

# F0 prominences (melisms) in French: a deeper insight about morphophonology

Geneviève Caelen-Haumont

International Research Center MICA  
Hanoi University of Technology - CNRS/UMI2954 - Grenoble INP  
(1 Dai Co Viet str. Hanoi, Vietnam)  
gcaelen@gmail.com

## Abstract

This study about the melism stuff (or melodic substance) is the continuation of previous ones since several years [for instance 1, 2]. Those studies dealing with a global analysis, this present one will examine in details the processes used by 4 speakers to highlight F0 prominences at the morphophonological level of the melisms. As the major interest of the study lays upon the repetition patterns, the differences between them and between speech strategies across speakers will be questioned.

**Index Terms:** melodic prominence, morphophonology, melisms, duplication, symmetry, tonal targets, tonal syllables.

## 1. Introduction

Since Bolinger [3], Chafe [4], Brown [5], many works in the domain of information structure have been undertaken in conversational speech to explore the domain of the prosodic word salience regarding the semantic and pragmatic approaches. On the other hand much less works concerning the melodic morphology of the word prominence have been done. It is nevertheless a domain of a great interest as it concerns not only linguistic analysis, but also automatic processing of speech (recognition and speech synthesis).

It is then the matter of this present paper, in the suite of the previous ones [1, 2]. For a precise description of the MELISM tool, the data bank, and the method of analysis, please refer to these papers. Let us simply precise that the analysis is concerning 4 female speakers, issued from the same family (from the daughter LR2, to her mother LR1, her grandmother SP1 and her great-grandmother HV1), so 4 generations on the whole, all living in the same house and working in the context of wine growing near Bordeaux.

## 2. About the data

The MELISM tool, running with MOMEL [6], is integrated under Praat software (<http://www.praat.org/>), and supplies a thorough analysis of the F0 values in 9 levels with a set of 9 symbols (and the related Hz and semi-tones values): Acute (= 9), Supra (= 8), High (= 7), elevated (= 6), middle (= 5), centred (= 4), bottom (= 3), infra (= 2), grave (= 1). It is important to know that this scale was computed on the overwhole of each speaker speech, and not only on the melism data. MELISM is used to automatically code absolute levels corresponding to fractions (on a logarithmic scale) of the speaker's pitch range. These levels are corresponding to the turning points of the modelled F0 slopes, and are named the *tonal targets*. In our application, in the frame of each word previously segmented and labelled under the Praat TextGrid tool, MELISM segments and labels these successive targets, and the space between 2

successive targets receives the name of *tonal syllables*. So the melismed word with its suite of tonal syllables takes for instance this structure: speaker LR1, *vrai ~ true* /bc cH HS SS Sm/.

All the words which presented the highest values (/A, S, and H/ this last one under some restrictions) were extracted as melisms and listed with their linguistic context in an xls data bank file, with their prosodic features.

## 3. The F0 structuration of the melisms

On the ground of the 400 melisms (100 x 4 speakers) of the data bank, through the great variability of the melism forms and of their tonal syllables (from 1 to 9 tonal syllables at the most per melism with an average of 3.55 syllables), we could extract prototypes [1] of a rising / falling melism structuration (see below Figure 1).

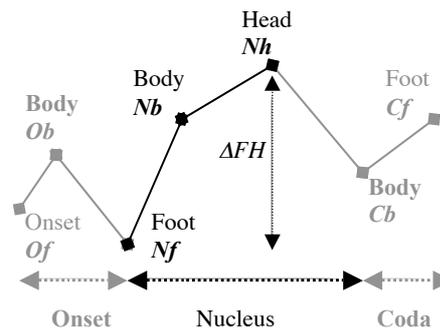


Figure 1: *The rising melism structure prototype.*

In such a prototype, only the Nucleus foot (Nf) and the Nucleus head (Nh) are compulsory (i.e. always present). The other are optional, but may be present most of the time (the Nucleus Body Nb, the Coda foot Cf, and the Onset foot Of), or often absent (the Coda Body Cb and the Onset body Ob). The aim of this present study is to get a deeper insight of the melism stuff. For instance some questions arise such as: are the successive targets always different or not? Have generally great contrasts between their successive values? For this analysis, we shall focus not on the tonal targets but on the tonal syllables: though the values are the same, the tonal syllables are more interesting because they supply a dynamic perspective. It is that perspective that we'll keep for the most part of this present analysis.

