The Realization of Early and Late Rises in French Intonation: A Production Study

Pauline Welby

Department of Linguistics
The Ohio State University
`welby@ling.ohio-state.edu`

Abstract

The results of a production study of two speakers provide evidence against models of French intonation in which there is a one-to-one correspondence between a syllable and each of the L (low) and H (high) tones of a rise. The results also fail to confirm claims about the influence of the consonant or vowel status of an accentable syllable on the placement of the H of the early rise. The L tone of the early rise consistently straddled the boundary between the last function word and first content word of the phrase (e.g., for *un ENfant SAGE*, the L tone would be realized very late in *un* or very early in *EN*), with a low leading plateau extending leftward to the beginning of the phrase. This timing is accounted for in the current proposal by a double association of the L tone of a LH- edge tone to the left edge of the phrase and to the left edge of the first content word of a phrase, similar to the double association of phrase accents discussed in Grice et al. (2000) for a number of languages.

1. Introduction

In French, accent is a property of the phrase, rather than the word. Discussions of French intonation commonly assume that there is an obligatory rising accent assigned to the final full syllable of a phrase (e.g., *un enfant SAGE* 'a good child') and an optional early rise that occurs somewhere before the late rise (e.g., *un ENfant SAGE*). A number of factors have been observed to influence the appearance of the early rise, including number of syllables in the phrase, rhythmic context, morphological status, and segmental characteristics of a word's first syllable ([6] *inter alia*).

All current accounts of French intonational structure agree that the late rise demarcates a level of prosodic constituency smaller than the intonational phrase. Following Jun & Fougeron, I will call this unit the "accentual phrase" (AP) [4], [5].[1] Accounts differ, however, in several im-

[1]Although I use the Jun & Fougeron term, I am not committed to all aspects of the Jun & Fougeron model. Jun & Fougeron's AP coincides in many cases with other proposed units, including
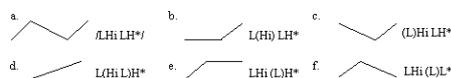

Figure 1: *Six surface realizations of Accentual Phrase (AP) (Jun & Fougeron (to appear, 2002)). In Post's (2000) model, these would be transcribed: a. %L H* L H*, b. %L H* (with spreading of initial boundary tone), c. %H L H*, d. %L H*, e. %L H* H*, f. %L H* H* L%.*

portant respects. For example, Di Cristo & Hirst (1993) treat the early and late rises as fundamentally the same phenomenon, a demarcative accent that defines a tonal unit below the rhythmic unit [1]. By contrast, Post (2000) [8] and Jun & Fougeron (to appear, 2002) argue that the two rises are different in nature. Post treats both rises as sequences of L (low) tone targets followed by H (high) pitch accents, but argues that the L of the early rise is a boundary tone. Jun & Fougeron (to appear, 2002) argue that the early rise is a bitonal LH phrase accent inserted at the left edge of the AP. The late rise is a bitonal LH* pitch accent with the H tone associated to the last full syllable of the phrase. The LH phrase accent and the LH* pitch accent together form the basic unit of French intonation, the accentual phrase (AP), which has the structure /LHiLH*/. The model accounts for five surface shapes in addition to /LHiLH*/ by undershoot and tone deletion when there are not enough syllables to realize all four tones of /LHiLH*/. A chart of the six total patterns is given in figure 1. An earlier Jun & Fougeron model (2000) posits the same underlying and surface forms of the AP, but claims a one-to-one correspondence between tones and syllables.

In Post's (2000) model, an Intonation Phrase can begin with either a %L or %H boundary tone and end with a L%, H% or 0%. The initial boundary tone is followed by an optional early H* pitch accent and by an obligatory late H* or H+H* pitch accent. If there is a sequence of two H tones (if there is both a late rise and an early rise, for example) a L tone may be inserted between the two. Post claims that the "low tone is usually aligned with the pre-accentual syllable . . . " (p. 153), although she later states that "[a]lthough the low target is often realised in the accented syllable in IP-final rising and IP-internal rising-falling movements, . . . it can also occur in the penultimate syllable . . . " (p. 168).

