

# The Realization of French Rising Intonation by Speakers of American English

Scott Lee

<sup>1</sup> Program in Linguistics, University of Georgia, Athens, Georgia, United States

gte577z@uga.edu

## Abstract

This study examines the realization of French intonational rises by adult native speakers of American English. Production data were gathered using a discourse completion task and a storytelling task from eight American college students beginning a semester-long study abroad program in Southern France. Results suggest that speakers struggled with two particular aspects of French intonation: the grouping of words into Accentual Phrases, and the phonetic realization of phrase-final rises. In particular, the probability distribution for the alignment of the late L elbow was bimodal for L2 speakers but unimodal for L1 speakers, suggesting the use in the learner speech of two distinct tonal patterns instead of the single French LH\*. Mean values for overall pitch range and the scaling of continuative rises were significantly lower and less variable than French L1 values as well.

**Index Terms:** intonation, L2 prosody, French

## 1. Introduction

Most studies on second language acquisition have focused on the production and perception of segmental material [see e.g. 1, 2, 3]. To gain a better understanding of the phonetic characteristics of foreign accent, though, it is necessary to explore prosodic as well as segmental characteristics of learner speech. A number of recent studies have addressed this issue, exploring features like tonal alignment [4], tonal phrasing [5], and pitch scaling [6]. This study continues this trend by exploring the intonational characteristics of L2 French spoken by native speakers of American English. In prosodic terms, the two languages have a number of important differences, notably that stress is distinctive in English but not in French, and that intonation phrases (IPs) are built from accentual phrases (APs) with a default tonal pattern in French, instead of from a series of pitch accents as they are in English. The main goal of this study is to examine how L1 English speakers navigate these differences in their spoken L2 French, which will hopefully lend support to previous research findings, as well as shed light on interesting directions for future research on the acquisition of L2 prosody.

### 1.1. Intonation models

Although the scope of this paper is primarily phonetic, it is informed by the intonational phonologies of the languages under investigation. The models of intonation assumed by this paper are the Autosegmental Metrical (AM) systems developed for English by Pierrehumbert and Beckman [7] and for French by Jun and Fougeron [8, 9]. The two systems have a few structural differences that are relevant to this study. First, IPs in French are built from APs, which have a default /LHiLH\*/ tonal sequence. The phonetic realization of this

sequence depends on a number of factors, like speech rate, speech style, and how many syllables are in the phrase. By contrast, IPs in English are built from a sequence of pitch accents on stressed syllables, with nuclear or phrase-final accents being followed by a phrase accent and boundary tone. A second important difference is that English has two bitonal pitch accents, while French only has one. A second rise has been proposed for French [10], but the accent spans two syllables, with the L target appearing on the penultimate syllable and the H target on the final syllable. In English, however, the targets for both pitch accents may associate with a single stressed syllable, producing a contour that is sometimes similar to the French contour (in the case of L+H\*) but sometimes not (in the case of L\*+H). For the latter, the L target aligns with the stressed syllable, and the H target is realized somewhere near the following syllable boundary, producing a “scooped” shape with a relatively pronounced dip in  $f_0$ .

#### 1.1.1. Alignment of the late rise in French

In her study of French tonal structure, Welby [11] reported values for the alignment of tonal targets in the early and late rises. In general, late H\* was aligned with the last full syllable of the phrase, and the late L was realized close to the preceding syllable boundary. The position of the late L was not correlated with the duration of either syllable (penultimate or final), but it appeared in the final syllable 82% of the time. Her conclusions support an AM analysis of French intonation, and they support earlier analyses [see e.g. 12] of the late rise as a bitonal pitch accent.

## 1.2. Previous research

### 1.2.1. Korean and English

Jun and Oh [5] looked at how native speakers of American English acquired the phonology and phonetics of Korean intonation, which shares important phonological characteristics with French intonation. As in French, Korean IPs are built from APs, and the APs are built from rising sequences of L and H tones, with the exact tonal specification depending on the number of syllables in the phrase. To explore the relationship between the speakers’ level of experience with Korean and their spoken proficiency, Jun and Oh designed a set of 40 sentences to test two main intonational features: the grouping of words into APs, and the phonetic realization of tone sequences. Interestingly, their results show that although advanced speakers produced more correct phrasings than intermediate and beginning speakers, they were generally not more successful at realizing the underlying tonal sequences phonetically. Specifically, they found that pitch range in AP-initial HL sequences was significantly smaller for learners than for native speakers, and that AP-final H tones were easier for learners to accurately produce than those elsewhere in the phrase, presumably because they are produced more regularly in L1 Korean and are thus more perceptually salient.

