# Polish-accented French prosody in perception and production: transfer or universal acquisition process?

*Anna Kaglik[1], Philippe Boula de Mareüil[2]*

[1] UMR7023 Université Paris 8, Saint-Denis, France
[2] LIMSI-CNRS, Orsay, France

akaglik@univ-Paris8.fr, mareuil@limsi.fr

## Abstract

This study addresses the mastery of prosody in French second language (L2) from the twofold point of view of perception and production. Speech samples from Polish learners, late and early bilinguals were examined. Their native-likeness was assessed by experts in prosody on both the original signal and synthesis voices using prosody transplantation. Results suggest that the acquisition of the L2 prosody is not constrained by the age of first exposure to the L2 as is the acquisition of segments. Yet, intonational breaks may result in an impression of jerky utterances interpreted as non-native. This phenomenon could be attributed to a prosodic transfer, but a comparison with French and Polish monolinguals suggests that overemphasis observed in L2 is better explained by difficulties in managing the structural aspects of discourse organisation as long as the production process remains costly.

**Index Terms**: second language acquisition, mastery of L2 prosody, structural organisation, accent perception.

## 1. Introduction

It is generally accepted that an impression of foreign accent is caused by unusual realisations on both segmental and prosodic levels. A proper articulation of phonemic segments is not enough to erase this impression [12] [2]; prosody also plays an important role [6][4]. A pilot study [7] revealed that influential criteria to identify a foreign (Polish) accent in French spoken as a second language (L2) relate to the intonational organisation: an impression of jerky productions which this article will try and elucidate.

From an acquisition viewpoint, this work seeks to provide answers to the following questions. What is the contribution of prosody to the perception of a foreign accent (here a Polish accent in French)? Is it possible to achieve a native command of prosody in an L2, whatever the age of first exposure to the L2? Does this impression of a "jerky" non-native prosody result from a prosodic transfer from L1 to L2 or from a universal acquisition process?

The age of first exposure to the L2 is a factor which immediately comes to one's mind when dealing with foreign accents. It is the basis of the critical period hypothesis, which defines a temporal window beyond which the acquisition of a given linguistic behaviour is no longer possible

The L2 mastery may also be facilitated or complicated by the L1/L2 configuration [3][14][15]. In particular, languages may convey the information hierarchy in terms of topic/ comment throughout various formal correlates such as cleft sentences in French, determining the nature of the prosodic structuring. On the one hand, syntax in French seems to code what prosody codes in Polish (a language without determiners), especially in the introduction of the referent [8]: the informational function of prosody in Polish is rather assured by morphosyntactic means in French. On the other hand, studies in the acquisition field notice a general tendency towards overemphasis in the L2 prosody, while the speech production process remains cognitively costly [13]. To test the

validity of these hypotheses, recordings were collected among bilinguals in French and Polish, monolinguals of these two languages, as well as native Polish speakers who could be considered as learners of French, for comparisons. After a presentation of the corpus, the methodology and results of a perceptual experiment will be reported. A contrastive analysis of French and Polish prosodic phrasing will follow, with further details for learners and a few other Polish speakers of French.

## 2. Corpus

The corpus analysed in this study is composed of narratives in which participants told a movie scene after watching it: 35 native Polish speakers and 5 native French speakers were recorded. Among Polish speakers, 5 were monolinguals, and were only requested to speak Polish; 3 speakers could be considered as skilled learners and 27 as Polish-French bilinguals. In this article, we will only study their narratives in French. The bilinguals' length of residence in France was between 10 and 74 years (28 years on average), and their age of first exposure to French ranged from 3 to 32. Both bilinguals and French monolinguals had so-called prestigious occupations (lawyers, doctors, teachers, writers, etc.). The expositional parameters shared by the bilinguals were therefore optimal in terms of language access and above all communicational needs associated with social pressure, integration and professional motivation [9]. On the opposite, the group of learners, despite a prolonged stay in France (8 years on average), only had few daily contacts with Francophones: these speakers used Polish almost exclusively in both their personal and professional lives.

In addition to French natives and learners, the speakers were split into six categories according to their age of first exposure to French (<6, 6–10 years-old, etc.). The number of speakers, the mean age per group, the length of residence in France and the age of first exposure to French are given in Table 1 for Polish speakers of French.

