

CAN A PROSODIC PATTERN INDUCE/REDUCE THE PERCEPTION OF A LOWER-CLASS SUBURBAN ACCENT IN FRENCH?

Philippe Boula de Mareuil^a & Iryna Lehka-Lemarchand^b

^aLIMSI-CNRS, Orsay, France; ^bUniversité de Rouen, Rouen, France

philippe.boula.de.mareuil@limsi.fr; iryna.lehka-lemarchand@univ-rouen.fr

ABSTRACT

Previous sociophonetic work suggested that an atypical prosodic pattern, namely a word-final sharp pitch fall, could characterise the French vernacular of youth living in working-class suburbs (the so-called “suburban accent”). A question we investigate in this study is whether the presence/absence of these prosodic patterns increases/decreases the perception of the suburban accent. Using prosody modification and re-synthesis, perceptual experiments were conducted. Results involving listeners from two French regions are rather robust. They show that utterances with (respectively without) high-low pitch falls are perceived as presenting a higher (resp. lower) degree of accentedness.

Keywords: prosody, perception, sociophonetics, French accents

1. INTRODUCTION

In France, some linguistic features may be specific to lower-class youth [2, 6]. Because of the empirical difficulties to carry out phonetic fieldwork, the lexicon (little demanding in terms of tools) is the favoured level of analysis [4]. Also, most folk comments relate to lexical innovations, even though pronunciation peculiarities may be perceived more readily. The latter are often referred to under the term *accent de banlieue* (‘suburban accent’) because many youth perceived as having this accent live in suburban areas. In French (at least in France), the term *banlieue* does not have the middle-class connotations that it does in the Anglo world; it defines a very different socio-economic space, often associated with violence, unemployment and immigration.

Phonetic studies usually focus on the segmental aspects of the *banlieue* accent: e.g. the pharyngealisation of /ʁ/, the palatalisation/affrication of /t/ and /d/ before high front vowels [1, 7]. Particular prosodic patterns have also been reported: a relative lengthening of the penultimate syllable [2, 3, 6] and/or a pitch rise followed by an abrupt fall on the final syllable before a prosodic boundary [5, 8]. This final fall was observed in the suburbs of Paris [3] and Rouen (located 80 miles

west from Paris, in Normandy), where it was examined in detail [8]. Most often, it starts from a high (H) tone and is not associated with lengthening.

The present study investigates how this prosodic pattern in adolescents from the suburb of Rouen is perceived by listeners living in Normandy and listeners living in the Paris region. The following questions are addressed. What is the contribution of this sharp fall (hereafter referred to as HL fall) to an impression of suburban accent? Does its presence (resp. its absence) entail the perception of a higher (resp. a lower) degree of accentedness? To what extent does perception depend on subjects’ region and social proximity to the suburb of Rouen? If a final HL fall triggers the perception of a suburban accent, whether listeners are from Normandy or from the Paris region, we will be in a position to claim that this prosodic pattern is a genuine ingredient of the French suburban accent.

This study is based on spontaneous speech collected among a ten suburban adolescents. For each speaker, we selected prosodically marked and unmarked utterances (i.e. exhibiting HL falls or not) and generated their counterparts (i.e. prosodically unmarked and marked utterances, respectively) by prosody modification and resynthesis. The corpus and method used are introduced in further detail in the next section (Section 2). Based on this material, a perceptual experiment was conducted, as described in Section 3. Results are presented in Section 4 and discussed in Section 5 from a sociolinguistic perspective.

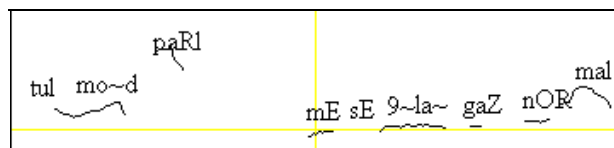
2. CORPUS AND METHOD

The corpus consists of recordings from 11 male adolescents aged from 14 to 17. All the speakers lived and went to school (a middle/junior high school where the data were collected) in the north-eastern suburb of Rouen. Two speakers had metropolitan French parents, another two were from French overseas departments, and the others were of African descent. The recordings analysed here were made during interviews directed by one of the authors, who lived in Rouen and asked the participants to talk about their lives.

