Durational Cues and Prosodic Phrasing in French: Evidence for the Intermediate Phrase

Amandine Michelas[*], *Mariapaola D'Imperio*[*]

[*] Université Aix-Marseille I et Laboratoire Parole et langage, CNRS, Aix-en-Provence, France
michelas@lpl-aix.fr, mariapaola.dimperio@lpl-aix.fr

Abstract

Studies addressing prosodic constituency in French generally agree on two levels of phrasing (accentual phrase, AP, and intonation phrase, IP), while the existence of an intermediate level of phrasing (intermediate phrase, ip) is still controversial. In this study we examine durational cues in a read speech corpus at normal and fast rates in which the target syllable was either adjacent to a prosodic boundary or word-internal. Additional evidence for the existence of an intermediate level of phrasing between the AP and the IP was found: the vicinity to an ip-boundary is signaled by durational cues that are stronger than the ones associated to an AP-boundary, yet this lengthening is weaker than the one found in the vicinity of an IP boundary.

Index Terms: intermediate phrase, prosodic phrasing, prosodic boundary, degree of lengthening, speech rate, French.

1. Introduction

Authors generally agree on the existence of two prosodic units in French (though they are labeled differently according to the approach and author). A smaller unit called rhythmic unit [1], accentual phrase (AP, [2]), phonological phrase [3], or prosodic word [4, 5], which is characterized by an obligatory F0 rise on the last syllable of the phrase (LH* in the the autosegmental-metric model of Jun & Fougeron [2]) plus a small degree of preboundary lengthening [6, 7]. This constituent is also characterized by the presence of an optional initial rise (Hi) on its first syllable. Authors also agree on the existence of a larger unit, i.e. the Intonation Phrase (IP, [2, 3, 8], see also intonation unit [9, 10, 11]). The intonation phrase (IP) is demarcated by a major intonation contour (marked by a phrase final boundary tone realized on the last syllable of the phrase) and a marked preboundary lengthening [7, 12]. Thus IP-final syllables show a greater degree of lengthening than AP-final syllables [2, 12, 13]. Additionally, IP can be optionally followed by a pause. From a syntactic point of view, the IP is the domain of the root sentence corresponding to the grouping of a noun phrase (NP) and a verb phrase (VP) [14]. According to this definition specific syntactic constructions such as dislocations or parentheticals, as in (2) may form an autonomous IP [15].

(1) *Le mari d'Amanda réclamait sa bicyclette.*
[ ]IP
"Amanda's husband asked for his bicycle."

(2) *Le mari d'Amanda, d'après ce qu'on m'a dit, était marin.*
[ ]IP [ ]IP[ ]IP
"As far as I was told, your husband was a sailor"

Speech rate has also been shown to affect phrasing in a significant way [16], [17], [18]. Specifically, Fougeron and Jun [18] show that speakers reduce the number of phrases within an utterance at fast rate, by deleting or reducing the strength of prosodic boundaries. The study uncovered three specific examples of reduction at fast rate: the IP boundary can be reduced to an AP boundary, (2) two APs can be grouped at fast rate and (3) the AP initial Hi tone can be suppressed. .

An intermediate level of phrasing, the intermediate phrase (ip) or major phonological phrase (MAP, [19]) has been shown for several languages, including English [20], Italian [21, 22], Catalan [23], and Cairene Arabic [24]. Arguments for the existence of an intermediate level smaller than the IP but bigger than the AP in French have been previously proposed even if its status remains controversial. Di Cristo and Hirst [9] remarked the existence of an intermediate constituent (called an intonation phrase segment, or S.UI) which would be stronger than the rhythmic unit boundary yet smaller than an intonation unit boundary in specific prosodic structures (such as tag-questions, dislocations or postpositions). A similar domain (which can be found in tag or wh-questions, dislocations as well as vocatives, lists, or implicative contours) has been uncovered by Jun and Fougeron [8].

In this study we focused on the durational cues of prosodic constituents of different sizes at both normal and fast speech rates (for tonal phonetic and phonological cues, see [25]). Specifically, we predicted that the prosodic cues associated with an ip right boundary would be stronger than the ones which are associated with an AP-right boundary and yet weaker than those associated with an IP-right boundary.

