

Uncovering the effect of imitation on tonal patterns of French Accentual Phrases

Amandine Michélas & Noël Nguyen

Laboratoire Parole et Langage & Aix-Marseille Université, Aix-en-Provence, FRANCE

michelas@lpl-aix.fr, noel.nguyen@lpl-aix.fr

Abstract

French accentual phrases (APs) are characterized by the presence of a typical final *fo* rise (LH*) and an optional/additional initial *fo* rise (LHi). This study tested whether between-speaker speech imitation influenced the realization of APs tonal patterns. The experiment was based on APs containing a function word plus a bisyllabic content word, whose tonal patterns differed in the potential placement of an optional/initial high tone (Hi). In two shadowing tasks (without/with explicit instructions to imitate the speaker's way of pronouncing the stimuli), participants produced more initial high tones when they heard a stimulus including both initial and final high tones relative to stimuli which only a final high tone was present. Thus, imitation influences the realization of APs tonal patterns in French.

Index Terms: imitation, convergence, tonal pattern, accentual phrase, initial high tone, intonation, French.

1. Introduction

In the last few years, a growing number of studies have shown that, in a conversational interaction, the behavior of each talker may become more similar to the other talker's behavior. This tendency shown by speakers to imitate each other is referred to as convergence or imitation. Imitation seems to occur at every level of the conversational exchange, and that includes the phonetic level. It has been shown that convergence affects both segmental features (e.g. [1], [2]) and suprasegmental features such as rate of speech, pauses [3], vocal intensity [4], pitch curve [5], [6], level pitches [7] or *fo* extremum timing continua [8], [9]. To characterize imitative patterns in speech, several studies [5], [10], [11], [12] have used the shadowing paradigm. Thus, Bosshardt et al. [5] showed that German nonstutterers and stutterers correctly reproduced contrastive stress occurring in the stimulus they had to repeat. This result seems to suggest that the speakers' prosodic structure is influenced by the prosodic structure of the speaker they heard. In an imitation task where subjects were asked to mimic a set of English intonation contours and then to imitate themselves in several successive sessions, Braun et al. [13] showed that the produced *fo* contours gradually converged towards a limited set of basic English intonation patterns. According to these results, the convergence did not seem to occur immediately but after several iterations of the same contour. While previous research on convergence in the intonational domain has focused on stress-accent languages such as English or German, no data are yet available for French. Hence this study represents the first quantitative description of accentual phrases tonal patterns in French in a shadowing context and establishes a previously undocumented link between imitation and initial high tone in French.

It is generally agreed that the French prosodic structure includes two organizational levels: the intonation phrase and a lower level in the prosodic hierarchy called the accentual phrase or AP [14], [15]. The AP is characterized by

the presence of a typically final *fo* rise (LH*) on the last syllable of the phrase which is lengthened. In addition, the AP optionally includes an initial *fo* rise (LHi) near the beginning of the AP (i.e. generally on the first syllable of the first content word occurring in the phrase) which is not associated to final preboundary lengthening. The final rise could be compared to the pitch accent associated with metrically strong syllables or moras in English [16] whereas the initial rise is less clearly linked with overall prominence and could be considered as a phrase accent which indicates the beginning of the AP [17]. In this work, we analyzed the AP tonal patterns within the model of French intonation proposed by Jun & Fougeron [14], [15]. In this model, developed in the framework of the autosegmental metrical theory of intonation [16], [18], the early and late rises of the AP can be described as a sequence of Low and High tones. The initial rise consists of a sequence of LHi tones (i indicate the initial rise) while the final rise consists of an LH* tone sequence (* indicates that the tonal sequence is associated to a metrically strong syllable). When all four tones are realized, the tonal pattern of the AP is LHiLH* (i.e. the default pattern postulated for the AP). The model also predicts five additional surface shapes formed by the absence of one or more tones of the default pattern: LLH* when Hi is undershot, LHiH* when the second L is undershot, HiLH* when the initial L is undershot and LH* when both Hi and the following L are undershot.