### 3.1. Data overview

Before examining the melism stuff in detail, we will first give some indications about the main melisms features.

HV1	367	LR1	386
SP1	343	LR2	325

Table 1: *Number of the tonal syllables*

Calculated respectively over 467 (HV1), 443 (SP1), 486 (LR1) and 425 (LR2) tonal targets, the Table 1 above presents the number of tonal syllables per speaker (400 melisms on the whole, and 100 per speaker). LR1 who is LR2's mother, supplies the most numerous tonal syllables, and LR2, the least ones. The average for the 4 speakers is 3.55 tonal syllables. As it can be noted, the age factor has no effect.

Lowest / highest targets	Nf	Nh	Range Nh-Nf
HV1	3.33	7.6	4.27
SP1	3.28	7.82	4.54
LR1	2.92	7.35	4.43
LR2	3.38	7.77	4.39

Table 2: The average height (semi-tones) of the lowest (Nf) and highest (Nh) targets in the melism, and range.

Another source of information is based on the compulsory targets which are composing the melism: the Nucleus foot (Nf) and the Nucleus head (Nh). The Table 2 above shows that for Nf, the span spreads out from 2.92 to 3.38 average semi-tones, and for Nh, from 7.35 to 7.82. Though the LR1 range is not particular, this speaker is still having the lowest heights for Nf and Nh. The following table (Table 3 below) explains why.

The highest targets	H = 7	S = 8	A = 9
HV1	57	26	17
SP1	45	28	27
LR1	71	23	6
LR2	42	39	19

Table 3: The number of the tonal syllables per speaker across the highest values of the targets

In fact LR1 presents the greatest number of melisms which culminate at 7, which number is balanced conversely by the smallest one for the melisms culminating at 9.

We can thus conclude that 1° there is no effect of the age on the results 2° LR1 is a speaker who seems to behave differently from the other ones, presenting the greatest number of tonal syllables in her melisms. As she presents the fastest speech rate for the speech and pauses (4.8 lexical syllables /s) such an effort might be balanced by a lesser effort concerning the upper F0 values.

### 3.2. Difference between two successive tonal targets

Another domain of interest concerns the difference calculated in semi-tones between two successive targets. This study is still grounded on 1424 tonal syllables composing the 400 melisms.

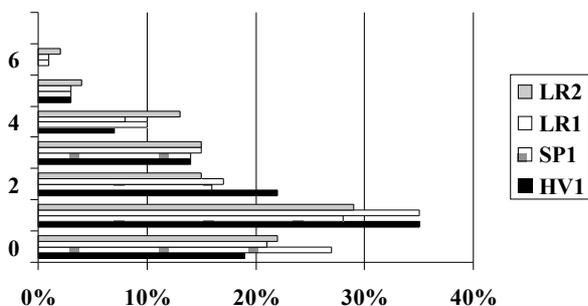


Figure 2: Distribution of the tonal syllables across the different steps between two successive targets as a function of speaker (from the eldest (HV1) to the youngest one (LR2)).

Over 1424 tonal syllables samples, the figure 2 above shows that globally the speakers do not differ much in the way of producing two successive targets: 64% of the tonal syllables differ by 0, 1 or 2 tonal levels (over 8 possible on the whole).

