

# Tonal Template for Background Information: the Scaling of Pitch in Utterance-Medial Parentheticals in French

Zsuzsanna Fagyal

French Department  
University of Illinois, Urbana-Champaign  
zsfagyal@uiuc.edu

## Abstract

This paper shows that utterance-medial parentheticals elicited in controlled conditions in French have a typical tonal template, in which pitch is scaled down to some extent in non-intonational phrase-final syllables. Downscaling is implemented primarily as pitch register lowering, while pitch range compression seems speaker-dependent. Results also show remarkable inter-speaker consistency in lowering pitch in edge syllables of minor prosodic units. It is suggested that parentheticals represent a case of ‘extreme’ backgrounding of information, which leads to a more uniform treatment of otherwise gradient pitch range variations.

## 1. Introduction

Intonational studies in a wide variety of languages had demonstrated that the scaling of pitch is largely influenced by pragmatic factors. Systematic variations in pitch range were related to discourse structure [1], information status [2], speakers’ emotions [3], and paralinguistic factors, such as speaking loudly [4].

Parentheticals (henceforth, PAs) seem to particularly well illustrate the role of pitch in signaling information status, since syntax does not account for their formal representation. From a strictly syntactic point of view, PAs are independent syntactic constituents inserted in another constituent, with which they bear no obvious formal relationship [5]. Easily replaced by a vocative, such as *mon ami* ‘my friend’, or even an utterance, such as *m’a dit le facteur* ‘the mailman told me’, the adverb *normalement* ‘normally’ in (1) is unattached to the highest node of its host sentence, and is considered a PA, because it is uttered as a hearer-oriented comment conveying secondary information.

- (1) l’adresse, *normalement*, c’était 2 rue Boulard ([6]:159)  
“The address, normally, [it] was 2 Boulard Street”

Secondary or ‘background’ information is commonly associated with lesser pitch range variations. Bolinger [7], for instance, observed that the relative pitch height of tonal targets is “compressed within the narrow range of the parenthetical” in English (p.34). Studies in pragmatics of intonation also report such pitch range compression in PAs in French [6][8], but they focus on the general lowering of pitch with respect to the rest of the utterance. This lowering is modeled by using distinct ‘levels’ (henceforth, registers) within the speaker’s overall ‘span’ [9] of pitch variations. Mertens [10], for instance, defines a separate tonal inventory of stressed and unstressed syllables (h-, H-) in *registre super-bas* ‘supra-low register’, reserved for both *incises*, i.e. utterance-medial PAs, and *appendices*, i.e. utterance-final

PAs. In Morel and Danon-Boileau’s [6] pragmatic model of French intonation, utterance-medial and final PAs are treated separately because of their different pragmatic meanings. However, there is considerable overlap between the two constituents in their phonetic realizations. One of the common features is low pitch register, referred to as *la plage basse* ‘the low range’. It corresponds to the lowest portion of the speaker’s overall *plage mélodique* ‘melodic range’. It is reserved for comments that are not oriented towards the addressee *co-énonciateur*, and are, therefore, withdrawn from mutual consideration by the speakers. According to this model, PAs uttered with such pragmatic meaning are characterized by both pitch range compression (expressed in terms of *absence de modulation de f<sub>0</sub>*), and register lowering or *abaissement de f<sub>0</sub>* (p.59).

The aim of this study is to determine whether previous conclusions on pitch range in PAs also apply to utterances elicited in controlled conditions. As briefly illustrated above, studies on the phonetic realizations of PAs are numerous, but those carried out in comparable discourse settings, and similar phonetic and phonological contexts are scarce in English [11], and seem inexistent in French. Inspired by Bolinger’s proposal that tonal targets in PAs undergo pitch range compression, the following pilot experiment is using a phonological model of French intonation [12][13] to examine whether pitch in different tonal targets is, indeed, compressed in PAs in French. It also investigates whether register lowering, if existent, is also uniformly applied to all tonal targets. If, contrary to the null hypothesis, the data show no unified strategy in handling pitch range in utterance-medial PAs, the conclusions will yield support to the representation of pitch range as an essentially gradient phenomenon (see [14] and [15] for discussion).

## 2. Experiment

### 2.1. Corpus

Three native speakers, one male and two female, in their thirties, born and raised in Northern France, volunteered to participate in the experiment (cf. [16][17]). The subjects were first asked to read a fictional narrative (Panel 1), and were then told that they would read several sentences, which were variations of a statement about a character in the story. None of the sentences appeared in the narrative, which was intended to provide a discourse context for backgrounding or foregrounding of information in the utterances. In order to make sure the context was understood, the subjects were asked to summarize it. No other instructions were provided.