2. Method

2.1. Corpus

In this study we investigated the properties of SVO utterances in which the target syllable /na/ occurred in four different contexts: 1) within an AP, 2-at an AP boundary (ip-internal), 3) at a potential ip boundary (at the boundary between the subject NP and VP) and 4) at an IP boundary (i.e., before a parenthetical). The target vowel was always the low front unrounded vowel /a/, in the context of a CV syllable, in which the onset consonant was always voiced and could be either a nasal or a liquid, while the following onset consonant was always /d/. The 4 contexts are shown in table 1.

	Target vowel /a/
1- within a Prosodic Word	Les grenadiers]AP de marrakesh]ip ne poussent pas bien vers chez nous.]IP
2- at an AP boundary	Le sauna]AP de Paolo]ip deviendra incontournable.]IP
3- at an ip boundary	Le sauna]ip deviendra incontournable.]IP
4- at an IP boundary	Le sauna]IP d'après ce qu'on m'a dit]IP n'est pas très loin.]IP

Table 1: *Example of corpus items.*

20 fillers items were added to the 20 experimental sentences. The order of presentation of the sentences was randomized separately for each subject.

2.2. Procedure

Two native speakers of French (both male: ages 28 and 27 years old) read the sentences aloud four times at both normal and fast speech rate for a total of 1600 experimental sentences (20 sentences x 4 repetitions x 2 speech rates x 2 speakers = 320). Speakers were instructed to read each sentence as naturally as possible, first at a normal rate and then at a fast speech rate. The sentences were visually presented on a computer screen. Participants were instructed to press the space bar on the keyboard in order to advance between items. Speakers were recorded onto digital audio tape (DAT), using a Shure SM10A head-mounted microphone in a sound-attentuated roomat the Laboratoire Parole et Langage. Before the experiment, participants read six sentences at both normal and fast speech rate in order to test the material and the procedure. The resulting-sound files were segmented and each utterance was saved as a separate file. Spectrograms were created using Praat [27]. Target vowels, target syllables and target AP boundaries were labeled by inspecting both waveforms and spectrograms.

2.3. Measures

Both syllable and target vowel durations were measured. The total duration of the utterance was calculated in order to verify that the rate manipulation was significant [F(3.252)=345.2, p<0.001]. Rate was similarly calculated for each target AP in order to test whether the adjustments to speech rate affected the target AP and the overall utterance in the same way. Results of two factor ANOVA (rate and speaker) show that there is a strong effect of speech rate (see Figure 1 and 2 [by utterances: F(3,252)=345.2, p<0.001; by target APs: F(3,252)=85.8, p<0.001)]. The effect of speaker was not significant [for sentences F(3,252)=3.8, p>0.001; for target APs F(3,252)=3.6528, p > 0.001], nor was the rate by speaker interaction [for sentences: F(3,252)=7,9, p>0,001; for target APs: F(3,252)=0,1 p>0.001]. Average speaking rate in syllables per second was calculated for each speaker according to the procedure used by Welby & Loevenbruck [28]. Each utterance was inspected auditorily, and the base count was adjusted on the actual pronunciation (the syllable base count of each sentence depended on the context condition: 15 syllables for the AP-internal vowel condition, 10 syllables for the AP-final vowel condition, 13 syllables for the ip and IP-final vowel conditions. This syllable count was divided by the utterance duration (including pauses) to obtain a rate measurement for each utterance. These results confirmed that the two speakers successfully augmented their rate of speech in going from a normal to a fast speaking rate condition, for both the target AP and the overall sentence. The rate increase was approximately the same for the two speakers.