In the study, I address aspects of the following interrelated questions: Is there a one-to-one correspondence be-

Di Cristo & Hirst's Rhythmic Unit, Post's Phonological Phrase, and Pasdeloup's (1990) mot rythmique.

tween tones and syllables? The Jun & Fougeron (2000) model explicitly claims a one-to-one correspondence, and the Post (2000) model makes more limited claims of such a correspondence. What are the details of the text-to-tune alignment? A number of detailed phonetic studies have examined the alignment facts of languages such as Greek, Spanish, and English, but there is a lack of such studies for French. The French alignment facts also bear upon the issue of whether the early and late rises are instances of the same phenomenon (as claimed by Di Cristo and Hirst) and whether they are pitch accents associated to syllables or edge tones associated to the edges of prosodic units.

I also test conflicting claims about the influence of segmental structure on the placement of the H tone of the early rise. Pasdeloup (1990) and Post (2000) both note a tendency for the early H to be realized on the first syllable of the first content word in cases where the content word is consonant-initial and on the second syllable in cases where the content word is vowel-intial. Post maintains that the status of a word as consonant- or vowel-initial is determined by its underlying segmental composition. Pasdeloup, however, claims that even a word which only becomes consonant-initial in its surface form through the process of enchaînement (e.g., cet<u>t'im</u>prudence)), will pattern with consonant-initial words with respect to the placement of the early H.

2. Methods

2.1. Subjects

Two native speakers of French participated, a 27 year old female speaker from Marseille (speaker 1) and a 33 year old female speaker from Tours (speaker 2).

2.2. Materials

Sentences were constructed with 15 proper name subjects. The names were 2, 3 or 4 syllables in length, and were preceded by 0, 1 or 2 monosyllabic function words (*et, ma, mon*). Half of the names were vowel-initial (always /a/) and half began with a consonant (always /m/). Combining the vowel-initial names with the possessive adjective *mon* created a liaison environment. Fifteen verb phrases (VPs) were constructed, each of which began with the auxiliary verb *allait* 'was going to'. This decision was made to minimize the effect of tone clash on the realization of tones early in the sentence, since the function word *allait* is unlikely to be assigned a tone. Different VPs rather than a single carrier phrase were used to minimize the likelihood that the sentence would be read with constrastive focus on the subject, since focus is known to affect the realization and timing of rises. Each of the 15 proper name/function word sets was paired with each of the 15 VPs to produce 15 lists of 15 sentences for a total of 225 sentences. An example sentence is "Ma Marianna allait rembobiner la laine irlandaise." 'My Marianna was going to wind the Irish wool'. Example materials and a longer version of the paper are available at http://www.ling.ohio-state.edu/~welby/sp2002.html.

2.3. Procedures

Speakers were instructed to read the 15 lists aloud at a normal speaking rate. Speakers were prompted to repeat disfluent or misread productions. The speech was recorded using a head-mounted microphone and a DAT recorder at 44.1 kHz in a quiet room or sound-attenuated booth.

2.4. Data preparation

Recordings were transferred to computer using a digital-to-digital cable and downsampled to 22.05 kHz. The sound-files were segmented and each utterance was saved as a separate file. Fundamental frequency (F0) was extracted every 5ms and spectrograms were created using Entropic xwaves speech analysis software. The analysis of the data thus far has been limited to utterance-initial APs (which almost always corresponded to the proper name/function word set), since the early rise is claimed to be very common in utterance-initial position (e.g., Pasdeloup (1990)). Word and syllable boundaries were tagged for each proper name/function word set. A number of prosodic features were also hand-tagged: initial low tone (L1), early high tone (H1), and late high tone (H2). Adapting a procedure described in D'Imperio (2000) [2], the position of inflection points or elbows between L1 and H1 (Elbow 1) and H1 and H2 (Elbow 2) was automatically calculated and inserted into a label file. Only APs with preceding function words typically have Elbow 1.

3. Results

3.1. Accent patterns

There was an inter-speaker difference in the AP patterns that were produced. Speaker 2 produced about 2/3 of her utterance-initial APs with the LLH pattern and 1/3 with the LHLH pattern. Speaker 1 produced many tokens with the LLH pattern and LHLH patterns, but she also produced about 1/4 of the utterance-initial APs with either a LHL or LHH pattern and a small number with the LH pattern.