Then the subjects were presented with the target sentences printed on separate cards in different random orders. The sentences were composed of the syntactic constituents listed in square brackets in (2). They contained a

PA (2b) preceded by a Noun Phrase, (NP, 2a), and followed by a Verb Phrase, (VP, 2c). The length of each phrase was increased by adding one syntactic constituent to the right. NPs and VPs were of three, six and nine syllables. PAs were of four, seven, ten and thirteen syllables. Their total combinations resulted in thirty-six different utterances, which contained only voiced sounds.

*Panel 1: English translation of the fictional narrative used in the elicitation task.*

“Mama Lamana is a former nanny who used to work in the richest villas of Lima, the capital of Peru. She is a nice woman, but she is very talkative. She loves to tell secrets, for instance, about children she used to baby-sit, and who became celebrities as adults. The opera singer, Lavilla Marina, who likes to be called « The Diva », is one of most famous celebrities in Peru. She is one of the children Mama Lamana used to baby-sit. Mama Lamana tells us, among other things, that Lavilla Marina was not an easy child. She often fought with her little sister, Nina, from whom she stole different things.

In the following, you will read several versions of a sentence published by a journalist in Paris Match about this topic.”

Although all utterances had an approximate 1:3 ratio of high to low vowels to control for intrinsic pitch variations in the vowel, only tonal targets containing low and low-mid vowels were included in this study (see also [16][17]).

- (2) a. [[*La diva*] [*Lavilla*] [*Marina*]]  
           ‘the diva’ ‘Lavilla’ ‘Marina’  
           “‘The diva Lavilla Marina”,  
       b. [[*m’a dit*][*Mama*][*Lamana*][*des villas*][*de Lima*]],  
           ‘me told’ ‘Mama’ ‘Lamana’ ‘of the villas’ ‘of Lima’  
           “Mama Lamana from Lima’s villas told me”,  
       c. [[*me vola*] [*les babas*] [*de Nina*]].  
           ‘me stole’ ‘the babas’ ‘of Nina’  
           “stole the babas\* from Nina”. (\*French delicacy)

The subjects were recorded in a sound booth. They were asked to read the sentences twice in two sessions, in a neutral way, and at their average tempo. Given the difficult task of reading maximally similar segments, in case of mistakes or hesitations the entire sentence was repeated.

## 2.2. Measurements

Pitch tracks were interpreted in terms of high and low pitch targets, corresponding to levels of  $f_0$  peaks and valleys (pitch and  $f_0$  used synonymously). One  $f_0$  value was extracted manually at the highest peak or lowest valley in each syllable in the utterance. When the peak or the valley occurred on obstruents, the closest  $f_0$  value in the vowel was taken. Values were measured and computed in semitones to facilitate the interpretation of fine-grade pitch-level differences.

Contrary to usual practice,  $f_0$  values were not averaged across repetitions. This was due to the lack of narrow control over the pragmatic interpretation of utterances. Since speakers’ interpretation can vary from one reading to another, averaging  $f_0$  values could eliminate differences in pitch height

this study set out to investigate. Consequently, only  $f_0$  values from the first reading were used.

Syllables were divided in three prosodic positions, based on a previous study conducted on prosodic phrasing in these utterances [16]. Final syllables of major and minor prosodic phrases were referred to as Intonation Phrase-final (IP-final) and Accentual Phrase-final (AP-final) [12][13], respectively. Since the subjects produced neutral reading, all non-phrase final tonal targets, including potentially initial-accented syllables, were considered ‘AP-non final’.

Pitch register was expressed in terms of mean  $f_0$  values, which were expected to be significantly smaller in IP2s than in IP1s and IP3s. Pitch range compression corresponded to the maximum range of  $f_0$  variations (max  $f_0$ -min  $f_0$ ). (For alternative definitions of these terms in the literature, see [6].)

## 3. Results

### 3.1. Pitch register lowering

Figures 1, 2, and 3 show the mean  $f_0$  values ( $\pm$  one standard deviation) for all syllables and prosodic positions for the three speakers. NPs, PAs, and VPs are referred to as IP1, IP2, and IP3, respectively. One-factor ANOVAs computed for each IP show an overall significant interaction at  $p < 0.001$  between pitch height and prosodic position for all speakers. Within-group comparisons using Scheffe F post-hoc tests, however, reveal important differences. F values for comparisons between IPs are shown in Tables 1, 2 and 3. Statistically non-significant (ns.) or perceptually negligible ( $\sim$ ns.) differences within each IP are shown in the figures.

