ProSeg: A comparative corpus of spoken L2 French
Elisabeth Delais-Roussarie, Tanja Kupisch, Paolo Mairano, Fabian Santiago, Frida Splendido

To cite this version:

HAL Id: halshs-02109337
https://halshs.archives-ouvertes.fr/halshs-02109337
Submitted on 24 Apr 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
1. Introduction

Advantages of speech corpora in L2 acquisition studies [2]
- They represent the oral dimension in L2 in an ecological way (i.e. with spontaneous speech);
- They put in perspective several explanatory factors affecting the L2 acquisition process and competence, such as L1 transfer, speech style (reading vs. spontaneous speech), age of acquisition. [5, 6]
- They can be used for cross-comparisons between language pairs, especially when using the same recording protocol [6]
- Given these advantages, many current studies are based on speech corpora. But different procedures may be used for data collection, each having its advantages [3]
  - Experimental data targeting specific linguistic phenomena;
  - Naturalistic speech from an oral corpus. ProSeg (Prosody Segments) combines the two procedures.

2. Main features of ProSeg

General features
The ProSeg Corpus has been designed to allow for :
- prosodic and segmental descriptions of L2 French
- comparison of potential effects of L1 transfer
- contrastive analyses of the oral productions in L1 and L2 with a comparable set of data
- analyses of task effects

Procedure
1. Background questionnaire
2. Placement test (DIALANG, yes-no vocabulary task, [1])
3. Reading tasks in L2 French (several texts including a dialogue)
4. Picture description task
5. Free speech (summary: book/movie)
6. Reading task in the L1

Participants
- 65 participants (so far)
- Recorded both in L2 French and in their L1
- All learners are university students

<table>
<thead>
<tr>
<th>L1</th>
<th>University</th>
<th>N</th>
<th>Age (SD)</th>
<th>Proficiency Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Italian</td>
<td>Turin</td>
<td>25</td>
<td>25.2 (SD = 3.7)</td>
<td>B1-C1</td>
</tr>
<tr>
<td>German</td>
<td>Konstanz</td>
<td>21</td>
<td>24.6 (SD = 6.5)</td>
<td>B1-C2</td>
</tr>
<tr>
<td>Swedish</td>
<td>Lund</td>
<td>5</td>
<td>25 (SD = 5)</td>
<td>B2-C1</td>
</tr>
<tr>
<td>French (controls)</td>
<td>Paris 8</td>
<td>14</td>
<td>26 (SD = 5)</td>
<td>Native</td>
</tr>
</tbody>
</table>


(Non-native) gemination in L2 French
- Analysis of <CC> vs. <C> spelling, e.g., inmigrés, (immigrated) vs. imitére (imitate)
- Results: Italians produce geminate consonants in L2 French due to L1 transfer and, specifically, of L1 orthography [4]

Foreign Accent Rating
- Two sentences extracted from read texts with no self-repairs or hesitations
- Judgments by native and non-native speakers of French (including teachers of French) in terms of
  - (i) accent strength
  - (ii) accent source (L1)

Prosodic phrasing and rhythm
- Prosodic analysis of the segmentation, durational and tonal patterns in accentual phrases (APs)
- Data taken from the long text and the monologues. Comparison of L1 German, L1 Swedish, L1 Italian and L1 French speakers.
- Hypothesis: L1 Swedish learners have advantages in producing e APs in L2 French due to the existence of a tonal accent in Swedish.