### 1.2.2. Dutch and Greek

Mennen [1] examined bi-directional intonational transfer in Dutch non-native speakers of Greek. Prenuclear or non-final rises have the same phonological structure in both languages, but they differ in their phonetic realization, with peak alignment not only occurring earlier but also being affected by vowel length in Dutch. The study consisted of two experiments, one looking for transfer from the L1 to the L2, and one looking for transfer from the L2 to the L1. Results for the first experiment showed that only one of the speakers was able to produce native-like L2 rises; the remaining four speakers produced L2 rises with alignment patterns from the L1. Results for the second experiment showed that the same four speakers produced L1 alignment patterns that were significantly different from those produced by monolingual L1 speakers, indicating an effect of exposure to the L2. Together, the experiments support the claim that intonational transfer can go both ways, i.e. from the L1 to the L2 and from the L2 to the L1. Crucially, they also support Jun and Oh's findings that the phonetics of L2 intonation are difficult for speakers to acquire.

### 1.3. Research questions

Using AP-final rises in French as a test case, the production experiment was designed to answer a number of questions raised by previous studies about the acquisition of L2 intonation. First, do learners acquire some aspects of L2 intonation (e.g. phrasing) more easily than others, and if so, which ones? If Jun and Oh's results hold true for the speakers in this study, then we would expect them to have more success grouping words into APs than realizing tonal sequences with native-like phonetics. Second, are phonologically similar tonal sequences realized differently in the learners' L2 than they are in their L1, and if so, how? French provides a good platform for testing both of these questions, since it is phonologically and phonetically different from English in terms of its intonation.

## 2. Methods

### 2.1. Participants

Production data were gathered from eight adult native speakers of American English participating in a study abroad program in Montpellier, France. This paper reports the results of the pre-departure pilot study that was conducted immediately before the participants left. Although their proficiency with spoken French varied, all speakers had completed the equivalent of four semesters' (i.e. two years') worth of undergraduate French. Their language backgrounds also varied, with two speakers having completed the same level of coursework or higher in Spanish, and one having working knowledge of Spanish, Italian, and Portuguese; the remaining five had only French as their L2. Production data were then gathered from four of the participants' native French host families for comparison.

### 2.2. Materials

All participants were asked to complete two tasks: a storytelling task, and a discourse completion task (DCT) as outlined by Prieto [13]. The DCT consisted of 31 scripted sentences elicited from the participants, which they read in

response to situational prompts described to them by the interviewer. A sample item is given in (1), where the response is italicized and bolded.

- (1) Tu as acheté de la glace à la vanilla et à la noisette pour ta fête. Demande aux invités s'ils veulent de la glace à la vanilla ou à la noisette.

***Vous voulez de la glace à la vanilla ou à la noisette?***

'You bought two flavors of ice cream for your party, but you're not sure which flavor to serve. Ask the guests at your party which flavor ice cream they'd prefer.

***Do you all want chocolate or vanilla?'***

31 sentences were recorded per speaker in the first task and approximately two minutes of extemporaneous speech in the second. Learners completed the tasks in both languages, while the host families completed them only in French.

### 2.3. Procedures

For the discourse completion task, the interviewer read the situational prompt for each item, and the participants read the target sentence in response. For the storytelling task, participants were asked to tell a story about a social event they recently attended, like a family gathering or party. The tasks were completed during a single interview session and were separated by approximately five minutes of time during which the participants could review the instructions for the upcoming task. Instructions for both tasks were printed on a prompt sheet and were explained by the interviewer at the beginning of the session.

#### 2.3.1. Data analysis

Speaker responses to both tasks were recorded using a Shure SM51 condenser microphone and digitized at a sampling rate of 44.1kHz and a depth of 24kbit/s. Responses to the DCT were segmented by hand and saved as separate files, yielding a total of 496 (16 x 31) phrases of varying length and tonal structure. Responses to the storytelling task were saved as a single wav file. Sound files for both tasks were then loaded into Praat [14] and segmented into phrases, words, syllables, and phones. Praat scripts were used to align phone boundaries to the text, and the boundaries were adjusted by hand to ensure accuracy.