Several kinds of factor such as phrase length [17], [19] or speech rate [20], [21] affect the realization of AP tonal patterns. More specifically, it appears that APs in Subject position within the utterance, containing a bisyllabic content word preceded by a function word, are preferentially realized with a LH* pattern at normal speech rate while tonal patterns including an initial H tone (LHiLH*, LHH* or HiLH*) are less frequent in the data. Figure 1 shows the AP "Le mari" *the husband* (extracted from the sentence "Le mari d'Amanda réclamait l'institutrice" *Amanda's husband was asking for the teacher*) realized with an LH* tonal pattern (including only a final rise, left) or with an additional/optional initial rise (LHiLH* tonal pattern, right).

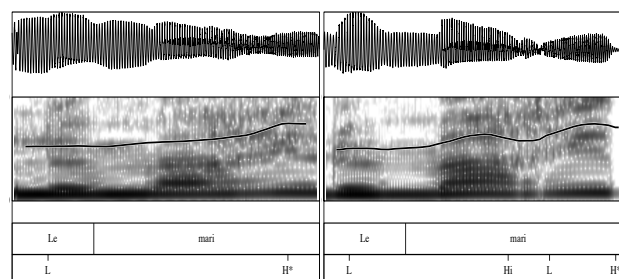


Figure 1. Example for two renderings of the first AP "le mari" (*the husband*), within the utterance "Le mari d'Amanda réclamait l'institutrice" (*Amanda's husband was asking for the teacher*) realized with either only a final rise (LH*, left) or with an additional/optional early rise (LHi, right).

In this experiment we examined the effect imitation on the realization of APs in French. We asked speakers to shadow APs differing in the potential occurrence of an initial high tone on the first syllable of the content word, first without any instructions to imitate auditorily presented stimuli (repetition task), and then with explicit instructions to imitate the speaker's way of producing the stimuli (imitation task). Our hypothesis was that speakers would correctly reproduce the tonal pattern which was present in the original input both in the repetition and imitation tasks. We expected a stronger imitation effect when speakers were explicitly asked to imitate phrases than when they were not.

2. Method

Pairs of sequences, whose segmental structure was identical but which differed in the potential placement of an initial high tone near the beginning of the first AP were presented to listeners. Consider the two French noun phrases shown in Figure 2. The two sequences are identical regarding their segmental structure but differ in the occurrence of an optional/additional initial high tone on the first syllable of the content word “maison” house of the first AP.

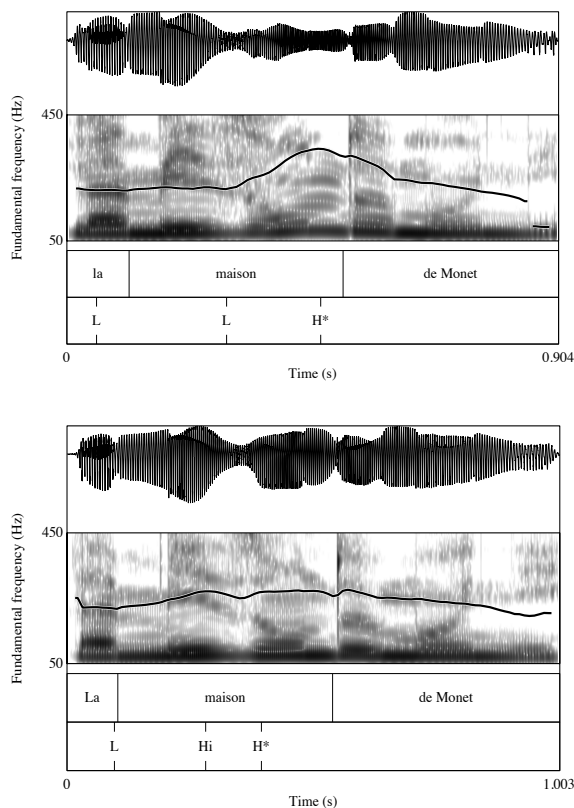


Figure 2. *Fo* contour for two noun phrases “La maison de Monet” (Monet’s house) pronounced in isolation (as an Intonation Phrase) with only a final rise (LH*, top) and with an additional/optional initial high tone (LHiH*, bottom).

