, into the corresponding vocal message.

In a text–to–speech synthesis system the lexical stress is very important not only for the correct and unambiguous pronunciation of every word, but also for the rhythm and intonation of the whole sentence to be synthesized. In fact, the duration of stressed vowels is at least twice as long as for unstressed ones. Furthermore, the number and the position of the stressed syllables of a sentence are of great importance for the determination of fundamental frequency discontinuities: prosodic rules locate f0 peaks around the position of some stressed syllables.

Lexical stress in Italian can be located on any of the last seven (graphemic) vowels and it needs to be graphically indicated in the text only when it falls on the final vowel (e.g. "*città*", town).

The form of the written word gives no indication of where the stress occurs, nor is there a set of exhaustive phonological rules governing the position of the stress accent, since this is determined by historical–etymological factors.

To localize the lexical stress of any word of a text two approaches are possible [1], [2], [3], [4], [5]:

1) access to a data base containing all existing words with the indication of the lexical stress position: this solution is very expensive because it requires a large data storage;
2) creation of models taking into account some regular phenomena which can be found in the lexical stress scheme of Italian words.

The method described in this paper uses this second approach; the idea is to analyze, from right to left, the word to be stressed until it becomes possible to determine the lexical stress position.

SOME OBSERVATIONS ABOUT THE LEXICAL STRESS IN ITALIAN

An electronic Italian dictionary containing 56,097 lemmas was used to determine the statistical distribution of the lexical stress in Italian. It gave the following results:

- 3.26% words with stress on the final vowel;
- 73.60% words with stress on the penultimate vowel;
- 22.84% words with stress on the antipenultimate vowel;
- 0.30% words with stress on the last vowel but three.

It was noted that the lexical stress in an Italian word can only fall on one of the last four vowels, as in "*àquila*" (eagle), whereas in inflected verbal forms the stress may sometimes fall on the fifth last vowel, as in "*chiàcchierano*" (they chat). Finally, in the verbal form with enclitics it can even fall onthe seventh last vowel, as in "*àuguraglielo*" (wish him that). This is due to the fact that in Italian a verbal form is composed of the following parts:

verb = (*prefix*) + *stem* + *flexion* + (*enclitic*)

- a *prefix* is not stressable and is optional;
- a *stem* may or may not be stressable and is always present;
- a *flexion* may or may not be stressable and is always present;
- *enclitics* are never stressed and are optional.

The morphological analysis of the last example is:

àuguraglielo = *àugur–a–glielo* = *stem* + *flexion* + *enclitic*

CONSTRUCTION OF RULES

To determine the set of rules it was necessary to generate a dictionary of Italian forms. An electronic Italian dictionary (Il nuovo Zingarelli minore, Zanichelli) containing over 44,000 headwords was used to generate inflected forms, giving feminine and plural forms for all nouns and adjectives; a typical entry was:

dottòr –e, –i, –èssa, –èsse; (doctor).

It was expanded by an automatic program into:

dottòre (singular, masculine)

dottòri (plural, masculine)

dottorèssa (singular, feminine)

dottorèsse (plural, feminine)

As far as the verb conjugation is concerned, each infinitive verb correspond to one of a given set of models: this allowed an automatic program to create all the inflectional forms of a verb (more than 50 flexions per verb were generated) starting from the infinitive form. A total of 110 inflectional models were created and used to generate the stressed inflectional forms of all regular and irregular verbs (more than 6,000 infinitive verbs were conjugated giving more than 300,000 inflected verbs).

The resulting lexicon, containing over 460,000 stressed forms, was subsequently broken down into a series of lists, one for each separate stress pattern (each form being stored with its spelling reversed, i.e. right to left).

These lists, containing words with the same ending and the same stress position already grouped together, were used to create the set of rules and its exceptions.

To give a clearer idea of what establishing a rule involves, let us provide an example: we could see from the lexicon that there were 63 words ending in "*–fera*", "*petrolìfera*" (oil) for example, with the stress on the antipenultimate vowel; so we created the rule "if the word to be stressed ends in "*–fera*" the lexical stress is to be put on the antipenultimate vowel". We subsequently observed from the lexicon that there were 17 words that contravened that rule because they ended in "*–fera*" but their lexical stress fell on the penultimate vowel, so we treated them as exceptions. These were the word "*bufèra*" (storm) and 16 words ending in "*–sfera*" as "*atmosfèra*" (atmosphere) for example. Instead of treating these 16 words as single exceptions we expressed them in the form of a rule, which then became a counter–rule. The final rule, valid for any word ending in "*–fera*" was: if the word is "*bufera*" or ends in "*–sfera*" then the lexical stress is to be put on the penultimate vowel, otherwise it should be put on the antipenultimate vowel.

A rule is a pair <*string, number*> (<"*fera*", 3> in the example above) where:

- *string* is a sequence of characters: a word matches a rule if it ends with *string*;
- *number* is the ordinal number, from right to left, of the stressed vowel (if the rule is matched and the word does not belong with the exceptions).

Each exception is a pair <*string, number*> (<"*–s*", 2> and <"*bu*", 2> in the example above) where:

- *string* represents the exception: counter–rules are the ones with the minus sign, the others are the single exceptions;
- *number* is the ordinal number, from right to left, of the stressed vowel (if the word is matched by the exception).