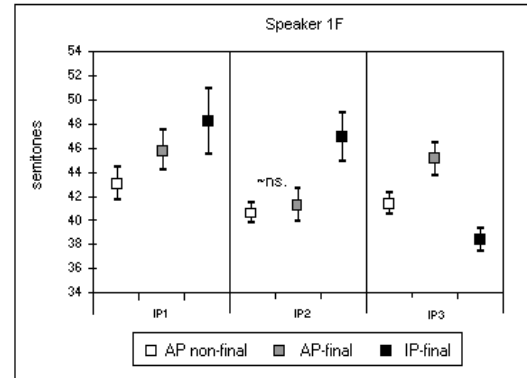


Figure 1: Mean  $f_0$  values ( $\pm 1$  s) for speaker 1F in AP non-final, AP-final, and IP-final prosodic positions.

Pitch in speaker 1F’s reading (Figure 1, and Table 1) is scaled significantly differently in most prosodic positions. There are, however, important variations within and between IPs. Between IP1s and IP2s, i.e. NPs and PAs, the pitch register is lowered significantly in all but the prosodically strongest, IP-final positions (‘ns’, Table 1). 1F’s strategy confirms the observation that final syllables of major prosodic units are not necessarily affected by register lowering [6]. The lowering of pitch is the greatest (4.4 semitones) in AP-final positions, i.e. for edge syllables of minor prosodic units, and also important in AP-non-final positions (2.4 semitones). The 0.7 semitones (about a quarter tone) difference between the two means in IP2s, however, might not be perceptible to all listeners (‘ $\sim$ ns’, Figure 1.).

Table 1: Values of Scheffe F for pitch height and IPs for speaker 1F (\*\*p<0.01; \*p<0.05)

Prosodic position	IP1-IP2	IP1-IP3	IP2-IP3
AP-non final	122.34**	21.46**	5.93*
AP-final	78.16**	ns	73.79**
IP-final	ns	245.13**	201.57**

The gradual downdrift of pitch over the course of the utterance, also known as  $f_0$  declination, cannot entirely account for the significant lowering of pitch in IP2s. In fact, in 1F's reading,  $f_0$  declination seems negligible for syllables in AP-final position. In IP1s and IP3s, the difference between the mean  $f_0$  values in this prosodic position (0.5 st) is not significant (Table 1). This means that after substantial register lowering in the parenthetical IP2s, pitch in AP-final syllables is set back to the speaker's initial register. AP-non-final syllables are also scaled up from IP2s to IP3s, although the difference between the means is very small (0.8 st). Significant final lowering seems to only affect IP-final syllables in IP3s.

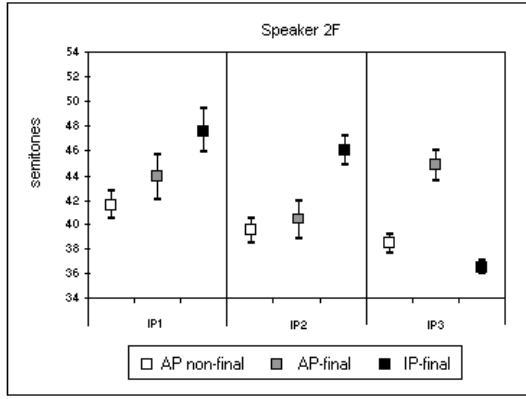


Figure 2: Mean  $f_0$  values (+/- 1 s) for speaker 2F in AP non-final, AP-final, and IP-final prosodic positions.

Speaker 2F's reading (Figure 2 and Table 2) yields similar results. Between IP1s and IP2s, register lowering does occur in all prosodic positions, including IP-final syllables, but it is the greatest in AP-final positions (3.5 st).

Table 2: Values of Scheffe F for pitch height and IPs for speaker 2F (\*\*p<0.01; \*p<0.05).

Prosodic position	IP1-IP2	IP1-IP3	IP2-IP3
AP-non final	92.12**	86.21**	11.97*
AP-final	42.68**	ns	92.88**
IP-final	14.82**	728.33**	535.37**

In 2F's reading, resetting of pitch from IP2s to IP3s only occurs in AP-final syllables. As in 1F's reading, the difference between  $f_0$  values in this prosodic position is not statistically significant (1 st), although it might be relevant perceptually. Final lowering also occurs, affecting not only IP-final, but also AP-non-final syllables in IP3s.