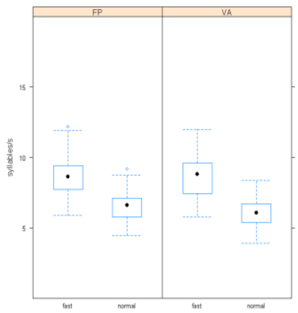

Figure 1: *Speaking rates in syllable/s for entire utterances for both speakers..*

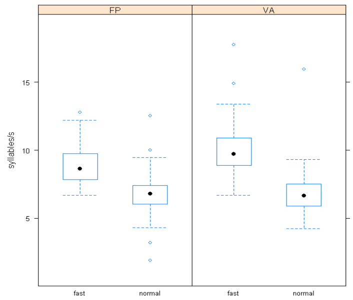

Figure 2: *Speaking rates in syllable/s for target APs for both speakers.*

Out of 320 utterances, 256 (16 sentences x 4 repetitions x 2 speech rates x 2 speakers = 256) were used for the analyses. The items containing the syllable /la/ were excluded because these items presented some segmentation difficulties (despite the liquid nature of the consonant there were not an abrupt change in intensity).

2.4. Hypothesis

We examined the effect of the four contexts on both vowel and syllable duration for the two speech rates. At normal speech rate, we predict that units higher in the prosodic hierarchy will undergo a greater degree of preboundary lengthening. If our hypothesis is correct, we will obtain a four-step vowel lengthening as shown in Figure 3.

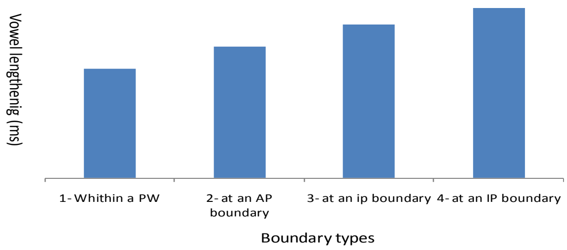

Figure 3: *Expected target vowel duration (ms) as a function of boundary type at normal speech rate.*

We also examined the effect of speech rate on vowel and syllable duration. Our hypothesis was that a fast speech rate would result in a modification of the prosodic organization of an utterance. Specifically we predicted that IP boundaries would be demoted to ip or IP boundaries.

3. Results

Two mixed models were separately performed for syllable duration and vowel duration. Both models included rate (normal/fast), prosodic boundary type (Unaccented, AP-final, ip-final, IP-final) and speaker (VA and FP) as fixed effects and the type of preceding onset consonant as random effect. These statistical models were intended to normalize the speech rate variability found both within and across speakers and allowed us to evaluate the effect of speech rate on prosodic phrasing (see [29] for more information). Duration measurements were log-transformed in order to obtain a normal distribution. Vowel and syllable duration (in ms) by boundary type at both the normal and fast speech rate are shown for both speakers in Figures 4 and 5.

Consistent with our predictions, the statistical analysis showed that at the normal speech rate, both vowel and syllable length increases with prosodic boundary strength. At the normal speech rate, AP-final vowels and syllables were significantly longer than unaccented vowels (for vowel lengthening: t=4.40, p<0.05, effect size: 19 ms; for syllable lengthening: t=2.15, p=0.05, effect size: 10ms) AP-final vowels and syllables were significantly longer than ip-final syllables (for vowel lengthening: t=2.64, p<0.05, effect size: 14ms; for syllable lengthening: t=2.4, p<0.05, effect size: 15ms) and ip-finals vowels and syllables were significantly longer than IP-final vowels (for vowel lengthening, t=3.83, p<0.05, effect size:24ms; for syllable lengthening, t=4.81, p<0.05, effect size:30ms). The large distribution of the duration values observed for the IP-final syllables boundary may be explained by the optional presence a silent pause after the IP boundary

At the fast speech rate, vowel durations were much more similar, so that no significant effect of boundary type is found between the levels of IP, ip and AP. The mixed model analyses showed that only the IP-final vowels (t=2.47, p<0.05, effect size: 9ms) and ip-final vowels (t=2.87, p<0.05, effect size: 10.5ms) were significantly longer than unaccented vowels. For syllable lengthening we similarly observed a significant difference between ip-final syllables and unaccented vowels (t=4.75, p<0.05, effect size: 18ms) and between IP-final syllables and unaccented vowels (t=2.60, p<0.05, effect size: 10ms). Durational differences were not significant for the comparison between unaccented vowels and AP-final vowels (t=1.06, p=0.2895) nor for the comparison between AP-final syllables and unaccented syllables (t=1.57, p=0.118, size effect: 10ms).

