EXPLOITING THE SECONDARY ACCENT
IN A PROSODIC MODEL FOR FRENCH SYNTHESIS

V. Pasdeloup

Institut de Phonétique, Université de Provence, 13621, Aix-en-Provence, permanent address: 8 rue du Château-Landon 75010 Paris, France

ABSTRACT

Unlike most other existing French speech synthesis systems where prosodic organization depends mainly on syntactic structure, we adopt an approach where the phonotactic criteria (syllable count, length of stress group etc.) are taken in account. We hypothesize that secondary accent takes a prominent part in stress structuration. Based on acoustic and perceptual analysis of read corpus, we present:

- a general description of secondary accent which includes: distribution rules, phonotactic constraints which rule its occurrence, and acoustic description;
- the linguistic and phonotactic criteria for prosodic structure analysis;
- a model for sentence prosodic organization.

1. INTRODUCTION

Traditionally, French language is described as having an accent at the end of lexical words (primary accent). However some sporadic works done in the last two decades mention the presence of a secondary accent which is not on the last syllable of lexical words and not relate to the rhetic or enunciatve accent (focalization) [1, 2, 3, 4].

Prosodic and perceptual analysis of 400 read utterances (5 speakers, 40 sentences, 2 repetitions) has been carried out [6].

2. SECONDARY ACCENT

Secondary accent should be taken into consideration in French accentual structure: the analysis of 400 read utterances shows that, in polysyllabic words, almost one accent out of three is a secondary accent [5, 6].

Secondary accent distribution is the following:

- on the first or second syllable (often first syllable starting with a consonant) of a word; this word is not necessarily marked by final primary accent (stressed syllables are coded I, unstressed syllables -);

un cher man ge con

- on the antepenultimate of a word ending with a primary accent;

une douce inflammation

- at the boundary of a morpheme in a polymorphic word:

c'est un ti / c'est un ti tien and

The secondary accent is located mainly at the beginning of a word or of a group. The occurrence of a secondary accent in a word depends on various constraints [5]:

- the accentual context (among others the number of unstressed syllables between secondary accent and the preceding or following accent);
- the word position in the sentence;
- phonetic nature of word first segmental unit (consonantic or vocalic);
- the number of syllables of the word;
- accentual strategy or speaker’s regional and individual characteristics.

We hypothesize that secondary accent have a regulatory function in the production of stress. Thus the stress pattern in a given sentence would be consistent with the biological and psychological standards of the production of rhythm [7, 8]: the secondary accent generally occurs so that a series of unstressed syllables never exceed the count of 4, and is distributed to an average of one accent (primary or secondary) per 3 syllables.

In the following sentence, if solely the final primary accents occurs, the series of 5 unstressed syllables would break the perceptive-motor unit [7]. Secondary accent on the word “desintoxiqué” would allow avoiding this lengthy series of unstressed syllables:

il voyait des dîtes sin dix quiés

The secondary accent is acoustically distinct from primary accent. The secondary accent is best described by a rising pitch movement (with medium amplitude) and a short syllabic lengthening (generally non significant); the primary accent being best described either by a rising or falling pitch movement (with medium or large amplitude) and by a variable syllabic lengthening (generally significant). Focalization accent differs from secondary accent by a larger amplitude of the pitch movement and a stronger intensity.

3. CRITERIA FOR PROSODIC STRUCTURE ANALYSIS

Almost all prosodic models developed during the last two decades are originally based upon a syntactic analysis to generate accentuation or intonation: prosody is considered to be congruent to syntax [9, 10].

Recently, in work on prosodic models, some importance has been given to phonotonic (also called eurhythmic) phenomena such as: principles of accentual alternation (strong/weak) and syllabic balancing, constraints on the number of consecutive unstressed syllables [11, 12, 13, 4, 14, 5]. However, these phonotonic constraints are generally considered only after having derived the prosodic structure from the syntactic structure, therefore compensating for the weaknesses of these models [11, 13].

We define a three-level prosodic structure for more details refer to [6, 16]:

- rhythmic sequences: major intonation groups;
- rhythmic words: minor intonation groups;
- accentual groups.

The demarcation of prosodic groups and the stress structuration are defined with linguistic and phonotactic criteria. For example, intonation units (such as rhythmic words or sequences) can be for phonotactic reasons:

- either, if they are too small, regrouped to constitute a larger unit [17];
- or, if they are too long, divided in two to constitute smaller units.