In the first part of the experiment, participants listened to target phrases and were instructed to repeat them (repetition task). In the second part of the experiment, participants listened the phrases again in a different order and were explicitly asked to imitate the speaker that had been recorded (imitation task). Our hypothesis was that participants would reproduce an initial high tone if it was present in the

stimulus in both tasks. We also expected that participants would produce more initial high tones in the imitation task than in the repetition task.

2.1. Materials

A total of 42 noun phrases were constructed. In each noun phrase the target AP, which was in initial position within the sequence, was a 3-syllable AP including a monosyllabic determinant and a bisyllabic noun. Each sequence also contained a second AP, which was always a 3-syllable prepositional phrase (also containing a monosyllabic function word and a bisyllabic noun). This second AP was added to make target APs non-final within sequences. Target APs included only voiced segments in order to facilitate *fo* analyses.

A native speaker of French (the first author) produced the phrases employed for the stimuli. The speaker read aloud the list of phrases at a self-selected normal rate, first as naturally as possible and then with an initial rise associated to the first AP of each phrase. For each sequence we checked if the speaker produced the expected tonal pattern on target APs using a signal editor.

Fo analyses revealed that the first syllable of the target content word was produced with no H initial target when the speaker produced the sequences as naturally as possible (Without Hi stimuli) and with an H initial target near the beginning of the first content word of the target AP when she produced initial rises (Including Hi stimuli). Specifically *fo* analyses revealed that the first syllable of the target content word was produced with higher *fo* values in the Including-Hi condition than in the Without-Hi condition ($t(39) = 8.2$, $p < 0.0001$). Duration analyses revealed that the first syllable of target content words did not differ in duration between the two conditions.

On the basis of the phrases thus produced, two lists of sentences were constructed so that the two members of a pair appeared in different lists. All of the participants heard the two lists twice (first in the repetition task and then in the imitation task). In each list, 24 distractor phrases were added to the 21 experimental sequences. For each task, half of the participants had List 1 first and then List 2. The reverse was true for the other half of participants. Within each list the order of presentation was both random and different across participants.

2.2. Participants and procedure

Six native speakers of Southern French (3 males and 3 females) took part in the experiment. The participants ranged in age from 23 to 42 with an average of 28.6. Each participant was recorded individually in a soundproof chamber onto digital audiotape (DAT) at 64kHz downsampled to 20kHz. First, participants were told that they would hear phrases and would have to repeat them. They typically started repeating each sequence from its acoustic offset. In the second part of the experiment, participants were told that they would hear the same stimuli as in the first part but in a different order, and would have to imitate the speaker’s way of producing the stimuli as closely as possible but within a pitch range they felt comfortable with. Before the experiment began, participants received 6 practice phrases. The whole experimental procedure was controlled using E-Prime.

2.3. Data analysis

Items containing disfluencies or *fo* perturbations in the critical region were excluded (2% of the data). The sound files were

segmented and each sequence saved as a separate file. *fo* curves and spectrograms were created using Praat. Each target accentual phrase was tonally labeled according to the tonal pattern inventory proposed by Jun & Fougeron [14],[15]. Syllable boundaries of the first AP were tagged using waveforms and spectrograms to guide segmentation. Late H and Early H were also tagged by hand using the definition given in Welby [17]:

Late H: An *fo* peak in the last full syllable of the phrase (or just beyond it)

Early H: An *fo* peak in one of the first two syllables (excluding the second syllable of two syllable content words)

The two H tones (initial and final) corresponding to two *fo* maxima (respectively in one of the two first syllables of the AP and in the last syllable of the AP).