Pitch curves were generated using Praat's built-in pitch tracker, and files containing  $f_0$  perturbations or disfluencies were discarded. Several intonational features were then labeled by hand for each phrase, including  $f_0$  minima and maxima, position of the late H, and position of the late L elbow. AP and IP boundaries were also marked in the French data to allow for the comparison of L1 and L2 phrasing. For the L2 speech, boundaries were marked both when speakers produced grammatical phrase-final pitch accents, and when they produced ungrammatical phrase-medial pitch accents.

Numerical data were exported to a spreadsheet and then analyzed using the R statistical package [15].

2.3.2. ToBI labeling

Learner varieties of intonation have been shown to be phonologically distinct from native varieties [16]. In this study, the speakers produced rises in the L2 that were sometimes phonetically similar to analogous rises in an L1, but sometimes not. Because of this ambiguity, tonal categories were not labeled in the L2 and were analyzed phonetically instead.

3. Results

3.1. Phrasing

Data from the discourse completion task show that learners produced more sentence-medial AP boundaries than native speakers. Mean counts for both boundary types are presented in the table below and organized by speaker group. Values for the French L1 speakers were tightly clustered ( $\sigma=3.7$ ) and are thus reported as an aggregate mean, but values for the L2 speakers were more widely dispersed ( $\sigma=11.38$ ). A Welch's *t*-test indicated that the difference between the means was statistically significant ( $p<0.0002$ ), supporting Jun and Oh's findings for Korean that non-native speakers of French are generally less successful than native speakers at grouping words into phrases.

Speaker	French L1	French L2
1	78	85
2	77	117
3	80	108
4	73	115
5	74	96
6	82	94
7	75	100
8	71	112
	$\mu=76.3$ $\sigma=3.7$	$\mu=103.4$ $\sigma=11.38$

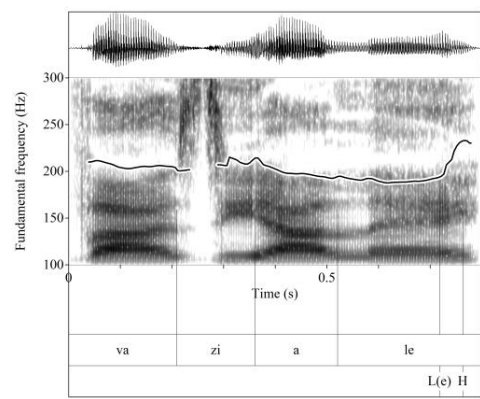
Table 1. Number of sentence-medial APs produced by native and non-native speakers in the DCT.

These results are likely due to two factors: the tendency for the non-native speakers to produce ungrammatical phrase-medial pitch accents, and their tendency to speak slower and with more disfluencies than the native speakers.

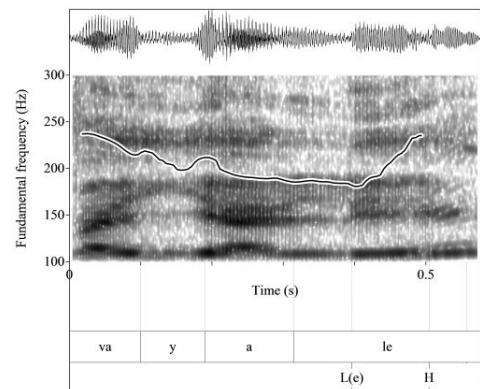
3.2. Tonal alignment

Results from both tasks show significant differences in the phonetic realization of rising tone sequences between the native and non-native speakers. In general, the L2 speakers realized the elbow in the late rise later than the L1 speakers. The placement of this elbow was also more variable. Figure 1 shows a spectrogram and pitch curve for the final four syllables of a sentence-medial AP in the phrase *Comment tu vas y aller* 'How are you getting there?' produced by a native (a) and a non-native (b) speaker. The low elbow in each figure is marked by an L(e), and  $f_0$  is shown in hertz (Hz).

Figure 1. Alignment of late rises in L1 and L2 French speech. a.



b.



The pitch curve in (b) is representative of a typical alignment pattern for late rises reported in [11], with the L elbow being located near the onset of the final syllable and the H target being reached near its end. The slope of the curve increases sharply about halfway through the syllable, but it is clearly rising throughout. By contrast, the pitch curve in (a) is marked by a much later rise in the final syllable, with the L elbow falling close to the H target near the end of the syllable. This pattern was common in the L2 speech, accounting for 45% of the total rises, and was in many cases produced with an even more pronounced delay.