It means that the tonal transitions are not abrupt and moreover that they are following a main global F0 pattern (ascending slope or descending one). It means also that the gap between Nf (the lowest target) and Nh (the highest one) rarely concerns two successive targets, or if so, that it does not contrast more than with 3 or 4 levels.

### 3.3. The processes to fill up the melism structure

In a previous study [1], we pointed out 2 main repetition processes used by the 4 speakers to fill up the melisms: *duplication* and *symmetry*, and we explained their global mechanism. In this study, we wish to expand this analysis and make comparisons between the 4 speakers.

First of all, a general remark is necessary: the number of duplications in a melism is a function of the number of the tonal levels chosen for splitting up the speaker melodic range. Thus for 9 tonal levels the number of tonal syllables duplications is greater than for 4 levels, and with 9 levels, some interesting details can come to light. In those conditions the main interest for this study is based on the nature of processes used by the 4 speakers in the frame of these 9 levels, and the comparisons between them.

#### 3.3.1. Main characteristics

The Figure 3 below presents the relative percentages of the 400 melisms (100 melisms x 4 speakers) with and without any repetition (symmetry and duplication).

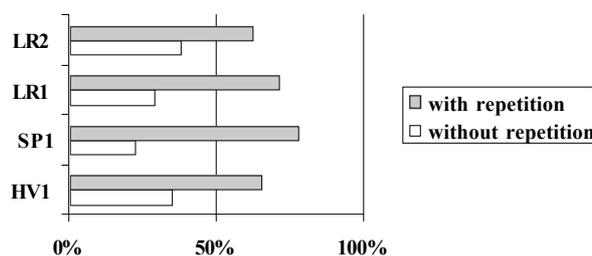


Figure 3: Relative percentages of the 400 melisms with or without repetition (duplication and symmetry), as a function of the 4 speakers.

One can see that the melisms with repetition are the most numerous (276 over 400 on the whole i.e. 69%, and 62% to 78% across the 4 speakers, the youngest one, LR2, producing the less numerous ones).

#### 3.3.2. Study of the Duplication process

Composing the duplication process, according to the place where this phenomenon occurs, 3 labels were given:

- **Appeal (A)**: when the duplication occurs at the beginning of the melism, for instance HV1, *midi ~ noon /bb bH/*,
- **Echo (E)**: when it occurs at its end, ex: LR1, *régulièrement ~ regularly /ig gS SS/*,
- **Murmur (M)**: when it occurs in-between, ex: LR2, *remarquer ~ to note /ib be ee eH/*.

It can be noted that Appeal, Murmur and Echo duplications may occur in the same melism (for instance SP1, *autres ~ other /ii ic cc cH HH/*): in fact 240 melisms are sharing 307 duplications. The next figure (Figure 4 below) shows the distribution of the 307 target duplications (240 melisms

on the whole, i.e. 60% of the 400 melisms), along each kind of duplication (Appeal, Murmur, Echo) as a function of the speakers. The Echo (final) position occurs generally a little bit more often for the eldest speakers than the Appeal one, while the Murmur one occurs less frequently for everybody. The Appeal position offers a number rather stable across the speakers.

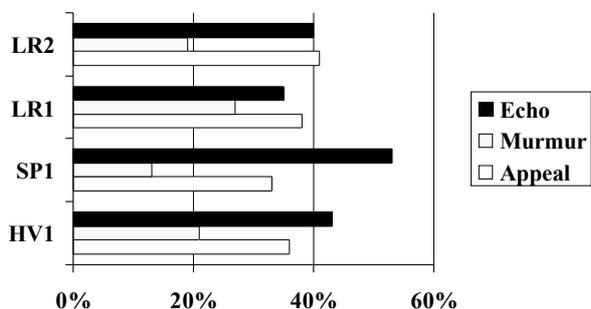


Figure 4: Duplication - Relative percentages of the number of melisms (with 1, 2, 3 target repetitions) as a function of their position across the 4 speakers

The figure 4 shows also that, respectively to other people or to other positions, the speaker SP1 places the greatest amount of her melisms in the final (Echo) position, and conversely, the smallest amount in the middle one (Murmur). As LR1 shows the weakest variations among the 3 positions, HV1 and LR2 supply a mean behaviour between these two opposite trends. From those data, we can still conclude that the effect of age cannot be observed.

