Impact of Vowel Reduction in L2 Chinese learners of Portuguese

Catarina Realinho1, Rita Gonçalves1, Helena Moniz1,2, Isabel Trancoso2,3

1FLUL/CLUL, Universidade de Lisboa, Lisboa, Portugal
2Human Language Technologies, INESC-ID Lisboa, Lisboa, Portugal
3Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal

Abstract

Connected speech processes have a great impact on word recognition. This is particularly impactful for L2 learners. Vowel reduction is a very productive process in European Portuguese and plays an important role within and across word boundaries. Our goal is to understand the influence of these phonetic-phonological phenomena in L2 Chinese learners of European Portuguese in word identification tasks, in continuous speech, in a classroom environment. We designed a perception experiment involving these phenomena in increasing degrees of difficulty: isolated word identification without (i) and with vowel reduction (ii); word identification with simple (iii) and complex connected speech processes (iv). This study took place in a classroom setting and it was applied to a group of B1 students of European Portuguese and a group of native speakers (control group). Data from L2 oral productions were also collected to compare it with perception data. The rate of correct answers for each task matched our expectations: (i) 94%; (ii) 65%; (iii) 31%; (iv) 16%. The results reveal that word recognition is challenging for L2 learners due to the connected speech processes. Vowel reduction brings difficulties for learners, even when the word is isolated. Production analysis reveals that learners do not produce vowel reduction, nor any other phonetic-phonological phenomena across word boundaries.

Index Terms: Second Language; Phonology; Perception; Production; Vowel Reduction; Connected Speech Processes

1. Introduction

Vowel reduction within and across word boundaries is usually described as a very frequent process in European Portuguese (EP). In continuous speech in informal contexts, it can be so extreme that may even be perceived by a non-native European Portuguese speaker as a Slavic language, due to consonant clusters. L2 perception studies usually use lab-based input. However, in (semi)spontaneous speech, sounds are altered, deleted, contracted, and this may affect L2 speech perception. The interplay between vowel reduction and connected speech processes in both perception and production tasks is not studied in a context of L2 learners of EP. To this end, in this paper we describe the impact of these phonetic-phonological phenomena in L2 Chinese learners of EP. The experiments were conducted in an ecological setting of an intensive Portuguese course, intermediate level (B1), at the University of Lisbon. A control group of native speakers also performed the experiments. The main goals of this study are threefold: (i) to describe the impact of vowel reduction and continuous speech processes in perception tasks; (ii) to analyze the presence/absence of vowel reduction phenomena in production tasks; (iii) to understand the interplay between the complexity of the linguistic structure and vowel reduction patterns, at the intermediate level (B1). We started by designing a series of perception experiments with increasing degrees of difficulty: from isolated word identification, without and with vowel reduction, to word identification in continuous speech with simple and complex contexts across word boundaries. The production data from students was also analyzed, to map the presence/absence of vowel reduction or/and other phonetic-phonological phenomena with the oral competences pertaining to that intermediate level. This paper is structured as follows: Section 2 summarizes the related work; Section 3 presents the experimental design of our perception and production tasks; Section 4 analyzes the perception results, whereas Section 5 deals with the production data. Finally, Section 6 presents our conclusions and future work.

2. Related work

Listening in second language is a complex process, as extensive literature has been pointing out. Efficient L2 listening is described with similar phases to native listening: segmenting continuous speech, distinguishing phoneme contrasts, activating words from vocabulary and dealing with other phonetic-phonological phenomena to achieve a successful listening[1]. An L2 learner is biased by his/her L1, which acts as filter conditioning both perception and production. Speakers use language-specific strategies in speech segmentation and word recognition [2], [3]. Thus, problems arise when both L1 and L2 are phonologically distant [4], [5], [6]. The vocabulary of an L2 learner is much smaller than his/her L1 vocabulary. Thus, it may trigger issues in L2 vocabulary activation and may even cause a competition between words and “non-words” activation [7]. Furthermore, the form of a spoken word may be altered due to connected speech processes [8], [9]. As a consequence, learners may recognize spoken isolated words but may not recognize that same words in continuous speech [2]. Thus, segmenting continuous speech requires knowledge of word-boundary patterns and distinct phonological inventories may trigger a wrong segmentation [10], [11], [12]. Moreover, meaning is a strong cue in word recognition. Research on this topic also revealed prosody plays an important role in detecting L2 syntactic boundaries [13].