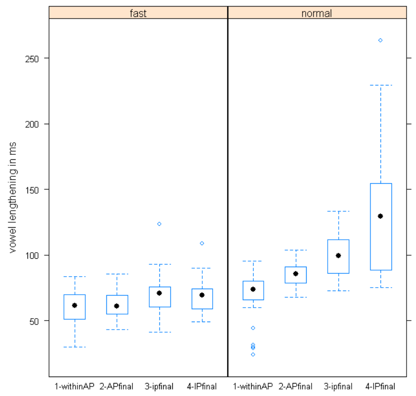

Figure 4: *Vowel duration (in ms) by boundary type for both speakers at fast and normal speech rates.*

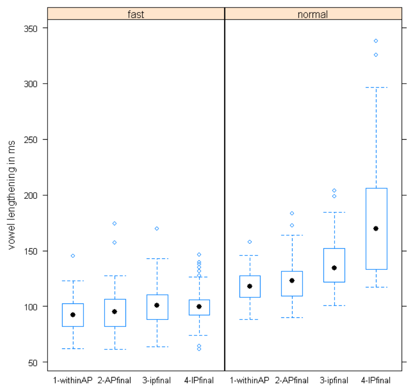

Figure 5: *Syllable duration (in ms) by boundary type for both speakers at fast and normal speech rates.*

4. Discussion

Consistent with the predictions of Jun and Fougeron [2] and Pasdeloup [7], our results showed that, at the normal speech rate, AP-final vowels were significantly longer than AP-internal syllables, while they were significantly shorter than IP-final syllables. As Pasdeloup showed, the lengthening of AP-final vowels does not exceed 50% of mean AP-internal vowel duration (the mean for the two speakers was 71 ms for the normal speech rate and 84 ms for the fast rate). Vowel lengthening associated with an IP-boundary exceeded 50% of mean AP-internal vowel duration (the mean for IP-final vowels for the two speakers at the fast speech rate was 128 ms). Nevertheless, the results also seem to support the existence of an intermediate level of phrasing (ip) which would be higher than AP but lower than IP. Indeed at the normal speech rate, ip-final vowels were significantly longer than AP-final vowels while also significantly shorter than IP-final syllables. This ip would not be restricted to marked constructions as it was originally proposed by Jun & Fougeron [8], since the syntactic construction employed here is unmarked. We propose that an alignment constraint between syntactic and prosodic structure conspires to place an ip boundary to the right edge of a major syntactic break (see [25]). Specifically, we propose that an ip boundary can appear also in all-focus utterances and that its right boundary will be signaled through significant preboundary lengthening relative to the AP-final domain. Moreover, our results show that the degree of vowel lengthening associated with an ip-boundary is about 50% of the mean for AP-internal vowels.

Although we showed that preboundary lengthening associated with an ip-right boundary is significantly different from that associated with an AP-right or IP-right boundary, the question of the nature (discret/gradual) of the phonetic cues associated with the ip-boundary arises. Most studies addressing prosodic constituency have been conducted in the Prosodic Phonology framework, in which prosodic constituency is viewed as a hierarchy of domains. We know that length affects prosodic phrasing. If the preboundary lengthening associated with the ip-right boundary reflects phonological structure, it should not vary relative to the length of the ip. In a study to appear [30], we manipulated the length of the ip in order to test the hypothesis that this manipulation would affect the “relative strength” [31] of the duration cues associated with the boundaries. The results showed that the phonetic and phonological cues associated with an ip break were independent of the length manipulation, and they generally support the existence of an intermediate level of phrasing in French.

At the fast speech rate, preboundary lengthening seems to differentiate only AP-internal vs. ip and the IP-final vowels. Despite the absence of significant lengthening of AP-final vowels at fast rate, this level nevertheless appears to be tonally marked by the presence of a LH* accent. Indeed in 75 out of 80 utterances in our AP-final condition, we observed a final LH* rise associated with the primary stressed syllable (/na/ in Figure 6).