Despite these differences, both speakers frequently produced the LHLH pattern. The LHLH was realized on APs as short as 2 syllables for Speaker 1, who produced about 30% of her utterance-initial 2 syllable APs with the LHLH pattern. Speaker 2 produced about 20% of her utterance-initial 3 syllable APs, although none of her 2 syllable APs, with the LHLH pattern. The percentage of APs realized with the LHLH pattern increased steadily for both speakers as the number of syllables increased. The observed occurrence of the LHLH pattern versus the occurence of the pattern expected in a model where there is a one-to-one tone-to-syllable correspondence, such as Jun & Fougeron (2000), is shown in figure 2. The precise numerical value of the bars representing the expected occurence of the pattern is not crucial. The important point is that we see a steady increase in the number of LHLH patterns as the number of syllables increases rather than a sharp jump from no or very few LHLH patterns in 3 syllable APs (which do not have enough syllables for a 4 tone pattern) to an equally high number of LHLH patterns in 4, 5, and 6 syllable APs. Note that an account that treats the early H as rhythmic in nature (e.g., Pasdeloup (1990)) would not predict the appearance of the early H in 2 syllable APs.

3.2. Text-to-tune alignment

The H tone of the late rise was almost always realized on last syllable of the AP. In a small number of cases, it was realized just after the end of this syllable. As expected, the realization of the H tone of the early rise was more variable, but
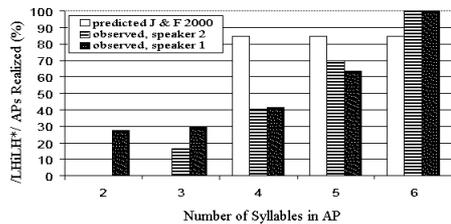
Figure 2: Realization of /LHiLH/, observed vs. expected (Jun & Fougeron (2000))*

Figure 3: Placement of Elbow 2 (Inflection Point before Late H

always occurred on one of the first two content word syllables. The two speakers had a general tendency to realize the early H on the second syllable of the content word (proper name). A comparison of APs containing only a proper name (e.g., *Marianella*, *Arianella*), showed that consonant-initial names were more often realized with an early rise than were vowel-initial names (54/89 cases *vs.* 36/89 cases). This difference points to the influence of duration rather than (or perhaps in addition to) number of syllables in determining the presence of an early rise. The consonant-initial names and the vowel-initial names were segmentally identical except for the onset /m/ of the consonant-initial words. Accordingly, the consonant-initial names had a slightly longer average duration (538ms *vs.* 518ms).

Both the L and the H tones of the late rise were often (in approximately 80% of cases) realized on the same syllable, the AP-final syllable. This is not predicted by the Jun & Fougeron (2000) model. Both the L and the H tones of the early rise were realized on the same syllable in about 15% of cases. This is not predicted by the Jun & Fougeron (2000) model, but is allowed by the Post (2000) model.

An examination of the influence of segmental structure on the realization of the early H fully supported neither Pasdeloup's (1990) nor Post's (2000) claims about the treatment of consonant- *vs.* vowel-initial words and the influence of underlying *vs.* liaison consonants. Even for consonant-initial words, the early H was most often realized on the second syllable of the proper name. Function word/proper name sets with a liaison environment (e.g., *Mon Arianella*) had the early H in the first content word syllable 23 out of 53 times. The corresponding sets with underlyingly consonant-initial names (e.g., *Ma Marianella*) more rarely had the early H in the first syllable (15 out of 54 times). This difference is significant ($\chi^2 = 6.35, p < .05$). Post (2000) predicts that there should be a difference in this direction, but by her account, we would expect the early H to be realized on the first syllable in the majority of consonant-initial cases.