3. Results

Figure 3 presents the percentage of tonal patterns produced by participants which could include either only a final H target (Without Hi Response) or an additional/initial H target (Including Hi Response) during the two tasks for both Without Hi Stimuli and Including Hi stimuli.

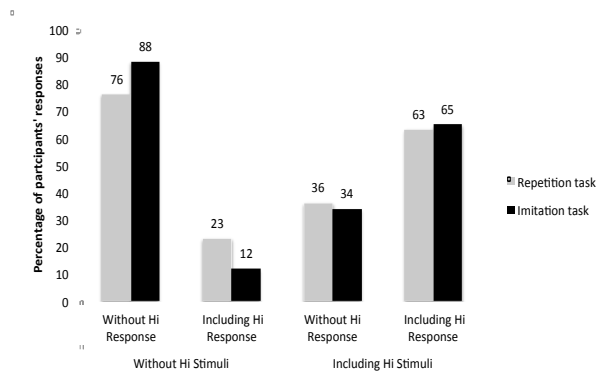


Figure 3. Percentages of responses including only a final H tone (Without Hi responses) or with an additional initial H tone (Including Hi responses) for stimuli including only a final H tone (Absence of Hi in the stimuli) or an additional initial H tone (Presence of Hi in the stimuli) for both repetition and imitation tasks.

A Mixed Logit Model (MLM) was used to examine the factors influencing the occurrence of an initial H tone in participants' responses. The dependent binary variable was the realization of the AP tonal pattern including an initial H tone with values coded as 1 (for occurrences including both initial and final H tones) or 0 (for occurrences including only final H tones). The fixed effects were the auditory stimuli (Without Hi/Including Hi) and the task (repetition/imitation). Random effects were included for speakers and for phrases. The MLM results showed that the tonal pattern of the auditory-presented stimuli has a significant effect on the tonal patterns they produce ($\beta=2.8$, $se=0.49$, $z=5.79$, $p < 0.0001$) while the kind of task had no significant effect ($\beta=1.0$, $se=0.5147$, $z=1.867$, $p=0.052$). The model estimated the probabilities for participants to produce a tonal pattern including an initial high tone for both auditory-presented stimuli and for the two tasks. The estimated probabilities are scaled from 0 to 1 and are given in Table 1.

Repetition task		
Auditory Stimuli	Without Hi Response	Including Hi Response
Without Hi	0.7101	0.1998
Including Hi	0.238	0.6349
Imitation task		
Auditory Stimuli	Without Hi Response	Including Hi Response
Without Hi	0.8010	0.1799
Including Hi	0.0987	0.6597

Table 1. Probabilities estimated by the MLM model for the occurrence of an initial H target in tonal patterns produced by participants. Probabilities are scaled from 0 to 1 relative to the type of auditory stimuli heard by participants (without Hi stimuli or Including Hi stimuli) and the task (repetition/imitation).

The model included 504 values. The interaction between the two fixed factors was not significant ($\beta=-1.2$, $se=0.6358$, $z=1.886$, $p=0.051$) and no other significant effect was found. Examples of typical repetitions of stimuli in the Without Hi Stimuli condition (top) and in the Including Hi condition (bottom) are shown in Figure 3.