We present an exemple illustrating the steps required to generate acceptable prosodic structures for two sentences having the same syntactic structure but different syllables count.

(A) Determination of rhythmic sequences and rhythmic words

(Al) Sentence segmentation into linguistic rhythmic sequences (\(A\))

Firstly, in text-to-speech synthesis, we used punctuation information to delimit linguistic rhythmic sequences. Then, after...
linguistic and phonotactic rhythmic words have been defined (refer to (A2)), we reconsider this segmentation (refer to (A3)). As no punctuation occurs inside the following sentences, each of them consists only of one rhythmic sequence. 1/ La souris est vue par un affreux chat. 2/ Le pélican est dévoré par un gigantesque rhinocéros.

(A2) Sentence segmentation into rhythmic words (1)
(A2a) Demarcation before the grammatical words
1/ La souris [est vue] par un a fireux chat. 2/ Le pélican [est dévoré] par un gigantesque rhinocéros.

(A2b) Grouping of two (or more) small groups
Conditions :
- one is composed of 1* or 2* syllables; - the grouping doesn't exceed the count of 7* syllables.

(A2c) partition of large group
Conditions :
- one is composed of more than 6* syllables; - the new group resulting from the partition must be composed of at least 3* syllables. (*These coefficients can be varied according with the speaking rate or the speakers individual characteristics). 2/ Le pélican [est dévoré] par un gigantesque rhinocéros.

(R2) Primary accents location
The last stressable syllable of each rhythmic word gets a primary accent. 1/ [La souris est vue] par un affreux chat. 2/ [Le pélican est dévoré] par un gigantesque rhinocéros.

(B2) Secondary accents location (B2a) Application of the rule of the sentence first accent
First stressable syllable of the first rhythmic word gets an accent. 1/ [La souris est vue] . . . . . . I 2/ [Le pélican est dévoré] . . . . . . I

(B2b) Secondary accent rules
These aim at effecting a secondary accent in conformity with the distributional rules so that : - a series of unstressed syllables never exceed the count of 4*; - the ratio of the total number of accents (primary and secondary) to the number of unstressed syllables in the sentence is included inside a range of 1/2 to 1/4 (average 1/3 : one accent per 3 syllables).

- each new group resulting from the partition must be composed of at least two rhythmic words (linguistic or phonotactic). 2/ [Le pélican] [est dévoré]/ 1 2 4 [par un gigantesque rhinocéros]/ 1 2

(B) Automatic positioning of accents
The intonation where the accentual rules are applied is the rhythmic word (phonotactic or linguistic). The accentual rules define the distribution and the occurrence of primary and secondary accents. The secondary accent occurrence inside the rhythmic word is based upon the phonotactic constraints described in sections 2-3 above.

The stress clash rule is valid inside rhythmic words : each time an accent occurs, the preceding and following syllables inside the same rhythmic sequence cannot be accented.

Firstly, all unstressable syllables (grammatical words, antepenultimate of lexical words composed of at least 3 syllables) are defined.

The accentual structure of sentence 2d has the characteristics of the journalistic speech style (systematic presence of a secondary accent at the beginning of a word). The accenstual structure of sentence 2a will better match the characteristics of slow speaking rate - reading for example - (lower ratio of number of accents per number of unstressed syllables). This prosodic model takes into consideration both linguistic and phonotactic constraints and not only, as in the past, linguistic constraints. Therefore two sentences having the same syntactic structure will composed of a different number of syllables will not be given the same prosodic structure.

5. CONCLUSION
Prosodic structure seems to be the result of a compromise between universal non-linguistic constraints (of biological and psychological type) and linguistic constraints relating to each language.

According to our hypothesis, in stress production, secondary accent has a phonotactic and linguistic function (demarcation of units), whereas primary accent has only a linguistic function. However further research on other aspects of phonotactic constraints in prosodic structuration is required.

ACKNOWLEDGMENTS
This research was supported by a research grant (Ministère de la Recherche et de la Technologie) and by a training grant (Réseau Francophone des Industries de la Langue). I would also like to thank Doctor D. O'Shaughnessey we gave me the opportunity to work in the Montreal INRS-telecom laboratory.

REFERENCES