In our research, the two languages are very distinct in terms of phonological rules, [14], [15], [16]. In European Portuguese, stress position may vary within the word [17] triggering lexical contrasts [18]. In addition, vowel reduction within and across word boundaries [17] and connected speech processes [19] are very productive in EP. The interplay of these phonological-phonetic phenomena results in the production of multiple consonant clusters [16] which has a great impact in the EP rhythm. In contrast, Mandarin Chinese (MC) has lexical tones [20] and a simple syllabic structure [14], [15] allowing the occurrence of only one segment at the onset position. The resulting
absence of consonant clusters contributes to a syllable-timed rhythmic structure [21]. Thus, Mandarin speakers may apply these patterns when listening to European Portuguese utterances [1] or when producing Portuguese utterances [13], which may be perceived and assessed as a non-fluent strategy.

Portuguese research on this topic [22] is related to spoken interlanguage rhythm of L2 Chinese learners of EP regarding an intermediate (B1) and an advanced level (C2). In the reading tasks, B1 students produced more vowel segments, contrasting with more advanced speakers (C2) who lowered the vowel production, resembling a fluent native Portuguese speaker, but still not in the same proportion of a native [16], [17]. This proves that at intermediate or advanced levels, speakers still transfer properties from their L1 to the L2 [13], [1].

A final note on speech data and its ecological usages. Second language teaching should provide sufficient training in order to improve students’ perception, production, and comprehensibility in the target language [13], [23], [24]. L2 teaching studies usually focus on learners’ errors, describing what kind of errors they produce [13]. If teachers have a clear understanding of the common difficulties that students experience [3], it will contribute to develop or adapt teaching strategies and didactic materials [2]. For instance, teaching students of how some words and sounds change in spontaneous speech may encourage their ability to segment continuous speech [8]. Thus, second language teaching must assure students are exposed to authentic speech materials and real communicative contexts [9], [23], [25], [1]. It is, therefore, important to include listening exercises with real interviews, for instance, and perform classroom activities or games to encourage oral interactions between students and teacher [24]. This was the bases o four work and the main distinction between what was scarcely been produced for L2 EP learning. We tackle read and spontaneous speech in an ecological classroom environment with exercises mimicking part of what the teacher would conduct with the class.

3. Experimental scenarios

3.1. Participants

To understand the acquisition trajectory of the phonetic processes, we selected a group of students at the intermediate level (B1) of EP (according to the Common European Framework of Reference for Languages (CEFR)). B1 students can understand the general message and the main points of clearly articulated standard speech, as in TV programmes, discussions, lectures, etc., provided the topics are familiar to them.

The class consists of Mandarin Chinese native speakers (N=12), aged between 20 and 25. Only a student also spoke Cantonese. Previously to the course, the students learned Portuguese in China and then took an intensive Portuguese course at the University of Lisbon. They were in Portugal approximately since the beginning of the course. In addition, a group of EP native speakers (N=12), aged between 23 and 27, also performed the experiments in order to set a baseline. They were either undergraduate or postgraduate students at the University of Lisbon.

3.2. Perceptual experiments

The time frame and the order of the experiments were planned so as to study the impact of vowel reduction in perception tasks, and simultaneously reinforce the didactic outcomes of the results. The experiments were integrated into the topics taught in the course, so that the experiments could be seen as a natural extension of the classroom dynamics.

3.2.1. Single word identification tasks

The first experiment consisted of two single words identification tasks. The first task encompassed 10 prosodic (stressed) words, all nouns, with an extension of 4 syllables on average (see Table 1, for examples). To ensure students knew the words, they were all selected from exercises performed within the same teaching unit, in their classes. The stimuli were previously recorded by a female European Portuguese native speaker, in a silent room, using the recording system from a Samsung Galaxy Note II smartphone. Recordings were converted to WAV format and edited in Praat software[26]. All words were produced without using vowel reduction. Students heard the stimuli of each isolated word, and wrote it down as soon as they identified it. Each word was presented twice. The task took around 10 minutes to be completed. The experiment took place in approximately a month after the beginning of the course.