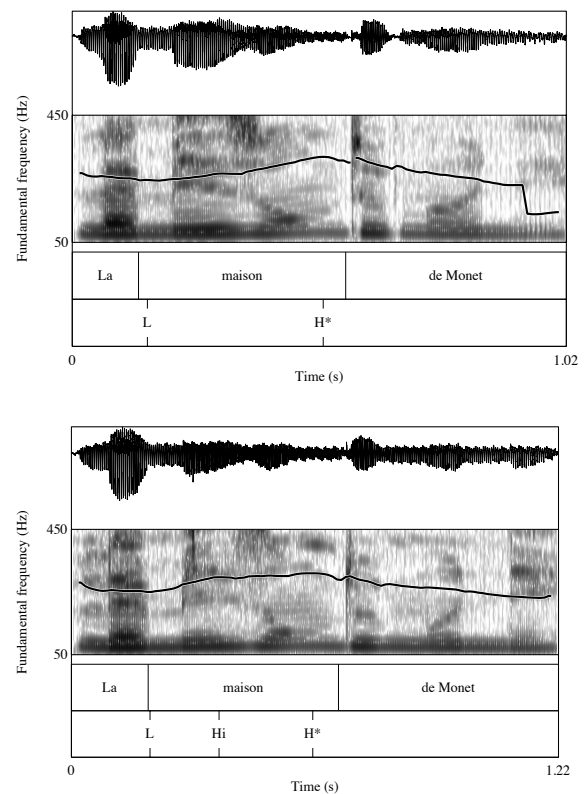


Figure 3. Examples of typical imitations of stimuli in the Without Hi Stimuli condition (top) and in the Including Hi condition (bottom). Shown are responses to the stimuli illustrated in figure 2 for subject MG.

4. Discussion

Our findings concern the types of tonal patterns of accentual phrases produced by participants in this experiment including both a repetition task and an imitation task. Participants readily differentiated among the stimuli including only a final H tone and stimuli including an additional/optional initial H tone and this in both tasks. Specifically participants produced more initial H tone when they heard a stimulus including both initial and final H tones relative to when only a final rise was present. Note that participants responded with more tonal

patterns including only a final H tone independent of the type of auditory-presented stimuli and independent of the task. This is in line with previous studies showing that APs containing a bisyllabic noun plus a monosyllabic function word are preferentially realized with a LH* pattern while the number of patterns including an initial H tone increase with the number of syllables within the phrase [17], [21]. Additionally, participants did not significantly produce more initial H tone when they heard a stimuli including initial H target in the imitation task compared with than the shadowing task. This result suggests that participants were influenced by the auditory stimuli in the same way in both tasks. While a previous study showed that convergence only occurs after several iterations of the same contour [13], we observed a mimicry phenomenon from the very first occurrence of tonal pattern. Indeed, all of our six participants correctly imitated the first auditorily-presented tonal pattern. Furthermore, note that while participants heard two lists of auditorily-presented stimuli for each task, the MLM model did not show a significant effect of list, and this suggests that participant imitate in the same way the tonal pattern they heard in both parts of each task.

These results lead to a number of conclusions about the way speakers generate prosodic structure. We know that tonal realizations are influenced by linguistic factors such as phrase length [17] or speech rate [21]. Our results confirm that tonal structure is also influenced by the prosodic structure of the speaker whose voice participants were exposed to. Specifically, participants correctly reproduced an initial high tone in their production when the original input also included an H initial target, hence suggesting that the presence of an initial high tone might also be conditioned by tonal patterns of phrases heard by participants. This result, in line with a growing number of studies, highlights the need for an integration of between-speaker accommodation mechanisms within models of speech production.

Moreover, we may wonder if the effect of imitation we obtained resulting of a mimetic phenomenon between the participants and the original input or if the auditorily-presented stimuli affected the participants' mental representation of intonation. Several factors (such as rate or number of syllables within the phrase) could induce a modification of the default AP tonal pattern. As it is the case for linguistic factors, we account for a modification of the AP tonal pattern depending on imitation phenomenon (i.e. more complex AP patterns such as LHiLH* would be produced when another speaker's antecedent speech contains a lot of complex patterns. We don't know yet if these modifications affect mental representation of the AP (yielding the effect that the AP patterns produced would only show complex patterns on the surface realization). The next step is to test if participants who repeated complex APs tonal patterns in a pre-task, would judge more plausible APs complex patterns than APs minimal pattern. This next investigation will allow us to study if auditory-presented stimuli affect participants' mental representation of intonation.