For each speaker, at least one prosodically marked utterance (exhibiting 1-3 HL falls within a duration of about 10 seconds) and one prosodically unmarked utterance (showing no such pattern also within a duration of about 10 seconds) were selected. Utterances with lexical indices such as slang, loans or word games (the so-called *verlan*) were avoided. HL falls were defined on the basis of strict acoustic criteria: they had to be word-final and present a slope steeper than -5 semitones between the end and the beginning of the syllable. Applying these criteria, prosodically marked utterances were rendered unmarked by smoothing their HL falls (especially by lowering the H tone). This was done by using the Praat software [9], and more specifically the implemented version of the PSOLA algorithm to synthesise the modified speech signal. In the same way, prosodically unmarked utterances were rendered marked by modifying their pitch contours: 1 to 4 word-final HL falls per utterance were generated (especially by raising the H tone in the syllable onset).

Two speakers were set aside for a familiarisation phase with a couple of sound examples at the beginning of the perceptual experiment. For the test proper, two prosodically marked and two prosodically unmarked utterances (plus their counterparts modified by synthesis) were kept for two speakers. Only one prosodically marked utterance and one prosodically unmarked utterance (plus their counterparts modified by synthesis) were kept for the remaining 7 speakers. Hence, the test corpus was composed of 44 utterances: 11 marked original (MO), 11 unmarked synthesised (US), 11 unmarked original (UO) and 11 marked synthesised (MS) utterances. There were as many high-low falls in the MO and MS utterances: 22 (i.e. 2 per utterance on average). An illustration of an MO pitch curve excerpt, with two HL falls on word-final vowel nuclei is provided in Figure 1.

Figure 1: Pitch curve of the end of the marked original (MO) utterance *on parle pas le langage que tout le monde parle mais c'est un langage normal*. The syllables [paRl] and [mal] correspond to pitch falls of -9 and -5 semitones (ST).



The HL falls of the MO and MS utterances had average slopes of -8.4 ST and -8.8 ST respectively. The counterpart slopes of the US and UO utterances were -2.0 ST and -0.7 ST respectively.

3. PERCEPTUAL EXPERIMENT

3.1. Task and protocol

The resulting stimuli were used in perceptual tests in order to determine how the prosodic patterns under investigation are evaluated by listeners. The participants were warned that the experiment dealt with the so-called “suburban accent” and that they would listen to excerpts of original or acoustically modified speech. They were advised to use headphones or earphones. After some autobiographical information (age, education level, place of residence, etc.), they were asked a couple of questions:

(A) Do you hear (or have you recently heard) young people from working-class suburbs speaking to each other or to other people (on the street, public transport, etc.)?

(B) Do you talk (or have you talked) to them?

In both cases, the proposed answers were: (0) never; (1) seldom; (2) a few times a month; (3) at least once a week.

After a familiarisation phase, during which the subjects were presented samples of MO, and MS utterances, the perception task proper consisted of:

- rating the degree of accentedness on a continuous 0–5 scale, and, optionally
- specifying if certain words or syllables sounded particularly marked by the suburban accent.

The proposed degrees were paraphrased in the following way: (0) no accent, (1) mild accent, (2) moderate accent, (3) rather strong accent, (4) strong accent, (5) very strong accent. Listeners used a mouse to move a slider from the default position which was 2.5. Also, they could type their comments on particular words or syllables in a small text window. They were explicitly instructed to judge the pronunciation — not the vocabulary nor grammar.

The perceptual experiment was conducted through a web-based interface to read the instructions, listen to the stimuli and capture the responses automatically. The test stimuli were presented in random order (different for each subject). This precaution was felt all the more necessary as subjects had no other reference point than the familiarisation phase and their everyday experience to assign a degree of accentedness. Participants could listen to each stimulus as often as they wished, but it was not possible to correct previous responses once a new stimulus was displayed.

At the end of the test, the subjects were requested to indicate the most characteristic pronunciation traits of the speakers they listened to. They were asked the following questions:

(C) Where do these youngsters live, according to you?

(D) Did many (at least 10) utterances sound artificial?

The test did not mention that the recordings came from the neighbourhood of Rouen. Nor did it specify that 22 stimuli were actually modified acoustically. Question C was deliberately open, but subjects had to answer it with the most precision possible. Question D was intended to check the naturalness of the prosody manipulations.