Table 1: *Number of speakers in each group of Polish speakes; mean age, length of residence in France (LOR) and age of first exposure to French (in years).*

| Group | < 6 | 6–10 | 11–15 | 16–20 | 21–26 | 27–34 | Learners |
|---|---|---|---|---|---|---|---|
| #speakers | 5 | 4 | 6 | 4 | 3 | 5 | 3 |
| Age | 41 | 27 | 40 | 47 | 54 | 53 | 30 |
| LOR | 39 | 19 | 28 | 29 | 31 | 22 | 8 |
| First exposure | 4 | 7 | 12 | 18 | 23 | 31 | 20 |

The perceptual experiment described below is based upon speech samples (of 36 seconds on average) extracted from the French narratives of the 5 French monolinguals and the 30 Polish bilinguals or learners of French (20 males and 15 females).
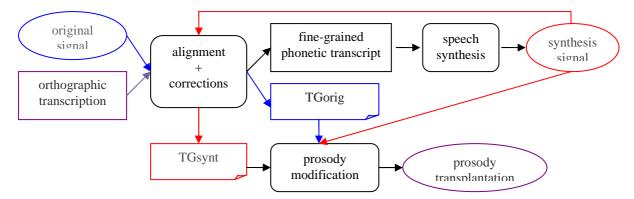
**Figure 1**: flow chart of the prosody transplantation procedure (TG refers to Praat TextGrids).

## 3.     Perceptual experiment

Several studies have pointed out the role of prosody in the perception of a foreign accent [6][4]. The authors of the latter study used two techniques, (1) based on text-to-speech synthesis relying on a prerecorded segmental base, and (2) modifying the prosody of the natural voice by the PSOLA speech processing algorithm implemented in the Praat software (www.praat.org). In the present work, these two approaches are combined by grafting the duration and fundamental frequency ($F_0$) parameters onto the male and female voices of the Acapela speech synthesis system (www.acapela-group.com).

### 3.1.     Experimental design

The 35 samples kept for the perceptual experiment were transcribed orthographically, from which a segmentation into phonemes was obtained by automatic alignment, using the EasyAlign system [5]. The resulting phonemic transcriptions were corrected manually (paying a particular attention to disfluency phenomena, schwas, liaisons and pauses), and given as input to the speech synthesis system. The synthesis output was also segmented by automatic alignment. Phoneme by phoneme, the duration and $F_0$ parameters of the originals were then transplanted onto the synthesis voices.

All in all, 8 specialists in prosody assessed the stimuli resulting from monolinguals, bilinguals and learners of French, presented in a different pseudo-random order for each subject. For both prosody transplantations and original stimuli, they had to estimate the native/non-native character of prosody on a 5-degree scale, from 0 (most likely non-native) to 4 (most likely native).

### 3.2   Results

The results obtained are presented in Table 2, averaged for the 8 groups of listeners defined in section 2.

Table 2: *Evaluation of the native-likeness for prosody transplantations and originals on a 0–4 scale.*

| Group | French natives | < 6 | 6–10 | 11–15 | 16–20 | 21–26 | 27–34 | Learners |
|---|---|---|---|---|---|---|---|---|
| Prosody | 2.5 | 2.4 | 2,1 | 2.0 | 2.5 | 2.1 | 2.0 | 1.1 |
| Originals | 3.3 | 3.5 | 3,6 | 1.6 | 1.5 | 1.7 | 2.7 | 0.1 |

Inspection of the prosody transplantation evaluations shows that monolinguals and bilinguals (whether early or late) have similar scores. Although there is a decline in naturalness

ratings of the prosody-transplanted speech from French natives to the next three groups, the bilinguals who began learning French between 16–20 years show a puzzling spike. Only the group of learners is clearly distinguished.

On the original stimuli (combining the segmental and prosodic levels), native speakers and early bilinguals (who began to learn French before the age of 10) are perceived in a rather similar way: they exhibit higher naturalness scores for original stimuli than for synthetic stimuli.  Late bilinguals (exposed to French after the age of 10) form a second group, whose behaviour is quite different from the first one. Learners form a third group, with a very low score (close to nil).

An analysis of variance (ANOVA) of the results on prosody transplantations reveals a significant effect of the group of speakers [$F(7,272) = 3.38$; $p < 0.01$]. However, pairwise *t*-tests only show a significant effect with the group of learners. Indeed, we have too few answers for each group of speakers, even though overall results suggest that the earlier speakers began to learn French, the more native-like their prosody sounds to the judges. On originals, the effect of the group of speakers is highly significant [$F(7,272) = 37.2$; $p < 0.001$]. However, the differences are not significant between early bilingual and native French speakers. These groups of speakers both achieve scores which are greater than 3 out of 4.