5. Conclusions

Results from two shadowing tasks showed that participants produced more initial high tone when they heard a stimulus including both initial and final high tones relative to when only a final high tonal target was present. These findings establishes a previously undocumented link between imitation and initial high tone in French suggesting that tonal patterns produced by speakers are also conditioned by tonal patterns of phrases they heard.

6. References

- [1] Pardo, J.S. On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America*, 119:2382–2393, 2006.
- [2] Delvaux, V. & Soquet, A. The Influence of Ambient Speech on Adult Speech Productions through Unintentional Imitation. *Phonetica*, 64, 145-173, 2007.
- [3] Giles, H., Coupland, N. & Coupland, J. Accommodation theory: Communication, context, and consequence. In H. Giles, N. Coupland & J. Coupland (Eds.) *Contexts of Accommodation: Developments in Applied Sociolinguistics*, 1-68. Cambridge University Press, Cambridge, UK, 1991.
- [4] Natale, N. Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *Journal of Personality and Social Psychology*, 37:790-804, 1975.
- [5] Bosshardt, H-G., Sappock, C. Knipschild, M. & Hölscher, C. Spontaneous imitation of fundamental frequency and speech rate by nonstutterers and stutterers. *Journal of Psycholinguistic Research*, 26: 425-428, 1997.
- [6] Gregory, S.W., Webster, S. & Huang, G. Voice pitch and amplitude convergence as a metric of quality in dyadic interviews. *Language and Communication*, 13: 195-217, 1993.
- [7] Xu, Y. & Sun, X. Maximum speed of pitch change and how it may relate to speech *Journal of Acoustical Society of America*, 11, 1399-1413, 2002.
- [8] Pierrehumbert, J., & Steele, S. A. (1989). Categories of tonal alignment in English. *Phonetica*, 46, 181-196.
- [9] Redi, L. (2003) Categorical effects in the production of pitch contours in English. In *Proceedings of the 15th International Congress of the Phonetic Sciences*, Barcelona, 2921-2924, 2003.
- [10] Goldinger, S.D. Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105:251-279, 1998.
- [11] Nye, W.P & Fowler, C.A. Shadowing latency and imitative: the effect of familiarity with the phonetic patterning of English. *Journal of Phonetics*, 31:63-79, 2003.
- [12] Dilley, L. & Brown, M. Effects of pitch range variation on F0 extrema in an imitation task. *Journal of Phonetics*, 35, 523-551, 2007.
- [13] Braun, B., Kochanski, G., Grabe, E., Rosner, B. S. Evidence for attractors in English intonation", *Journal of Acoustical Society of America*, 119(6) 4006–4015, 2006.
- [14] Jun, S.A., & Fougeron, C. A phonological model of French intonation. In A. Botinis (Eds.), *Intonation: Analysis, modelling and technology*. Kluwer Academic Publishers: Boston, MA, 209-242, 2000.
- [15] Jun, S.A., & Fougeron, C. Realizations of accentual phrase in French. *Probus*, 14, 147-172, 2002.
- [16] Pierrehumbert, J. *The Phonetics and Phonology of English Intonation*. Doctoral Dissertation, Massachusetts Institute of Technologies, 1980.
- [17] Welby, P. French intonational structure: Evidence from tonal alignment. *Journal of Phonetics*, 34(3): 343–371, 2006.
- [18] Ladd, D.R. *Intonational Phonology*. Cambridge: Cambridge University Press, 1996.
- [19] Astésano, C. *Rythme et Accentuation en Français : Invariance et Variabilité stylistique*, Paris, L'Harmattan,, 2001.
- [20] Fougeron, C., & Jun, S.-A. Rate effects on French intonation: prosodic organization and phonetic realization. *Journal of Phonetics*, 26, 45-69, 1998.
- [21] Michelas, A & D'Imperio, M. Multiple sources of phrase-final lengthening in French: the interplay between prosody, syntax and speech rate, *Journal of Phonetics*, submitted.
- [22] Di Cristo A. Vers une modélisation de l'accentuation du français : première partie. *French Language Studies*, 9, 143-179, 1999.