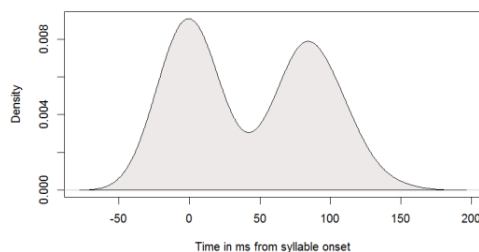
Table 2 presents the mean alignment in ms of the late elbow relative to the onset of the accented syllable. In calculating the means, positive values were entered for positions in the target syllable and negative values for positions in the preceding syllable. The French L2 values show late mean alignment relative to the French L1 values, which are closer to the syllable onset and similar to those reported by Welby [11].

Speaker	French L2 ( $\mu, \sigma$ )		French L1 ( $\mu, \sigma$ )	
1	32.1	39.3	10.5	8.3
2	51.6	58.6	15.6	10.3
3	45.3	33.2	8.5	9.6
4	39.8	37.4	21.3	11.4
5	52.5	49.3	12.3	9.3
6	61.8	56.7	11.2	14.2
7	33.1	36.5	9.3	16.3
8	71.0	58.8	18.1	15.8
All	48.4	46.3	13.35	11.9

Table 2. *Alignment in ms of the L-elbow in AP-final rises for French L2 and L1 speakers.*

The trend that emerges from these data is the close relationship between the mean and standard deviation for the L2 speakers' alignment patterns. In effect, the distribution was approximately bimodal, suggesting the speakers chose one of two options when realizing the rise: early alignment of the L, or late alignment of the L. In cases of early alignment, the elbow most often fell and occasionally before the syllable boundary. The alignment of the late H was relatively invariable, almost always occurring near the syllable boundary (mean latency from the right edge of the phrase was 1.2ms for all speakers, with  $\sigma = .53$ ). A kernel estimation was used to calculate the probability distribution function for the alignment of the late L, shown in Figure 2. In addition to the bimodality, the distribution highlights the fact that the speakers tended to produce more early- than late-aligned L targets, on the whole.

Figure 2. *Probability density function for the alignment of the late L elbow in L2 speech.*



### 3.3. Pitch scaling

Overall pitch range, measured as 80% quantal range in Equivalent Rectangular Bandwidths (ERB) was lower for the learners ( $\mu=2.1\text{ERB}$ ) than for the native speakers ( $\mu=4.3\text{ERB}$ ); the difference was statistically significant ( $p=0.002$ ). After z-score normalization, average values for the scaling of the late rises were shown to be significantly less variable for the learners as well, perhaps indicating a restriction in expressivity caused by a general processing constraint limiting the complexity of syntactic and pragmatic information that they can prosodically code.

## 4. Discussion and conclusion

The production data gathered for this study clearly suggest the presence of interlanguage effects in the non-native speakers' French intonation. In particular, the alignment and scaling of L2 AP-final rises differs significantly from that of the same rises in L1 French. One potential explanation for this difference is phonological. The L2 speakers essentially had two categorical choices when producing the AP-final rise: the scooped bitonal accent  $L^*+H$ , and either the  $LH^*$  or the  $L+H^*$  (the phonetic distinction between these two was not clear in the data). If adult speakers can acquire L2 intonational phonetics, then we would expect advanced learners to produce mostly  $LH^*$  and intermediate and beginning learners to alternate between the  $LH^*$  and the  $L^*+H$ . The results give some support to this hypothesis, with the learners, who self-rated their spoken proficiency in French as intermediate, producing the two contours fairly evenly.

The factors underlying this phonological interference are likely complex. Perceptual factors may play a role if the learners hear the French contour as a generic rise without perceiving (or perhaps recognizing as linguistically relevant) how consistently the tonal targets are aligned in L1 speech. However, pragmatic factors may also be involved, since the two pitch accents are generally assumed to have distinct meanings in English, with the scooped accent in particular signaling uncertainty or hesitation [see e.g. 17,18]; although the alignment of the late L elbow is potentially influenced by pragmatic factors [19], the variation does not appear to be as clearly categorical as it is in English. Perceptual and pragmatic factors are of course not mutually exclusive, and more investigation is needed to determine to what extent they both contribute to the phonological and phonetic characteristics of the learners' L2 intonation.