If now we consider the distribution of the 175 melisms (44% of the 400 melisms) with only one duplication (thus the number of melisms equals the number of duplications) placed therefore either on Appeal or Murmur or Echo position, we remark that there is not much variation between these two sorts of data: the main difference stands on a greater opposition between SP1 and the 3 other ones. If HV1, LR1 and LR2 behave in the same way (and also SP1 for the Appeal position), using about the same number of one duplication melisms whatever their position, SP1 overshoot the other speakers' threshold in the final part (Echo) with a greater number of melisms (60% of all of hers). SP1 is thus presenting the greatest dispersion among the speakers. As before, we cannot observe any effect of the age on the results.

Now the following figure (Figure 5 below) presents the distribution of melisms according to the combination of 2 target duplications per melism. On the whole 63 melisms (16% of the 400 melisms) are using such a melodic pattern. Among them, the 3 possible combinations are used:

- Appeal + Murmur: ex.: LR2, *étude ~ study* /**bb** bm **mm** mS Sc/,
- Appeal + Echo: ex.: HV1, *pain ~ bread* /**bb** bm mH **HH**/,
- Murmur + Echo ex.: SP1, *difficile ~ difficult* /cm mc **cc** cS **SS**/.

The Figure 5 below shows that the most frequent pattern of the double duplication is Appeal + Echo, i.e. concerning the opposite boundaries of the melisms, with an exception: LR1. Though the average percentages of melisms for Appeal + Murmur (24%) and Murmur + Echo (22%) for the 4 speakers on the whole are close, they hide nevertheless some substantial differences in the speakers behavior.

In the condition where the data are not so numerous (63 on the whole), it can be seen however that while the two eldest ones, HV1 and SP1, have quite the same distribution of their melisms across the different localisations of the duplications, the youngest ones differ. The youngest one,

LR2 is not so different from the eldest ones, but she presents, conversely to her grand-mother and great grand-mother, an inverted proportion of Murmur + Echo (very few) relatively to Appeal + Murmur. As for LR1, she is quite different from the others, the Appeal + Echo duplication having the same rate than the Murmur + Echo one, and the Appeal + Murmur one presenting the greatest proportion of melisms. Moreover LR1 presents the smallest variation about the distribution of the melisms.

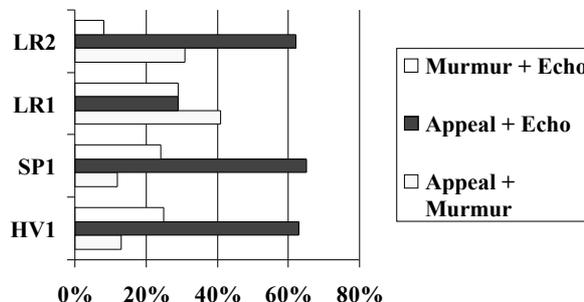


Figure 5: Duplication - Relative percentages of the number of melisms with 2 duplications, as a function of the position of the duplications across the 4 speakers.

Let us just mention that very few melisms present 3 pairs of duplication, only 2 examples: SP1, *autres ~ others* /ii ic **cc** cH **HH**/; LR2, *huitres ~ oysters* /cc cA **AA** **AA**/.

### 3.3.3. Study of the Symmetry process

The second process for filling up the melisms consists in using symmetry. On the whole, 76 melisms (19%) adopt such a kind of pattern. Among them, some are also using duplication. There are several kinds of symmetry among our data. Symmetry may either be simple with two syllables (LR1 *contact ~ contact* /**cm** **mc** cH He/), which may alternatively embed another one (HV1, *soirs ~ evenings* /cm **mH** **HH** **Hm**/), or spreads over three syllables (LR2, *pas ~ not* /**HS** **SH** **HS**/), or four (SP1, *collection ~ collection* /**cm** **mc** **cH** **Hc**/) in a double symmetry process.