As can be seen, the whole words "*–sfera*" and "*bufera*" do not need to be put with the exceptions because they are looked for only if corresponding rule is matched (i.e. the word ends in "*–fera*"). The search mechanism is a simple pattern matching algorithm.

The total number of rules created is 86, the number of counter–rules is approximately 600 and the single words total approximately 1,000.

SPECIAL CASES

A large number of verbal forms with enclitics were considered. These are atonic pronouns added to the verb in a single word (i.e. "*dìmmelo*", tell me that). Rules take into account the fact that clitics are never stressed and the stress does not shift from its original position. Enclitics are recognised as such and separated from the verb which is subsequently stressed according to the rules.

The rule method is based on single word analysis, therefore it cannot process words having two different stress patterns, such as "*ancòra*" (still, yet) adverb, and "*àncora*" (anchor) noun. A local analysis is performed to try to resolve these ambiguities; each ambiguity corresponds to a procedure which performs a grammatical analysis on the text (different ambiguities can use the same procedure).

Let us consider for example the words "*tùrbine*" (whirlwind) a masculine singular noun, and "*turbìne*" (turbine) a feminine plural noun; in this case the procedure corresponding to this ambiguity attempts to determine the gender and/or number of the word by analysing the

words preceding the homograph. A data-base containing functional words with the indication of the gender and the number is used.

There are cases where not even a full syntactic analysis of the text is able to resolve an ambiguity. Let us consider for example the phrase

"La formica è utile": it could means

"La formìca è utile" (the ant is useful) or

"La fòrmica è utile" (formiea is useful).

In fact, both *"fòrmica"* and *"formìca"* are feminine singular nouns and so you can use them in the same syntactic context; a semantic analysis alone can resolve this ambiguity.

When an ambiguity cannot be resolved by the corresponding procedure, the most frequent stress pattern is adopted (for example the adverb *"ancòra"* is more frequent than the noun *"àncora"*).

Another very important aspect in the lexical stress assignment is the necessary exclusion of those very frequent cases where the lexical stress, which in an isolated word is always pronounced, is omitted in a phrase context. For example in the phrase *"ti sento ma non ti vedo"* (I can hear you but I can't see you) the two verbs are the only words which are pronounced stressed. To take into account this aspect, lists of functional words (articles, prepositions, pronouns, negations, conjunctions, etc.) are used to determine whether a word is to be unstressed.

PERFORMANCE

The method described was implemented with a finite automaton and an in-memory exception mechanism in C-Language. It succeeds in positioning the lexical stress correctly for almost every word in the electronic dictionary used (just some foreign words and very infrequent words were not treated like exceptions).

When an Italian text is analyzed, the performance of the method will vary according to the nature of the text: if the words of the text are words of the Italian lexicon then the results of the automatic stress assignment will be 100% error-free; if the text contains words outside the lexicon (like for example foreign words or surnames) the error rate will increase.

The set of rules was also applied to the 6,000 most frequent Italian surnames (which cover about 50% of all Italian surnames) and on the 600 most frequent Italian first names (which cover about 90% of all Italian first names). It gave an error rate of 6% and 4% respectively.

A PC/386 20MHz was used to evaluate the execution time: about 500 words per second were stressed. We found that the time taken by the program to stress a text is less than 1% of the time taken to "read" (i.e. to pronounce) the same text by a human being. This means that a text-to-speech synthesis system using this module has more than 99% of the human reading time to work in real time.

The code size of the executable program is about 40 Kbytes: about 50% of this space is taken up by the exceptions.

CONCLUSION

This paper has described a method of coding a dictionary for lexical stress assignment in Italian words.

An Italian lexicon containing inflected forms of all nouns, adjectives and verbs was used to establish the set of rules. It can be easily extended to other areas, such as surnames and geographical names (which were not present in the lexicon used to generate the rules).

This method works well and is currently used to improve the quality of the Italian text-to-speech synthesis system developed at CSELT [6]. In fact, lexical stress plays an important role in the correct pronunciation of words, and has a fundamental function in prosodic rule application.

REFERENCES

[1] Balestri, M.; *"Localizzazione automatica dell'accento lessicale in Italiano"* (in Italian), degree thesis on Computer Science, Univ. of Turin, 1989.

[2] Sandri, S.; Vivalda, E.; *"Automatic stress assignment for Italian text-to-speech synthesis"*, CSELT Technical Reports, Vol. VIII, No 3, pp. 213-216, June 1981.

[3] Martin, Ph.; *"Automatic assignment of lexical stress in Italian"*, Proceedings of the ESCA Workshop on Speech Synthesis, Autrans (FRANCE), September 1990.

[4] Delmonte, R.; *"L'accento di parola nella prosodia dell'enunciato dell'italiano standard"* (in Italian), Studi di Grammatica Italiana, Vol. X, pp. 351-394, 1981.

[5] Vivalda, E.; *"Italian text-to-speech synthesis: the linguistic processor"*, Olivetti Res. & Tech. Review n. 7/1987, pp. 47-60, July 1987.

[6] Nebbia, L.; *"Text-to-speech synthesis system for Italian: an overview"*, Paper presented at VERBA 90, Rome, Italy, January 1990.