The second task, similarly to the previous one, was also single word identification, in which students were exposed to a sequence of 10 isolated words, but this time pronounced with vowel reduction. This task was performed in the same day, and took 10 minutes as well. The vowel reduction patterns closely followed the frequent ones described for EP, as established in [17], such as the deletion of [i], [u] and [i] in unstressed positions. Again, all words were selected from previous exercises (see Table 2 for examples). Recording and edition processes were replicated.

3.2.2. Cross word identification tasks with vowel reduction and other phonetic-phonological phenomena

The second experiment was more challenging than the first one, by means of combining vowel reduction with other very productive continuous speech processes in EP. For this reason, it was essential to use authentic speech material presenting real communicative contexts. This experiment (8 minutes) replicates common exercises students performed in the classroom. For the first task, we selected a YouTube news video, 01:29 minutes long, related to the topics the students were learning. Only the audio of the video was presented (twice). The students were given a form with an incomplete orthographic transcription, from which 9 word pairs were missing (totaling 18 different words). The selected missing words that the students had to fill in were prosodic words from varied lexical categories, with 3 syllables on average. The words were familiar to the students,
<table>
<thead>
<tr>
<th>Word</th>
<th>Translation</th>
<th>Transcription</th>
</tr>
</thead>
<tbody>
<tr>
<td>nove anos</td>
<td>nine years</td>
<td>['nOv'5nuS]</td>
</tr>
<tr>
<td>já estava</td>
<td>was already</td>
<td>['3af'tave]</td>
</tr>
<tr>
<td>jovens saem</td>
<td>youngster leave</td>
<td>['govjJf'sajJ3J]</td>
</tr>
<tr>
<td>países preferidos</td>
<td>preferred countries</td>
<td>['prEiz'pur'ridJ]</td>
</tr>
<tr>
<td>dois mil</td>
<td>two thousand</td>
<td>['dOj3'miõ]</td>
</tr>
<tr>
<td>depois vêm</td>
<td>then come</td>
<td>['dpoj3'viJ3J]</td>
</tr>
<tr>
<td>melhores oportunidades</td>
<td>country where</td>
<td>[mi 'azorpZptuni'daJfJ]</td>
</tr>
<tr>
<td>país onde</td>
<td>country where</td>
<td>['priz õdl]</td>
</tr>
</tbody>
</table>

Table 3: Examples of word sequences in Task 3.

<table>
<thead>
<tr>
<th>Word</th>
<th>Translation</th>
<th>Transcription</th>
</tr>
</thead>
<tbody>
<tr>
<td>vinte à hora</td>
<td>twenty per hour</td>
<td>['vitu'are]</td>
</tr>
<tr>
<td>uma antiga colega</td>
<td>an old colleague</td>
<td>['umã'tig5k 'lerge]</td>
</tr>
<tr>
<td>sabe que o</td>
<td>knows that the</td>
<td>['sabkJu]</td>
</tr>
<tr>
<td>penso que a</td>
<td>I think that the</td>
<td>['pe5skje]</td>
</tr>
<tr>
<td>resultado do encontro</td>
<td>meeting result</td>
<td>[ruZ'adwe'vã'oF'ra]</td>
</tr>
<tr>
<td>sede de trabalhar</td>
<td>thirst to work</td>
<td>['sodtrebe'ar]</td>
</tr>
<tr>
<td>tentar a sorte</td>
<td>try your luck</td>
<td>[tê'tarǝ'sorǝt]</td>
</tr>
<tr>
<td>mil euros</td>
<td>a thousand euros</td>
<td>['mil'eurofJ]</td>
</tr>
</tbody>
</table>

Table 4: Examples of word sequences in Task 4.

since they were also taken from exercises concluded in previous classes. Table 3 illustrates the 4 categories of our target processes: vowel encounters[19] (top row) and the production of /l/ as [j], [õ] or [z] due to voicing assimilation [17] (second, third and fourth rows, respectively).