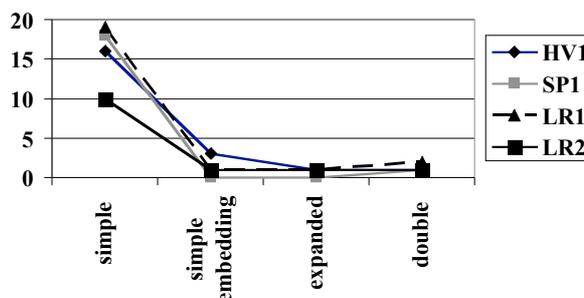


Figure 6: Symmetry - Distribution of the 77 melisms depending on the kind of symmetry.

The figure 6 above shows clearly that speaker's behavior is nearly the same. There is a great amount (63 or 16% of the 400 melisms or 83% of all the symmetry melisms) which only use a simple symmetry (ex: HV1, *sac ~ bag* /bm mS **SH** **HS**/).

The 13 remaining symmetry items may:

- embed a tonal syllable (5 examples, 3 speakers, HV1, LR1, LR2). In such a case the tonal syllable (TS) embedded is always a duplication, here /SS/ for the example above. The embedding symmetric syllables vary a lot, they may be constructed either on the pattern Appeal + TS + Murmur, Appeal + TS + Echo, Murmur + TS + Echo, or Murmur + TS + Murmur.

- be expanded over 3 tonal syllables (3 examples, 3 speakers), the item 1 and 3 being the same (duplication).
- be double (5 examples too, 4 speakers, for instance the longest melism: LR1, *draconien ~ draconian /ib bg gb bi im mm mH Hm mm/*). This example is a very complex one, 1° as there exists a double symmetry */bg gb/* and */mH Hm/*, and 2° the first one is embedded in another one */ib bg gb bi/*. Note also that the second one is also included inside a double duplication */mm ... mm/*, which is also a symmetric pattern.

If now we focus on one symmetry melisms (63 items), we remark (Figure 7 below) that the behaviors are quite different among speakers.

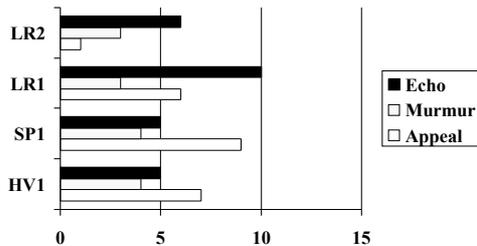


Figure 7: *Symmetry - Distribution of the 63 melisms providing a simple symmetry (2 tonal syllables), as a function of the position of the duplications across the 4 speakers*

Note that as the symmetry process is based on 2 following tonal syllables, “Appeal” “Murmur” or “Echo” corresponds in fact respectively to the first syllable location of the pair. The eldest speakers (HV1, SP1) tend to place their symmetry process at the beginning of the item (Appeal position), while the youngest prefer the final one (Echo). On the whole HV1 and SP1, the oldest speakers, present the closest patterns. LR1 present the greatest difference respectively to other speakers. In the whole, the position which contrasts the best the speakers is the Appeal one while Murmur one contrasts the least.

### 3.3.4. Duplication and Symmetry

As Symmetry and Duplication are both a process of repetition (of the same / different targets), it is thus possible to gather the data, especially in the context where the repetition is simple (2 tonal syllables). 238 melisms are then concerned.