In the light of these results, we may claim that prosody plays a determining role in the perception of Polish-accented French in the sense that it enables proficient speakers to be distinguished from learners. In addition, the mastery of the L2 prosody is not an exception, even in late bilinguals. Whereas speakers exposed to French after the age of 10 are rather well distinguished from French natives in their original utterances, they may perform as native speakers regarding prosody. This outcome questions the critical period hypothesis assuming that one cannot be identified as a native speaker in an L2 if the learning of this language begins late in the lifespan, due to biological and maturational constraints affecting articulatory and prosodic aspects of language competence [17]. Our results suggest that a kind of window is shut at the segmental level but not at the prosodic level.

When speech sounds as foreign accented even at the prosodic level, language practice-related factors should be taken into account. Another possible explanation leads us to consider the typological differences between the two systems (here French and Polish), as put forth in a number of studies [14]. This is what we are going to develop in the remainder of this paper, insofar as prosody is not employed to fulfil the same functions in the two languages. The 6 speakers whose prosody is perceived as least native will also be analysed in more detail: the 3 learners and 3 bilinguals whose average degree of native-likeness was assessed at 1.2/4 on prosody

transplantations and 1.0/4 on the originals, while their age of first exposure to French ranged between 11 and 29.

# 4. Contrastive analysis of prosody: tentative explanation

Is it possible to evidence prosodic organisation differences between French and Polish? Before undertaking to answer this question, our analysis framework will be presented.

## 4.1. Conceptual tool: the intonational period

As previously for the experiment excerpts, narratives (both in French and Polish) were aligned and annotated under Praat with the aid of EasyAlign. The alignment output was then transferred to Analor (www.lattice.cnrs.fr/-Analor-). This software enables the automatic detection of so-called *intonational periods*, and has already been applied to French and other languages [1][11]. The segmentation into intonational groups was corrected manually as well as the coding of accentual prominences, according to a language-independent feature matrix corresponding to the following paradigms. **intonational gesture contour** (rising, falling, rising-falling, falling-rising or static), **range** (over-high, high, mid, low, infra-low), **syllabic lengthening**, **presence of a pause** and **hesitations**.

Stemming from [10], the intonational period is defined as an autonomous macrostructure unit beyond which there are no more intonational constraints. Corresponding to van Dijk's [18] *buffer*, this prosodic unit is cognitively constrained and constitutes an integration cycle in semantic memory: a thematic break or the weak coherence degree with old information triggers the start of a new period. The period break criteria (the **pause**, the **gesture and jump amplitudes**) are evaluated as follows:

- The **pause** (or more precisely the interval between two defined portions of $F_0$) is longer than 300 ms.
- The **gesture amplitude** (i.e. the difference between the last $F_0$ extremum and the mean $F_0$ value on the whole portion preceding the pause) is greater than 4 semitones (ST).
- The **jumb amplitude** (i.e. the difference between the last $F_0$ value preceding the pause and the first $F_0$ value following the pause) is greater than 3 ST.

Empirically validated values are associated to different criteria, as shown in Table 3. Their sum must be ≥ 1 to delimit a period, the presence of a pause being obligatory.

Table 3: *Values corresponding to different thresholds of pause, melodic gesture and jump parameters.*

| PAUSE | GESTURE | JUMP |
|---|---|---|
| < 250 ms = -1 | < 3 ST = -1 | < 2 ST = -1 |
| [250ms – 330 ms[ = 0 | [3 ST – 5 ST[ = 0 | [2 ST – 4 ST[ = 0 |
| [330ms – 660 ms[ = 1 | [5 ST – 8 ST[ = 1 | [4 ST – 7 ST[ = 1 |
| ≥ 660 ms = 2 | ≥ 8 ST = 2 | ≥ 7 ST = 2 |

These acoustic cues also served to evaluate the nature of the relationship between intonational groups (IGs) within periods. We annotated and ranked the parameters which punctuate each end of group-final accent by applying grouping/segmentation principles proposed by [10]. These principles allow us to establish a typology of intonational groups which takes into account the degree of prosodic salience with respect to the left and right contexts: an IG which constitutes a period (an *exo-period*) by itself is the most salient, throughout the combination of cues which characterise it as well as the left and right detachment; an IG breaking to the right with what follows (within the period) is also salient but to a lesser extent, while an IG linked to the left and to the right is considered as not salient prosodically.

The example below illustrates a segmentation into intonational periods (within parentheses) and groups (linked to the right "_" or, on the contrary, in a break "/").

