

Duration and Energy indices : a Convergent Deep Organization of Grouping and Phrasing

Geneviève Caelen-Haumont
 Institut de la Communication Parlée, CNRS URA 368,
 INPG et Université Stendhal,
 46 avenue F. Viallet, 38031 Grenoble Cédex 1 France

ABSTRACT

The aim of this paper is to present the results of a study on duration and energy indices in a context of text readings with tasks. This study is not traditional because the data are not analyzed in the respect of the prosodic organization of all the linguistic units which compose the phrases. In fact the study is based on the deep relative organization of duration and energy values of the lexical items (internal and external relationships) at different phrase levels. On the whole the study takes into account 10 indices.

INTRODUCTION

Since a few decades, many papers have been devoted to duration organization. Whatever the perspective in phonology and in acoustic studies, most of these works took in consideration the internal organization of phrases. For instance in the perspective of phonetical studies, Bruce (1985) identifies two functions of prosody, i.e. weighting (stressed / unstressed syllables) and grouping. This paper concerns with the function of grouping in french. The new perspective which is reported here considers 1° the deep acoustical organizations of prosodic indices 2° the internal and external lexical organization of these indices in different types of phrases 3° the grouping and the phrasing functions of 7 duration indices and 3 energy indices (*Energy* and not *intensity* because this parameter is extracted from a model of ear).

EXPERIMENTATION, AIMS AND METHOD OF ANALYSIS

The study of duration and energy organization corresponds to the second part of a more important analysis concerning an assessment of the relations between prosodic indices (Fo, energy and duration), and linguistic structuration (syntax, semantic) and pragmatic conditions (phases of discourse, reading constraints...). For this purpose, a 50-word text was elaborated : "*D'éminents biologistes et d'éminents zoologistes américains ont créé pour des vers géants un nouveau phylum dans l'actuelle classification des nombreuses espèces vivantes. Ces longs vers prospèrent sur le plancher marin des zones sous-marines profondes. Des sources thermales chaudes y maintiennent une température moyenne élevée.*" The experiment bore on 3 readings (3 tasks: 1° natural and intelligible reading 2° very intelligible reading 3° extremely intelligible reading, relevant for man-machine interaction) of the text by 12 speakers. A data base was elaborated from these 36 utterances, and was tagged with about 40 000 labels related to the various linguistic and prosodic analysis levels.

ENERGY AND DURATION INDICES

Definition

Duration and energy indices were converted in a four-tiered space, in order to discard surface phonetic information in the search of underlying structures. The minimal and maximal numerical values were automatically detected in the context of each discourse. Duration indices are based on all phonetic segments of the lexical words, and energy ones on their vowels. Note that the analysis being only based on lexical words, the study does not concern the group duration and the group energy, but the internal and external organizations (duration and energy) of the lexical items at phrase, sentence and text levels.

Duration indices are seven. Calculated in the entire word, they are: whole duration (WWD), mean syllabic duration (mSD), maximum syllabic duration (MSD), absolute value of syllabic duration range (Δ SD). In the last syllable of the word (and mono-syllables), the whole duration (WSD), mean phonetic duration (mPD) and the whole syllable duration + the following pause (SDP) are found. As for energy indices, there are few of them. Only calculated in the whole word, they are the absolute value of vowel energy range (Δ El), the mean energy (mE), and the maximum of vowel energy (MWE).

Forms

On a general point of view, energy and duration indices are constructed on the same pattern: the values (1 to 4) are arranged in an order which is ascending for duration indices (lengthening), and generally descending for energy indices (less loudness). Though the results were calculated in previous studies, on the whole structure of phrases, these present findings concerning the lexical duration structure of groups support these previous results (Caelen-Haumont, 1978; Padeloup, 1992).

This dynamic process concerns duration and energy parameters. Note that the step between two successive values may not be proportional, but the order is respected. The resetting is more or less important between the final value of the precedent group and the first one of the following group. So this organization enables us to give each group an internal structuration based on this progressive order, and an external one based on the breaks of this process. Table 1 below shows an example of this pattern issuing from the 2 first phrases of the text and 12 speakers.

For the analysis it is useful to distinguish between 3 types of phrase units : the micro-phrase (mP), the phrase of the most superficial level (P) and the macro-phrase (MP). P concerns the syntactic one which just dominates the level of words, and also pseudo-syntactic one. The pseudo-syntactic group is constituted for prosodic reasons when the syllable number of one of the two phrases is less than 5 (ex: "*ont créé pour des vers géants*" ...). MP is the combination of n phrases and mP is a part of the phrase (P) made up by two sub-units which are strongly linked in the domains of syntax and semantics: for instance the units of pseudo-syntactic phrase that we described just above, or as a second example, a noun with an adjective. This process always occurs in the case of a compound noun, but it is not the only case. It is also used when an adjective is syntactically and semantically closer from the noun than a second one. See for example table 1 below, speaker IN for DL index, and speaker BR for Em index.