3.2. Participants

Forty subjects took part in the experiment, which lasted 20-30 minutes. They had no known hearing impairment and they were not paid for their participation. Half of them lived in Normandy, the other half in the Paris region. Their familiarity with the suburban accent was estimated from their responses to questions A and B reported above. Results are presented in Table 1, in terms of passive or indirect exposure (question A) and active or direct exposure (question B).

The 20 listeners from Normandy (4 males, 16 females, aged 19 on average) all lived in the department of Rouen (Seine-Maritime), and French was their mother tongue. They were all undergraduate students, except two of them.

The 20 participants living in the Paris region (12 males, 8 females, aged 28 on average) were all graduate native speakers of French, except one of them. For most listeners the passive exposure to the suburban accent was rather frequent, even though the active exposure was rarer.

Table 1: Listeners' passive/active exposure to the suburban accent (0 = never; 1 = seldom; 2 = a few times a month; 3 = at least once a week). The numbers of listeners (out of 20 from Normandy and 20 from the Paris region) are tabulated.

Listeners	Normandy				Paris region			
	0	1	2	3	0	1	2	3
passive	0	2	5	13	0	2	4	14
active	4	13	2	1	4	9	4	3

In age, educational background and geographical proximity, the listeners from Normandy were thus closer to the test speakers than were the listeners from the Paris region. Note however that the active exposure to the suburban accent was somewhat more frequent in the subjects from the Paris region.

4. PERCEPTUAL RESULTS

Let us first examine the responses to questions C and D. Question C: 8 listeners from the Paris region and 4 listeners from Normandy associated the accent they had listened to with the suburb of Paris. The other listeners from the Paris region usually had no idea or

referred to the northern half of France with no more precision. The listeners from Normandy were more accurate in their answers, referring to Normandy or the vicinity of Rouen. This result is interesting because, to some extent, it enables us to generalise our findings to other suburbs of the northern half of France, with Paris playing an emblematic role.

Question D: only 4 subjects from the Paris region and 11 from Normandy perceived over 10 stimuli as sounding artificial (whereas each subject listened to 22 acoustically modified stimuli). The latter result is reassuring as far as the naturalness of the prosody modifications is concerned.

4.1. Ratings

The results of the rating task are reported in Table 2. Analyses of variance (ANOVAs) were carried out on the listeners' ratings with the random factor Subject, the between-subject factor Group (Normandy or Paris region) and the within-subject factor Type of stimulus (MO, US, UO or MS). There is no main effect of listeners' Group, but the effect of the Type of stimulus is significant [$F(3, 114) = 30.767$; $p < 0.001$]. The interaction Group×Type is marginal.

Table 2: Degree of accentedness rated on a 0–5 scale by listeners from Normandy and the Paris region (and averaged), for MO, US, UO and MS stimuli.

Stimuli	MO	US	UO	MS
Normandy	2.65	2.43	2.14	2.40
Paris region	2.78	2.47	2.14	2.44
Average	2.71	2.45	2.14	2.42

In both groups of listeners, MO stimuli are perceived as more marked than US stimuli (average difference = 0.27), and MS stimuli are perceived as more marked than UO stimuli (average difference = 0.28). US stimuli are perceived as slightly more marked than MS stimuli, but the difference, contrary to all the other cross-type differences, is not significant according to pairwise *t*-tests.

Out of the 22 synthesised stimuli, only one MS stimulus is judged as less accented than its UO counterpart and 2 US stimuli are judged as more accented than their MO original counterparts. These exceptions come from 3 different speakers.

The MO–US and MS–UO rating differences are all consistently positive for the two speakers who had two MO stimuli and two UO stimuli each. MO–US and MS–UO rating differences are 0.34 on average (MO: 2.58; US: 2.26; UO: 2.27; MS: 2.62). This result suggests that for a given speaker the contribution of prosody is relatively independent of the segmental content.

The influence of the number of HL falls (1-3 for the MO stimuli, 1-4 for the MS stimuli) on the degree

of accentedness was analysed. Since perception was very similar across listeners from Normandy and the Paris region, listeners' ratings were pooled in Table 3.

Table 3: Degree of accentedness assigned to MO and MS stimuli as a function of the number of HL falls.