The timing facts were very different for the the L tone of the early rise and the L tone of the late rise. While the L target of the late rise (Elbow 2) was typically realized within the last full syllable of the AP, the L target of the early rise (Elbow 1) was typically realized late in the last function word syllable of the AP or early in the first content word syllable of the AP, that is, at the boundary between the last function word syllable and the first content word syllable of the AP. Figure 3 shows the location of Elbow 2, the low inflection point associated with the late rise. The x axis shows the duration of the hypothesized associated syllable, here the penultimate syllable of the AP, the syllable to which Jun & Fougeron (2000) predict the L of the late rise to be
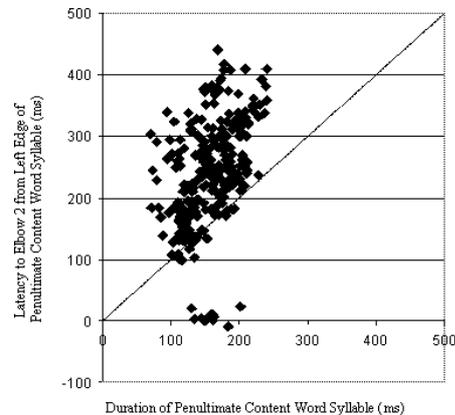
associated. The y axis shows the latency from the left edge of this syllable. The vast majority of the data points have positive values above the $x = y$ line, indicating that the L tone of the late rise was realized not in the penultimate syllable, but in the final syllable of the AP. The small cluster of outliers at the bottom of the graph indicate tokens for which the elbow of the late rise (Elbow 2) was located at the function word-content word boundary. All these tokens are single rise tokens produced by Speaker 1. They are interesting because the timing of the elbow is consistent with the timing of Elbow 1.

A very different pattern is shown in figure 4, a plot of the location of Elbow 1, the low inflection point in cases where the early rise is preceded by a leading low plateau.[2] The data points for Elbow 1 fall in a cluster on both sides of the $x = y$ line, suggesting that the tonal target is neither the last function word syllable of the AP nor the first content word syllable of the AP. One possibility is that the low tone immediately preceding the early rise is a leading tone of a bitonal LH*. If so, the rise time from the low tone immediately preceding the H of the early rise (Elbow 1 in tokens with preceding function words, L1 in other tokens) should precede the early H at some "fairly invariant" interval, following the standard assumption of the timing of leading tones ([7], p. 123). Rise times for the early rise are plotted in figure 5. As the positive slope of the regression line indicates, the low tone does not lead the early H at a fixed distance. Rather rise times increase as the duration of the proper name increases, indicating that the low tone is not a leading tone of a bitonal LH* unit.

An analysis that treats the L tone of Elbow 1 as the spreading of the AP-initial L tone predicts that Elbow 1 will appear in the last function word syllable. Since most of the data points in figure 4 do not fall below the $x = y$ line, that treatment is not supported. Vaissière (1997) argues that the low elbow is associated with a function word and serves to mark the function word-content word boundary (a "marque de frontière" 'boundary marker' ([10], p. 67)). It seems preferable, however, to treat the this elbow as a target associated to the left edge of the first content word rather than the right edge of the last function word. This allows a unified treatment of the low target immediately preceding the

---

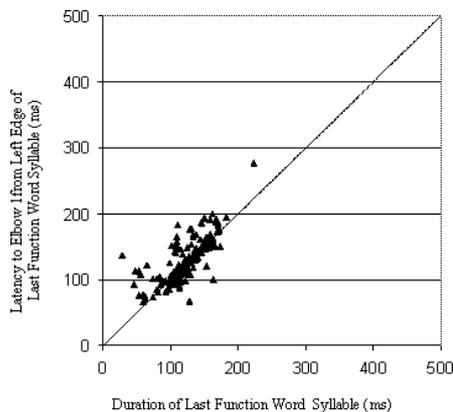[2]A leading low plateau is typical of APs with preceding function words (e.g., *Et Marianna*).

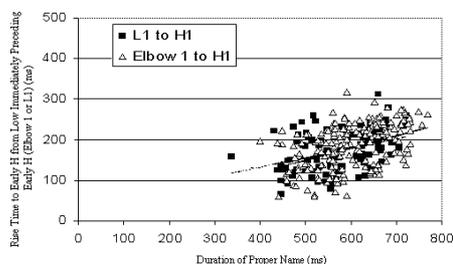*Figure 4: Placement of Elbow 1 (Inflection Point Preceding Early H)*

Figure 5: Rise Time to Early H

early H in cases with and without preceding function words.