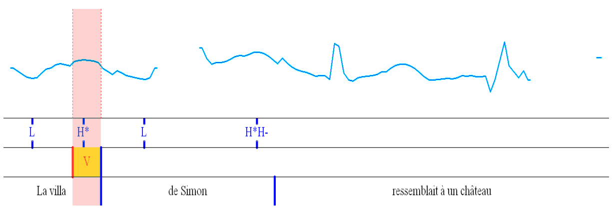

Figure 6 : *F0 curve for the utterance La villa de Simon ressemblait à un château. ‘Simon’s house looked like a castle’ produced by one of our speakers at fast speech rate.*

Since no significant vowel lengthening is observed here, it could be objected that the rise observed on /na/ is an instance of an initial LHi rise (and not a final LH* rise), which is an optional left edge marker of the AP [2] subject to conditions of rhythm, style or speaker (see also [32]). However two facts allow us to claim that the H tone we observed in our AP-final condition is a LH* pitch accent and not an instance of Hi. First, the H tone is often realized on a primary stressed syllable, which is a phrase-final full vowel (*accent primaire*) of a phrase and not on the initial stressed syllable of the first lexical word. Secondly, the L tone in the APs tends to be realized on the syllable preceding the H* marked syllable (as shown in Figure 6), and not at the onset of the same syllable, as it is usually the case for LHi.

Thus, in contrast with previous results [18], we conclude that in our corpus, fast speech rate did not induce complete AP-boundary. Moreover, we did not observe ip-boundary erasure, since a slight lengthening was always found for this level as well as a tonal marker of this level (i.e., a significant return to the phrase register line, cf. [25]). Finally, the duration cues are less marked at the fast speech rate than at normal speech rate. These results seem to support the idea of mixed marking for prosodic boundaries, in that there appears to be a trade-off between lengthening cues and tonal (and possibly even spectral) cues in order to induce a perceptible phrase break. We plan to test our proposal through future perception experiments testing the reality of the phrasing levels addressed in the present analysis.

5. Conclusion

In this paper, we have shown that, in French, an ip-boundary may occur in all-focus utterances showing a non-marked syntactic structure. The right edge of an ip boundary appears to be marked by significant vowel and syllable lengthening which are both stronger than those observed for AP-final syllables and yet weaker than those observed for IP-final syllables. Our results suggest that prosodic cues are reinforced when there is an alignment between prosodic and syntactic boundaries, and they support the existence of an intermediate prosodic level in French. Finally, our results show a tendency for prosodic organization to be modified at a fast speech rate, since duration cues can be weaker than at a normal speech rate, while being compensated for other cues, such as phrase accents and/or edge tones.

6. References

[1] Di Cristo, A., "Intonation in French". In D. Hirst, & A. Di Cristo (Ed), Intonations systems: A survey of twenty languages, Cambridge: Cambridge University Press, 195-218, 1998.

[2] Jun, S.A. and C. Fougeron.,"A Phonological Model of French Intonation", Probus, 14, 147-172, 2000.

[3] Post, B. "Tonal and Phrasal Structures in French Intonation". The Hague: Holland Academic Graphics, 2000.

[4] Martin, P., "Une théorie syntaxique de l'accentuation en français", Studia Phonetica, 15, 1-12, 1979.

[5] Vaissière, J., "Rhythm, Accentuation and Final Lengthening in French". In J. Sundberg, L. Nord and R. Carlson [Ed], Music, language, speech and brain, 59,108-120, Stockholm : Wenner-Gren, International Symposium Series, 108-120, 1992.

[6] Hirst, D. and Di Cristo, A. "French Intonation: a Parametric Approach". Die Neueren Sprachen, 83(5), 554-569, 223-226, 1984

[7] Pasdeloup, V., "Modèle de règles rythmiques du français appliqué à la synthèse de parole" . Doctoral thesis, Université de Provence, 1990.

[8] Jun, S.A. and Fougeron, C., "The Realizations of the Accentual Phrase in French Intonation". Probus, 14, 147-172, 2002.

[9] Di Cristo, A. and Hirst, D. "Vers une typologie des unités intonatives du français ". XXIème JEP, Avignon,219-22, 1996.

[10] Di Cristo, A., "Vers une modélisation de l'accentuation du français : première partie". French Language Studies, 9, 143-179, 1999.