#falls	1	2	3	4
MO	2.47	2.67	3.10	–
MS	2.57	2.23	2.47	3.06

On the whole, the perceived degree of accentedness increases as a function of the number of HL falls, for both the original and the synthesised stimuli. Despite a particularly high rating for the three MS stimuli exhibiting one HL fall (primarily one of them), this further supports the determining role of these prosodic patterns. Recall the sentences were different and other features could influence the listeners' ratings.

4.2. Salient cues

The subjects' second task consisted of indicating particularly salient words or syllables in a text window. This field was filled in for a third of the stimuli, in a manner balanced according to the different stimulus types. That is, over a hundred comments were collected for the MO, US, UO and MS stimuli. Most of them pinpointed one or two words per stimulus. These words were classified according to whether or not they were target words, whose pitch contours were modified in the synthesised stimuli (see Table 4). The table reveals that most comments focus on the target words, even in the prosodically unmarked stimuli. These target words may bear other hints of a suburban accent. However, they are more often pointed out in the MO stimuli than in the US stimuli and in the MS stimuli than in the UO stimuli, while the number of salient words other than target words remains rather stable across the 4 stimulus types.

Table 4: Number of salient words pointed out by the listeners for the different stimulus types.

#salient words	MO	US	UO	MS
Target words	195	136	129	177
Other words	107	114	111	105

Listeners' final comments confirm the importance of prosody. Almost all of them refer to prosody, especially word-final accentuation.

5. DISCUSSION

The sociolinguistic literature pays a lot of attention to the social meaning of linguistic features (including prosodic features) found in the French working-class suburbs, but their phonetic grounding is often fragile.

Using an experimental methodology based on prosody modification and resynthesis, this study concentrated on the perception of an atypical prosodic pattern (a final HL fall). Results suggest that the latter is indexical of the suburban accent in French, for listeners from both Normandy and the Paris region. This HL fall may function as a marker of identity.

Specific prosodic patterns allow lower-class suburban youth to distinguish themselves from the mainstream society, as their look (loose tracksuit, cap, hood and basketball boots). These forms witness these adolescents' need to be recognised and integrated within a peer group. They are more and more associated with the ethno-geographic origin, especially from the Maghreb, in folk comments [4], even though perceptually a suburban accent may be related to Magrebian immigration regardless of the actual origin of the speaker [10]. Note that in the present experiment no speaker could speak Arabic. One should remain cautious about the role of Arabic (or the influence of rap music, which has also been put forth [2]) in what characterises this suburban accent. Language contact is an attractive interpretation, but it should be handled warily.

6. REFERENCES

- [1] Armstrong, N., Jamin, M. 2002. Le français des banlieues: uniformity and discontinuity in the French of the Hexagon. In Salhi, K. (ed.), *French in and out of France*. Hamburg: Peter Lang, 107-136.
- [2] Calvet, L.J. 1994. *Les voix de la ville. Introduction à la Sociolinguistique Urbaine*. Paris: Payot.
- [3] Fagyal, Z. 2003. The matter with the penultimate: prosodic change in the vernacular of lower-class immigrant youth in Paris. *Proc. 15th ICPhS Barcelona*, 671-674.
- [4] Fagyal, Z. 2004. Action des médias et interactions entre jeunes dans une banlieue ouvrière de Paris. Remarques sur l'innovation lexicale. *Cahier de Sociolinguistique* 9, 41-60.
- [5] Le Gac, D., Jamin, M., Lehka, I. 2006. A preliminary study of prosodic patterns in two varieties of suburban youth speech in France. *Proc. Speech Prosody Dresden*, 241-244.
- [6] Gadet, F. 2002. Français populaire: un concept douteux pour un objet évanescant. *Ville-École-Intégration Enjeux* 130, 40-50.
- [7] Jamin, M. 2005. *Sociolinguistic Variation in the Paris Suburbs*. Ph.D. thesis, University of Kent, Canterbury.
- [8] Lehka-Lemarchand, I. 2007. *Accent de Banlieue. Approche Phonétique et Sociolinguistique de la Prosodie des Jeunes d'une Banlieue Rouennaise*. Ph.D. thesis, Université de Rouen.
- [9] Praat. <http://www.praat.org>
- [10] Stewart, C., Fagyal, Z. 2005. Engueulade ou énumération? Attitudes envers quelques énoncés enregistrés dans "les banlieues". In Bertucci, M.-M., Houdart-Merot, V. (eds.), *Situations de Banlieue: Enseignement, Langues, Cultures*. Lyon: INRP, 241-252.