Table 1. An example of DL index and EM index coefficients coded in a four-tiered scale. Data across 6 speakers, the first 2 phrases of sentence 1, task 1: "*d'éminents biologistes et d'éminents zoologistes américains ...*"

Idx	DL	DL	DL	DL	DL	DL	Em	Em	Em	Em	Em	Em
speakers	BR	CA	CO	FA	FO	IN	BR	CA	CO	FA	FO	IN
G1 é	2	2	2	2	2	2	3	3	4	3	2	3
b	3	3	4	3	3	3	2	2	3	2	1	2
G2 é	2	2	2	2	1	2	3	2	4	3	2	2
z	3	3	3	3	2	4	2	2	3	3	1	2
a	3	3	3	3	2	3	3	2	3	2	1	2

In order to explain the following results, we make a distinction between the notions of *lexical phrase segmentation* and *lexical phrase demarcation*. The function of the segmentation is grouping, while the function of the demarcation is phrasing.

Lexical phrase segmentation

The notion of phrase segmentation is used each time the boundaries of a series of values in a progressive order (admitting plateaux) correspond to the limits of a phrase (P) and macro-phrase (MP), whatever syntactic or not the combination of phrases is. For

instance, whatever the grouping of the verb with the precedent or following phrase is, those cases were considered relevant for phrase segmentation tasks.

In those conditions the mean results, evaluated on the number of syntactic and pseudo-syntactic phrases (i.e. 396 phrases) and concerning the number of lexical phrases corresponding to these segmentation rules, is 86%. That mean is calculated over the 12 speakers, the 3 tasks, and the 3 energy indices. Most of energy indices generally overrun 90% in sentences 1 and 2, but not in sentence 3 because of the great number of micro-phrases which occur in speaker utterances, and make the results drop drastically. On the whole and even in sentence 3, $|\Delta E|$ produces the best scores (89%). As it concerns duration indices (all the indices except SDP), the mean number of syntactic and pseudo-syntactic phrases which were correctly segmented, is 83%. Among all of them, the whole duration of the word (WWD) has the best score (93%). In sentence 3, just like energy indices, the results concerning the ratio of phrases and macro-phrases drop steeply.

On the whole, these results indicate that the internal and external organizations of prosodic indices in the space of phrase lexicon is thoroughly constructed. While the targets of minimum and the maximum loudness are well positioned according to the internal space of the lexical item, to the relations of lexical items in the phrase, and to the relations of phrases at the sentence level, the speakers simultaneously organize at the same time the structuration of duration. This structuration is organized on the basis of relative lengthening of each successive lexical word in the phrase.

Another interest is that of determining what the phrase and macro-phrase distribution is. For energy indices, the amount of phrases is 60%, the amount of macro-phrases is 40%. For duration indices, the amount of phrases is 81%, the amount of macro-phrases is 19%. Among the list of these indices, some are more analytic than others. For energy parameter, the best index in this respect, is $|\Delta E|$ (82%); for duration parameter, the best ones are two indices which are very close, MSD (85%) and $|\Delta D|$ (84%). For those last 3 indices, the ratio phrase/macro-phrase is greater than for the other indices (about 3 times). These results show once more that the targets of the maximum loudness (and minimum loudness) and the maximum duration in the lexical word are fundamental : the prosodic organization of the lexical dimension is centered on them and their relative places in this hierarchical structure within the phrase.

On lexical phrase segmentation, the sentence effect is very strong. For energy and over the 3 tasks, the mean proportion of phrases increase in a great expense for sentence 1 to 3 (56->58->91%), and the same phenomenon occurs for duration (62->62->92%). On the other hand, the task effect is insignificant. At the most, the analytic perspective decreases of 6%, for the benefit of macro-phrases constituted of 2 minimal phrases. The macro-phrases of more than 2 phrases are very rare (about 4%) for both parameters. So concerning lexical organization of duration and energy in phrases of various length, the general characteristics between indices are very close whatever the parameter may be.

Lexical phrase demarcation

Lexical phrase demarcation imposes a new condition, namely a syntactic one. Thus a macro-phrase is considered syntactic if 1° the ascending or descending pattern of the value index (including plateaux) corresponds exactly to its syntactic boundaries 2° the successive phrases which compose the macro-phrase have the same hierarchical level in the syntactic structure and if they are issued from the same father. In addition a new index is introduced, the last syllable duration plus the following pause (SDP).