Speaker 1M's reading confirms the general tendency of lowering the pitch register before, and resetting it after the

parenthetical IP2s in AP-final positions (Figure 3 and Table 3). 1M, however, does not reset or lower pitch between IP2s and IP3s in AP non-final positions: the difference between the means is not significant (1 st).

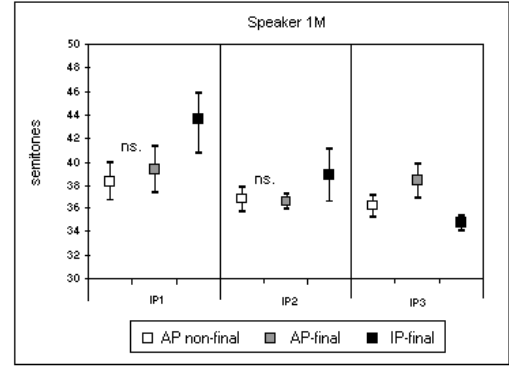


Figure 3: Mean  $f_0$  values (+/- 1 s) for speaker 1M in AP non-final, AP-final, and IP-final prosodic positions.

Similar to 1F's reading final lowering seems to only affect IP-final syllables in IP3s: the difference between mean  $f_0$  values of AP-non-final syllables is not significant (0.6 st). (For an interpretation of non-significant differences within IP1 and IP2, see [16]).

Table 3: Values of Scheffe F for pitch height and IPs for speaker 1M (\*\*p<0.01).

Prosodic position	IP1-IP2	IP1-IP3	IP2-IP3
AP-non-final	30.68**	25.05**	ns
AP-final	36.59**	ns	19.74**
IP-final	57.01**	201.46**	44.12**

### 3.2. Pitch range compression

Table 4 shows somewhat surprising results with respect to pitch range (max  $f_0$  - min  $f_0$ ) variations. Although pitch range is about 30-40% narrower in all speakers's readings when calculated without the  $f_0$  values of IP-final syllables (H% and L% boundary tones, following [12][13]), it is not systematically compressed within the parenthetical IP2s.

In 1F's reading, the maximal range of  $f_0$  variations is similar (~12 st) in all IPs when all syllables are counted, and remains similar in IP2s and IP3s (8.6 st and 8.9 st) when IP-final syllables are discounted. In 2F's reading, there is substantial narrowing of pitch range only in IP3s (6.9 st). The same is true for 1M. This speaker, however, is the only one compressing pitch range in parenthetical IP2s (5.5 st compared to 7.9 st in IP1s, and 7.3 st in IP3).

Figure 4 illustrates the pitch track of one of speaker 2M's utterances, a clear example of phrase-medial pitch range compression in this speaker's speech (excerpt from [17]). Although utterances as in Figure 4, in which both pitch register lowering and pitch range compression co-occur, are not typical in this corpus, it is likely that they represent a 'prototype' of utterance-medial parentheticals. Studies on pragmatics of French intonation might be referring to this prototypical tonal template when reporting on both pitch

register lowering and pitch range compression. It must be emphasized, however, that such template might not occur

systematically in all in utterance-medial parentheticals or *incises*.

Table 4: Pitch range (max  $f_0$ -min  $f_0$ ) in semitones for all speakers and IPs.

	Speaker 1F			Speaker 2F			Speaker 1M		
	IP1	IP2	IP3	IP1	IP2	IP3	IP1	IP2	IP3
max	51.4	50.3	47.8	52.2	49.9	42.6	48.9	43.5	42
min	39.5	38.3	36	39.1	35.5	35.7	35.6	34.2	32.8
range <b>with</b> IP % tones	<b>11.9</b>	<b>12</b>	<b>11.8</b>	<b>13.1</b>	<b>14.4</b>	<b>6.9</b>	<b>13.3</b>	<b>9.3</b>	<b>9.2</b>
range <b>without</b> IP % tones	6.9	<u>8.6</u>	8.9	7.4	<u>8.5</u>	9.8	7.9	<u>5.5</u>	7.3

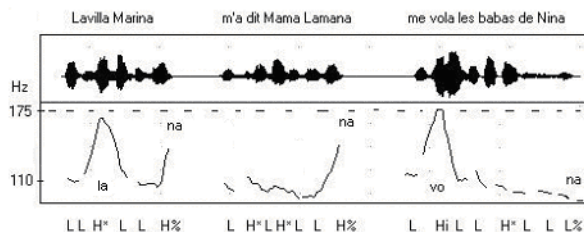


Figure 4: Speaker 2F's reading of the utterance "La diva Lavilla Marina (IP1), m'a dit Mama Lamana (IP2), me vola les babas de Nina (IP3)" containing a seven-syllable parenthetical in IP2.