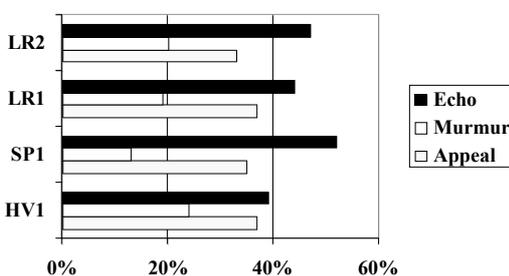


Figure 8: *Symmetry and Duplication - Distribution of the 238 melisms providing a simple repetition (2 tonal syllables) across the 4 speakers as a function of the repetition position*

In the figure 8 below, we note several interests: 1° that the strongest differences previously noted are smoothed. This means that each speaker manages the processes of duplication and symmetry of her own as their specific production strategy.

The process of balance is clear for LR2, SP1 and HV1, though it is different according to the age of the speakers:

thus LR2 balances the wealth of the Appeal duplication melisms by the rareness of the Appeal symmetry ones, while conversely, SP1 and HV1 balance the wealth of the Echo duplication melisms by the rareness of the Echo symmetry ones.

We remark then that across speakers 1° the Echo position provides the greater number of repetition processes, 2° the Appeal one is quite stable 3° the Murmur one is the least concerned with the repetition process: it is justly a dynamic space between 2 extrema, the lowest and highest F0 values.

## Conclusion

Since the previous studies, the present one proposes a deeper insight into the melisms morphology. To filling up the melisms stuff, the speakers are using 2 main repetition processes: duplication and symmetry.

The overall behaviour of the 4 speakers is nearly the same as they are using the same processes. While speaking, everybody has in fact to deal with the same problem: sparing his/her own energy while being understood at the best at the phonetic level as well as at the linguistic and pragmatic ones. Now making a big F0 contrast in one word is costly, as giving an F0 prominence to a word is a matter of duration and big contrast. Duration is concerned with duplication and symmetry, and contrast with the opposite short targets, Nf and Nh.

It seems that the Prosody level in comparison with the phonetic one, is kept more unharmed from variations in time. So when the variations oppose the eldest ones to the youngest ones, this opposition does not concern with the melism morphology, but because of a greater effort to supply, on the one hand with features and quantity, and on the other hand, with local speech strategy.

So if the overall process is the same among speakers, nevertheless the strategy about the management of duplication / symmetry differs between them. The speaker LR1 is varying the most: she provides the greater number of melisms ending at the level 7 (so the smallest one at the level 9), the same for the lowest average tonal height for Nf and Nh, the distribution variations about the Appeal, Murmur and Echo positions are the most balanced. She provides also the greater number of the tonal syllables within the melisms, so the greater number of the melodic variations, eventhough with a smaller contrast (undershoot), and in the same time she supplies also the fastest speech flow (4.8 lexical syllables/ second for speech and pauses). Thus one may conclude that these characteristics are an effect of her speedy speech flow. This matter might provide an interesting future development of research.

## References

- [1] Caelen-Haumont, G., “Emotion, emotions and prosodic structure : an analysis of the melisms patterns and statistical results in the spontaneous discourse of 4 female speakers from 4 generations”, book chapter, in S. Hancil [Ed], *The Role of Prosody in Affective Speech*, Peter Lang, li97, 95-138, 2008.
- [2] Caelen-Haumont, G., “Melodic prominences structures: exploring to what extent the speaker variability is spreading”, *Proceedings of the Acoustical Society of America (ASA)*, Paris, 3513-3518, 2008.
- [3] Bolinger, D.L., “Accent is predictable (if you’re a mind-reader)”, *Language*, 48: 633-644, 1972.
- [4] Chafe, W., “Givenness, contrastiveness, definitiveness, subjects, topics, and point of view”, *Subject and Topic*, 25-55, 1976.
- [5] Brown, G., “Prosodic structure and the given/new distinction”, in A. Cutler & D.-R. Ladd [Eds], *Prosody: Models and Measurements*, Berlin, Springer Verlag, 67-77, 1983.