4. Discussion

Accounts which assume a one-to-one syllable/tone association predict that only APs 4 or more syllables in length should have the LHLH pattern. Although Jun & Fougeron (2000) do note that the /LHiLH*/ pattern can be realized on APs of fewer than 4 syllables if a syllable is lengthened to accommodate a second tone, the model specifies a one-to-one relationship between syllables and tones. The revised Jun & Fougeron model (to appear, 2002), which discards the one-to-one tone/syllable correspondence, is better able to account for the observed patterns.

The timing of Elbow 1 suggests a model in which the early rise is analyzed with a double association of the L tone of a LH edge tone to the left edge of the AP and to the left edge of the first content word of the AP. This double association is similar to the secondary associations of phrase accents discussed in Grice et al. (2000) for a number of languages, including German, Hungarian, and Romanian, but here the L is associated to another phrase edge, not a stressed or stressable syllable [3]. The observed timing differences also support a model in which the early rise and the late rise are different in nature.

The results also support the long-standing observation of the variability of the realization of the early rise within and across speakers. There were a number of differences between the Marseille speaker and the Tours speaker in this study and between these two speakers and the speakers in other studies. For example, Rolland & Lœvenbruck (2001) report that speakers typically produced the LHLH pattern

for 4- and 5- syllable APs in utterance-initial position [9]. By contrast, the speakers in this study produced the LHLH for only 40% of 4 syllable APs. In addition, Speaker 2 produced 12 utterance-initial APs (of a possible 60) with 4 syllable stretches with no H tones. Such accent-less stretches are reported to be rare in utterance initial position (e.g., Pasdeloup (1990)). It is not clear whether these inter-speaker differences are due to dialect or to other factors.

5. Conclusions

A number of directions for future research present themselves. In order to test hypothesized syllabic association, we could systematically vary the duration of hypothesized target syllables (by varying low and high vowels, for example). We could also investigate claims about the nature of early rises when there is more than one early rise in a phrase. Researchers (e.g., Pasdeloup (1990)) and Jun & Fougeron (to appear, 2002)) have noted that long polymorphemic words are often pronounced with more than one early rise (e.g., *anTIconstiTUtioNEL*). Jun & Fougeron argue that additional early rises are LH phrase accents. If that is the case, we would expect the text-to-tune alignment of these additional rises to be similar to that of the early rises in unexceptional cases (e.g., alignment of the L tone to the edge of some unit, perhaps the morpheme).

6. References

[1] Di Cristo, A.; Hirst, D., 1993. Rythme syllabique, rythme mélodique et représentation hiérarchique de la prosodie du français. *Travaux de l'Institut de Phonétique d'Aix-en-Provence*, 9-24.

[2] D'Imperio, M., *The Role of Perception in Defining Tonal Targets and their Alignment.* Doctoral thesis, The Ohio State University.

[3] Grice, M.; Ladd, D.R.; Arvaniti, A., 2000. On the place of phrase accents in intonational phonology. *Phonology*, 17, 143-185.

[4] Jun, S.; Fougeron, C., 2000. A phonological model of French intonation. In *Intonation: Analysis, Modeling and Technology*, A. Botinis (ed.). Dordrecht/Boston: Kluwer, 209-242.

[5] Jun, S.; Fougeron, C., to appear, 2002. Realizations of accentual phrase in French intonation. *Probus.*

[6] Pasdeloup, V., *Modèle de règles rythmiques du français appliqué à la synthèse de la parole.* Doctoral thesis, Institut de Phonétique d'Aix-en-Provence, Université de Provence, Aix-Marseille.

[7] Pierrehumbert, J.B.; Beckman, M.E., 1988. *Japanese Tone Structure.* Cambridge: MIT Press.

[8] Post, B., 2000. *Tonal and Phrasal Structures in French Intonation.* The Hague: Thesus.

[9] Rolland, G.; Lœvenbruck, H., 2001. *Pertinence physiologique et perceptive du Syntagme Accentuel en français.* Paper presented at the Journées Prosodie, Institut de la Communication Parlée, Centre National de la Recherche Scientifique, Grenoble, October 10-11, 2001.

[10] Vaissière, J., 1997. Langues, prosodies et syntaxe. In *Traitement Automatique des Langues*, 38.1, 53-82. [published by ATALA, Paris].