In the context of phrase demarcation, it seems necessary to take into account the third linguistic unit, i.e. the micro-phrase. Let us consider the different cases where a micro-phrase was performed by speakers. Over 7 events in the text, 2 concern the case of pseudo-syntactic phrase, made up of two syntactic groups, one of them having too few syllables. So this micro-structuration takes into account syntactic organization. Over the next 5 cases, 4 are made up of compound-nouns or expressions the noun and adjective meanings of which are very dependent (*sources thermales, température moyenne* ...), and the last one is made up of two adjectives, one of them qualifying two coordinate noun

phrases. Thus one can say that all those phenomena are syntactic ones. They operate on less extended units, the relations of word dependency are stronger, but nevertheless syntax is always concerned. The other cases of micro-phrases which do not correspond to those cases are not counted as syntactic ones.

For the 3 energy indices, over the 12 speakers, the 3 tasks and the 3 sentences, among all the productions corresponding to syntactic phrase level, it appears that the phrases with the least extension are the more syntactic: 99% of decreasing or increasing values performed in the context of phrases correspond to a syntactic organization, and so are 83% of them corresponding to a micro-phrase, and 68% corresponding to macro-phrases. As for the best duration indices (WWD, Δ DI, mPD, SDP), the results are similar to energy ones: phrase level, 99%, micro-phrase level, 92%, macro-phrase level, 53%.

Over all sentences, tasks, speakers, and different phrase units, Δ EI appears to be the best index of phrase demarcation (91%), while for duration, the best indices are Δ DI (92%) and SDP (91%). From task 1 to 3, the ratio of phrases and micro-phrases with a syntactic organization of energetic values, is regularly increasing for the three indices (as the case may be, 3 to 18% more), while that of macro-phrases decreases (21 to 25% less). For duration, from task 1 to 3, the 4 best indices show a great steadiness for phrase level, a slight tendency to decrease for micro-phrase, and for all indices except SDP, a tendency for macro-phrases to increase (10 to 37% more). Then it seems that the function of energy indices is more analytic than that of duration indices when the speech rate becomes slower in tasks 2 and 3, with many pauses in addition for task 3.

It is interesting to note that DSP is the best index for macro-phrase demarcation when task 1 is running (76%) with relevant pause durations, but that its score becomes worse and worse (from 76 to 53%) as pauses get more numerous. On the whole, Δ EI seems to be specialized in the demarcation of the phrases of less extension (phrases and micro-phrases), and MWE is the more resistant index for the hardest tasks. On the other hand, the 4 best duration indices are all specialized in the demarcation of phrase level (99 or 100%), SDP both for micro-phrase (100%) and macro-phrase demarcations, though the score of the latter is not very high (66%). For energy and duration indices, the task has no significant effect: from tasks 1 to 3, the range spreads from 87 to 90%.

CONCLUSION

For both duration and energy parameters the two functions of grouping and phrasing exist. The grouping function is based on the search of cohesion in discourse. The phrasing function is built on the linguistic objectives of parsing.

The 7 duration indices and the 3 energy indices show a deep lexical organization very similar in their forms and functions. Within this complex (and at the same time, simple) prosodic organization, all phrases (and probably most of the micro-phrases) are demarcated. For the grouping function, the global duration of lexical items appears to be a sufficient cue, while the phrasing function requires within this global duration, more precise targets, such as those performed by SDP and Δ DI. Δ EI exerts both grouping and phrasing functions. For the phrasing function, the targets of minimum and maximum duration and energy in the lexical items are crucial, and most of the speakers control those targets with a great precision. Then the demarcative function is built, for both parameters, on the precise control by speakers of the minimum and maximum values. This control at both levels, i.e. energy and duration, is operated for each parameter, in the context of at least 2 intricated scales: the scale within the word, and the scale within the phrase.

REFERENCES

- G. Bruce (1985), Structure and Functions of Prosody, *Proc. of French-Swedish Seminar on Speech*, Guérin et Carré eds., Grenoble, pp. 549-559.
- G. Caelen-Haumont (1978, 1981), *Structures prosodiques de la phrase énonciative simple et étendue*, Thèse de doctorat de 3ème cycle, Université du Mirail, Toulouse, Hamburger Phonetische Beiträge, band 34, Hamburg Busk.
- V. Pasdeloup (1992), "Durée intersyllabique dans le groupe accentuel en français", *Proc. of 19èmes JEP*, Bruxelles, pp. 531-536.