- (euh il y avait donc une jeune **fi**lle / qui regardait dans une boutique _ apparemment une pâtisserie _ qui semblait avoir faim)
  'er there was thus a young woman / who was looking in a shop _ apparently a cake shop _ who seemed to be hungry'
- (qui a profité de ce que le livreur s'éloigne / pour euh _ voler un / une baguette)
  'who took advantage of the fact that the delivery man went away / to er _ steal an / a baguette'

As can be noticed, there are 8 IGs with 3 segmentation processes within only 2 periods. In comparison, in the following example in Polish, there are 7 IGs within 3 periods, with only 2 grouping processes:

- (Dziewczyna / ukradła bułkę _ z samochodu)
  'She stole / a bread _ from the car'
- (eee później / eee zaczęła / uciekać)
  'er then / er she began to flee)
- (z tą bułką)
  'with this bread'

## 4.2 Comparative results

On this basis, we analysed the whole of the narratives in order to determine what may account for this impression of "jerky" productions felt in some speakers. Results are given in the first lines of Table 4 for French and Polish monolinguals (from respectively 6 min 31 and 5 min 11 of speech which in total were analysed). They are reported per minute: number of syllables/minute for estimating speech rate, number of intonational periods (including exo-periods)/minute, number of IGs (including those which break to the right with the following IG)/minute. The speech rate is identical in the two languages, and the numbers of intonational periods per minute are comparable in both languages. Nevertheless, we notice that Polish resorts to exo-period strategies (i.e. periods made up of a single intonational group) more often than French does. Also, cross-language differences show up concerning intonational groups in a break with subsequent groups. Intonational groups breaking to the right only represent 33% of cases in French (13/40) as compared to 72% of cases in Polish (28/39). French and Polish thus differ as to the within-period organisation, the Polish language privileging segmentation processes whereas French rather employs grouping strategies.

Table 4: *Speech rate and prosodic structures in French L1 (for 5 monolinguals), Polish L1 (for 5 monolinguals) and French L2 (for 3 learners and the 3 bilinguals who are least judged as natives according to the perceptual test on prosody transplantations).*

| Number/min | Syllables | Periods | Exo-periods | IGs | IGs in a break |
|---|---|---|---|---|---|
| French L1 | 236 | 11.7 | 1.4 | 40 | 13 |
| Polish L1 | 236 | 13.3 | 4.8 | 39 | 28 |
| Non-natives | 234 | 13.2 | 6.4 | 36 | 29 |

What may be the implications for prosody acquisition? The same analysis was applied to the perceptual experiment samples of the 6 speakers who were least judged as natives on the basis of prosody transplantations: the 3 learners and 3 bilinguals (totaling 3 min 34 of speech). These two sets of speakers were grouped insofar as they present similar behaviours in terms of prosodic organisation. As shown by the bottom row of Table 4, their speech rate (in syllables/min) is a little bit slower than that of L1 speakers. More strikingly, their choices concerning the intonational period internal organisation are very far from what native French speakers do, that is, grouping. We observe that 81% of IGs (29/36) constitute groups segmented to the right. Additionally, in the majority of cases, these groups are marked up by the association of a falling contour and a syllable lengthening sometimes accompanied by a pause — a rare marking in French L1. They are slightly more frequent in learners (84%) than in the other speakers (77%). Moreover, the analysis of the speakers' narratives shows that intonational periods constitute relatively reduced-sized units from a syntactic-semantic viewpoint. Here is an example with, as above, periods within parentheses and intonational group breaks indicated by slashes:

- (une jeune fille)
  'a young woman'
- (apparemment pauvre)
  'apparently poor'
- (et affamée)
  'and starving'
- (voulait absolument euh / manger / et s'est arrêtée devant une vitrine / de boulangerie)
  'absolutely wanted er / to eat / and stopped in front of a baker's / shop window'

The oversegmentation observed could be attributed to processes typical of the Polish system. However, the segmentation rate is higher than for Polish monolingual speakers. Hence, prosodic transfer is not a sufficient explanation. The intonational period is cognitively constrained, especially by the production processing load. In accordance with Perdue and Gaonac'h's [13] hypothesis, we may state that the textual aspect of the message (in this case the prosodic aspect) is still beyond the control of the Polish speakers who pass least as French natives. Among these speakers, the reduced size of intonational periods, often limited to a single IG, reveals probable wording difficulties, impeding an overall utterance planning: the double focusing on both the content and the form still seems to represent too high a cognitive cost.

## 5.    Conclusion

A prosody transplantation-based technique was used to disentangle phonetic segments and prosody, in order to separate these two nested but autonomous levels. Polish-accented speech samples in French, stemming from speakers who started to learn French at the age of 4 to 31, were assessed and compared with speech samples from native speakers. Considering the age of L2 acquisition and in particular the prosodic competence learning, experimental results speak in disfavour of the critical period hypothesis as regards prosody. Concerning the very object of this study, prosody, our analysis brought to light intonational breaks which contrasted Polish speakers of French with French and Polish monolinguals.