#### 4. Conclusion

This study investigated the scaling of pitch in utterance-medial parentheticals in French. It concluded that parentheticals have a typical tonal template, in which pitch is scaled down to some extent in all but non-intonational phrase-final syllables. Neither phrase final lowering, nor  $f_0$  declination provided satisfactory explanations. Downscaling of pitch, interpreted as pitch register lowering, is the greatest in edge syllables of minor prosodic units, and shows remarkable inter-speaker consistency. Pitch range compression, on the other hand, seems speaker-dependent. Only one speaker showed a systematic tendency for compressing pitch range in all tonal targets in utterance-medial parentheticals. This suggests that narrowing of pitch range might be a less stable phonetic cue for background information in parentheticals than previously thought.

The uniform lowering of pitch register in parentheticals raises the much-debated issue of modeling pitch range as a gradient or a categorical phonetic/phonological feature. Although this study cannot deal with this vast question, it suggests that there might be an alternative approach to modeling pitch range on the pragmatics/phonetics interface, rather than in phonology (see [15] for discussion). Free from any interference with tonal inventories or other phonological constraints, backgrounding and foregrounding of information could be treated as more or less categorical cases of gradient pitch range variations at the pragmatics/phonetics interface.

#### 5. References

[1] Swerts, M., 1993. *Prosodic Features of Discourse Units*. Doctoral Dissertation, IPO, Eindhoven.

[2] Denzel, van, M., 1999. *Prosodic Aspects of Information Structure in Discourse*. Doctoral Dissertation, University of Amsterdam.

[3] Mozziconacci, S., 1998. *Speech Variability and Emotion: Production and Perception*. Doctoral Dissertation, University of Eindhoven.

[4] Ladd, D.R.; Terken, J., 1995. Modeling intra- and inter-speaker pitch range variation. In *Proceedings of the 13<sup>th</sup> ICPhS, Stockholm*, vol. 2, 386-389.

[5] Espinal, M.T., 1991. The representation of disjunct constituents. *Language*, 67(4), 726-62.

[6] Morel, M-A.; Danon-Boileau, L., 1997. *La grammaire de l'intonation*. Paris: Ophrys.

[7] Bolinger, D., 1986. *Intonation and its parts: melody in spoken English*. Stanford: Stanford University Press.

[8] Ashby, J.W., 1997. An acoustic profile of right dislocations in French. *Journal of French Language Studies*, 4, 127-145.

[9] Ladd, D.R., 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.

[10] Mertens, P., 1984. *Intonation du français: De la description à la reconnaissance automatique*. Ph.D. dissertation. Katholieke Universiteit Leuven.

[11] Kutik, E.; Cooper, W. E.; Boyce, S., 1983. Declination of fundamental frequency in speakers' production of parenthetical and main clauses. *Journal of Acoustical Society of America*, 73, 1731-1738

[12] Jun, S-A.; Fougeron, J., 1995. The Accentual Phrase and the prosodic structure of French. In *Proceedings of the ICPhS of Stockholm*, vol. 2, 722-725.

[13] Jun, S-A.; Fougeron, J., 2000. A Phonological Model of French Intonation. In *Intonation: Models and Theories*, Botinis A., ed. Amsterdam: Kluwer, 169-182.

[14] Beckman, M.; Pierrehumbert, J., 2000. Tonal elements and their alignment. In *Prosody: Theory and Experiment. Studies presented to Gösta Bruce*, Horne, M., ed. Amsterdam: Kluwer, 11-36.

[15] Ladd, R., D., 1994. Constraints on the gradient variability of pitch range, or, Pitch level 4 lives! In *Papers in Laboratory Phonology III*, Keating P., ed. Cambridge: Cambridge University Press, 43-63.

[16] Fagyal, Zs., (in press). "Prosodic boundaries in the vicinity of utterance-medial parentheticals in French." In *Probus. Special issue on the intonation of Romance languages*. Hualde, J. I. ed. Spring 2002.

[17] Fagyal, Zs., (in press). "Intonation in utterance-medial parentheticals and the syntax-phonology interface in French." In *Chicago Linguistic Society 37: The Main Session*. Andronis, M. et al., eds. Spring 2002.