The findings suggest a number of directions for future research. First, production data from English are needed to determine whether rising intonation in the learners' L1 is affected by their exposure to the L2. Second, perception experiments are needed to look for interference between the learners' L1 and L2, e.g. whether they perceive L2 rises as distinct from L1 rises. Production data from the participants in this study will be gathered after they return from France to examine the effects of prolonged exposure to the L2 on their speech and add to the growing body of longitudinal research on prosodic acquisition. Based on the findings in [5] and [16], the expected result is for the speakers to make improvements to their phrasing, but not necessarily to the realization of the tone sequences. This last component also has the benefit of indirectly testing the effectiveness of language immersion as a pedagogical technique for improving L2 prosody, which is a relatively under-researched area of second language acquisition.

## 5. Acknowledgements

I would like to thank Meghan Armstrong and Keith Langston for their guidance and patience in helping me design the experiments for this study. I would also like to thank Diana Ranson for nurturing my interest in French, and the University of Georgia Graduate School and Office for the Vice President of Research for their generous financial support. Finally, thanks to Pauline Welby for creating the Praat script used to draw the pitch tracks in this paper.

## 6. References

- [1] Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research*, 233-277.
- [2] Best, C. T. (1995). Learning to perceive the sound pattern of English. *Advances in infancy research*, 9, 217-217.
- [3] Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y. I., Kettermann, A., & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*, 87(1), B47-B57.
- [4] Mennen, I. (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics*, 32(4), 543-563.
- [5] Jun, S. A., & Oh, M. (2000, May). Acquisition of second language intonation. In *INTERSPEECH* (pp. 73-76).
- [6] Mennen, I., Schaeffler, F., & Docherty, G. (2007). Pitching it differently: A comparison of the pitch ranges of German and

- English speakers. *16th International Congress of Phonetic Sciences*, 1769-1772.
- [7] Beckman, M., & Pierrehumbert, J. (1986). Intonational structure in Japanese and English. *Phonology yearbook*, 3(1), 5-70.
- [8] Jun, S. A., & Fougeron, C. (2000). A phonological model of French intonation. In *Intonation* (pp. 209-242). Springer Netherlands.
- [9] Jun, S. A., & Fougeron, C. (2002). Realizations of accentual phrase in French intonation. *Probus*, 14(1), 147-172.
- [10] Delais-Roussarie, Elisabeth; Post, Brechtje; Avanzi, Mathieu; Buthke, Carolin; Di Cristo, Albert; Feldhausen, Ingo; Jun, Sun-Ah; Martin, Philippe; Meisenburg, Trudel; Rialland, Annie; Sichel-Bazin, Rafèu & Yoo, Hi-Yon (to appear: 2014). "Intonational phonology of French: Developing a ToBI system for French". In Frota, Sónia & Prieto, Pilar (eds.) *Intonational variation in Romance*. Oxford: Oxford University Press.
- [11] Welby, P. (2004). The structure of French intonational rises: A study of text-to-tune alignment. In *Speech Prosody 2004, International Conference*.
- [12] Post, B. (2000). *Tonal and phrasal structures in French intonation* (Vol. 34). Thesus.
- [13] Prieto, Pilar (2001). 'L'entonació dialectal del català: El cas de les frases interrogatives absolutes', in A. Bover, M.-R. Lloret, and M. Vidal-Tibbits (eds.), *Actes del Novè Colloqui d'Estudis Catalans a Nord-Amèrica*. Barcelona: Publicacions de l'Abadia de Montserrat, 347-377.
- [14] Boersma, Paul & Weenink, David (2014). Praat: doing phonetics by computer [Computer program].
- [15] R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- [16] Mennen, I., Chen, A., & Karlsson, F. (2010). Characterising the internal structure of learner intonation and its development over time. In *Proceedings of The 6th International Symposium on the Acquisition of Second Language Speech (Newsounds 2010)*.
- [17] Ward, G., & Hirschberg, J. (1985). Implicating uncertainty: The pragmatics of fall-rise intonation. *Language*, 747-776.
- [18] Hirschberg, J. (2004). Pragmatics and intonation. *The handbook of pragmatics*, 515-537.
- [19] Welby, P. (2006). French intonational structure: Evidence from tonal alignment. *Journal of Phonetics*, 34(3), 343-371.