[11] Di Cristo, A., "Vers une modélisation de l'accentuation du français seconde partie". French Language Studies, 10, 27-44, 2000.

[12] Di Cristo, A., "Des indices prosodiques aux traits perceptuels: Application d'un modèle d'analyse prosodiques à l'étude du vocative en français". Travaux de l'institut de Phonétique d'Aix, 3, 213-258, 1976.

[13] Di Cristo, A., "De la microprosodie à l'intonosyntaxe". Aix-Marseille : Université de Provence, 1985.

[14] Selkirk, E., "On Prosodic Structure and its Relation to Syntactic Structure". In T. fretheim [Ed], Nordic Prosody II, Trondheim: TAPIR, 1978.

[15] Delais-Roussarie, E. and Post, B., "Unités prosodiques et grammaires de l'intonation : vers une nouvelle approche », In Actes des XXVIIème JEP, Avignon, 2008.

[16] Bartkova, K., "Speaking Rate Modelisation in French : Application to Speech Synthesis". In Proceedings of the XIIth International Congress of Phonetic Sciences, 3, 482- 485, 1991.

[17] Keller, E. and Zellner, B., "A Statistical Timing Model for French". In Proceedings of the XIIIth International Congress of Phonetic Sciences, 3, 302-305, 1995.

[18] Fougeron, C. and Jun, S.A., "Rate Effects on French Intonation: Phonetic Realization and Prosodic Organisation". Journal of Phonetics, 26, 45-70, 1998.

[19] Selkirk, E., "The Interaction of Constraints on Prosodic Phrasing. In Prosody: Theory and Experiment (Ed), Merle Horne, 231-262. Dordrecht: Kluwer Academic Publishing, 2000.

[20] Beckman, M. and Pierrehumbert, J., "Intonational Structure in Japanese and English", Phonology Yearbook III, 15-70, 1986.

[21] D'Imperio, M., "Italian Intonation: an Overview and some Questions". Probus, 14(1): 37-69, 2002.

[22] D'Imperio, M., Elordieta, G., Frota, S., Prieto, P., Vigario, M. "Intonational Phrasing and Constituent Length in Romance". (Selected papers from the Phonetics and Phonology in Iberia Conference, 2003). Phonetics and Phonology Series. The Hague: Mouton de Gruyter, 59-98, 2005.

[23] Prieto, P., "The Intonational Phonology of Catalan". In S. A. Jun (Ed) Prosodic Typology 2, Oxford University Press, Oxford, to appear.

[24] Hellmuth, S., "The Relationship between Prosodic Structure and Pitch Accent Distribution: evidence from Egyptian Arabic". The Linguistic Review. 24: 291-316, 2007.

[25] D'Imperio M. and Michelas, A., "Embedded Register Levels and Prosodic Phrasing in French". Proceedings of Speech Prosody, 5, Chicago, 2010.

[26] D'Imperio M. and Michelas, A., "Interface entre structure syntaxique et structure prosodique: le syntagme intermédiaire en français". Actes, Interfaces Discours & Prosodie, 3, Paris, 2009.

[27] Boersma, P. & Weenick, D., "Praat: Doing Phonetics by Computer" (Version 4.5.18), 2007, http://praat.org/.

[28] Welby, P. and Lœvenbruck, L., "Anchored down in Anchorage: Syllable Structure and Segmental Anchoring in French". Rivista di Linguistica, 18, 74–124, 2006.

[29] Baayen, R.H., "Analyzing Linguistic Data: a Pratical Introduction to Statictics. Cambridge, Cambridge University Press, to appear.

[30] Michelas, A. and D'Imperio, M., "Syntactic and Length Effects on French Prosodic Phrasing", to appear.

[31] Gee, J. and Grosjean, F., "Performance Structures: a Psycholinguistic and Linguistic Appraisal", Cognitive Psychology, 15, 411-458, 1983.

[32] Welby, P., "The Slaying of Lady Mondegreen, Being a Study of French Tonal Association and Alignment and their Role in Speech Segmentation. Doctoral thesis, The Ohio State University, 2003.