According to this study, prosody acquisition does not seem to be constrained by the age of first exposure to the L2; nor does it seem to be out of the reach of late bilinguals. Cognitive processes used by adults in the late acquisition of an L2 (so-called "less specific" processes [16]) do not necessarily lead to a poor competence, just as early acquisition does not preclude the presence of a foreign accent in the L2. Nonetheless, we saw that non-native cues in the utterances of learners and even bilinguals may result in an overemphasis impression in French. This impression is not only due to the Polish-French L1-L2 pairing. According to our analysis, it is not so much a matter of prosodic transfer as an acquisition process-related phenomenon, attested regardless of source and target languages: language production difficulties in the management of prosodic aspects, still ill-controlled because of the planning and wording cognitive load [13]. Mastering rules governing prosodic groupings and detachments as well as the nature of the cues associated with these operations requires a very high level of competence, and seems to raise difficulties even for very advanced speakers. The too frequent segmentation leads to suppose that the planning unit is more local than macrostructural. In that sense, it relates to a universal acquisition process as long as speech production remains cognitively laborious.

## 6.    References

[1] Avanzi, M., Lacheret. A., Victorri, B. 2008. Analor, a tool for semi-automatic annotation of French prosodic structure. *Speech Prosody*, Campinas, 119–122.

[2] Birdsong, D. 2003. Authenticité de prononciation en français L2 chez des apprenants tardifs anglophones : analyses segmentales et globales. *AILE* 18, 16–36.

[3] Bongaerts, T. 2003. Effet de l'âge sur l'acquisition de la prononciation d'une langue seconde. *AILE* 18, 79–98.

[4] Boula de Mareüil, P. and Vieru-Dimulescu, B. 2006. The contribution of prosody to the perception of foreign accent. *Phonetica* 63, 247–267.

[5] Goldman, J.-P., Auchlin, A., Simon, A.C., Avanzi, M. 2007. *Phonostylographe* : un outil de description prosodique. Comparaison du style radiophonique et lu. *Nouveaux cahiers de linguistique française* 28, 219–237.

[6] Jilka, M. 2001. The contribution of intonation to the perception of foreign accent. PhD thesis, Universität Stuttgart.

[7] Kaglik, A. and Boula de Mareüil, P. 2009. Perception d'un accent étranger et part de la prosodie selon l'âge de première exposition à la L2 : transfert ou phénomène universel en acquisition ? *6es Journées d'Études Linguistiques*. Nantes, 7–13.

[8] Kaglik, A. 2009. La trace prosodique du mouvement référentiel et son acquisition en langues seconds (to appear).

[9] Klein, W. 1989. *Acquisition de langue étrangère*. Paris: Armand Colin.

[10] Lacheret-Dujour, A. 2003. *La prosodie des circonstants en français parlé*. Leuven/Paris: Peeters.

[11] Mettouchi, A., Lacheret-Dujour, A., Silber-Varod, V., Izre'el, S. 2007. Only Prosody? Perception of speech segmentation in Kabyle and Hebrew. *Colloque International Interface discours Prosodie*. Geneva, 207–218.

[12] Missaglia, F.: 1999. Contrastive prosody in SLA. An empirical study with adult Italian learners of German. *ICPhS*, San Francisco, 551–554.

[13] Perdue, C. and Gaonac'h, D. 2000. Acquisition des langues secondes. In *Acquisition du langage*, M. Kail, & M. Fayol (eds), Paris: Presses Universitaires de France.

[14] Rasier, L. and Hiligsmann, P. 2006. La focalisation en français, en néerlandais et dans l'interlangue d'apprenants d'une langue étrangère. In *La focalisation dans les langues,* H. Wlodarczyk & A. Wlodarczyk (eds). Paris: L'Harmattan, 173–186.

[15] Rasier, L. and Hiligsmann, P. 2007. Prosodic transfer from L1 to L2. Theorical and methodogical issues. *Nouveaux cahiers de linguistique française* 28, 41–66.

[16] Schachter, J. 1988. Second language acquisition and its relation-ship to Universal Grammar. *Applied Linguistics* 9, 219–235.

[17] Scovel, T. 1988. *A time to speak. A psycholinguistic inquiry into critical period for hman speech*. Rowley: Newbury House.

[18] van Dijk, T.A. 1985. *Handbook of Discourse Analysis*. London: Academic Press.