A second task performed in the same day was similar to the previous one, but targeted more complex connected processes. In this case, we tested haplologies (deletion of an equal or similar syllable as in "Cam[po] Pequeno"), occurrence of [J] and re-syllabification with liquids [r] and [õ]. The selected news video, of 02:36 minutes, included real interviews which means students were exposed to casual speech [1]. We selected 11 target word sequences, totaling 29 different words. In this task, we tested stressed and unstressed words, ranging between 1 to 6 syllables. Table 4 illustrates examples of the 4 selected categories [19], [16]: vowel encounters (row 1), occurrence of [J] (row 2), haplologies (row 3), and resyllabification with liquids [r] and [õ] (row 4). The experiment lasted around 10 minutes and was implemented two weeks after the previous one. The control group performed the same sequence of tasks in different moments. The experiments were applied in silent rooms such, as libraries and study or meeting rooms.

3.3. Production Data

Students’ speech productions were collected from their final oral exam. In this test, students had to express their opinion about a picture or a topic taught in class. This exercise lasted 12 minutes. We analysed a small sample, in order to detect any occurrence of vowel reduction or any other sandhi phenomena. Our aim in analyzing spontaneous speech data was to further understand if the students would acquire and produce vowel reduction or connected speech processes at the end of the course. The data is still undergoing a more thorough process of analysis,

<table>
<thead>
<tr>
<th>Tasks</th>
<th>Correct Answers</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Single word w/out reduction</td>
<td>94%</td>
<td>9,4</td>
<td>0,8</td>
</tr>
<tr>
<td>2 Single word with reduction</td>
<td>65%</td>
<td>6,5</td>
<td>1,2</td>
</tr>
<tr>
<td>3 Simple connected speech</td>
<td>31%</td>
<td>2,5</td>
<td>1,7</td>
</tr>
<tr>
<td>4 Complex connected speech</td>
<td>16%</td>
<td>1,6</td>
<td>1,0</td>
</tr>
</tbody>
</table>

which could not be fully performed for this paper.

4. Results

4.1. Results of the perception experiments

Students’ answers revealed they use several orthographic forms to transcribe the same word. We considered an answer correct when the word was orthographically correct or when the grapheme could correspond to the phonetic transcription of the intended grapheme, for instance: alternations between the grapheme “í” and “é” both corresponding to the sound [i] [23] and changes in the order of the segments to simplify consonant clusters [22]. The following tables show the percentage of correct answers calculated for each experiment for Mandarin speakers (Table 5) and the control group (Table6).

As shown in the Table 5, experiments with isolated word identification tasks had higher rates (94% and 65%) than experiments involving word identification tasks in continuous speech (31% and 16%), as expected [1], [8]. Task 1 revealed that students identified words correctly in a satisfactory way (94%). This means that the occurrence of all the phonetic segments was useful to listeners to recognize and identify most words. In task 2, the percentage decreased to 65%. Vowel reduction [17] is not a common process in Mandarin [14], [15]. So, this decrease revealed that vowel reduction within a word has a great impact on word identification process at this level.

The gap between tasks is even more evident in task 3 with only 31% of correct answers, showing that word identification is affected when vowel reduction, voicing assimilation and vowel encounters (common EP connected speech processes) occurred across word boundaries. Even though students knew the target words they did not recognize them in continuous speech [2].Thus, students might be using cues to segment speech that result in their L1 but not in EP [3], [8].

In task 4, we tested vowel encounters plus haplologies, insertion of [j] and re-syllabification with liquids [r] and [õ] which means students dealt with stressed and unstressed words and with the increase of the number of target words. The rate of correct answers was 16%. This decrease shows that the occurrence of complex connected speech processes plus vowel reduction, always present within words and across word boundaries, heavily compromised word identification. In addition, students
were exposed to casual speech which is less careful [1] than the speech in a news report, as in task 3. As shown in the Table 6, control group had 100% for tasks 1 and 2. This highlights the fact natives do not have problems recognizing isolated words with or without vowel reduction. In addition, native participants rated 93% for task 3 and 98% for task 4, showing they did not have greater difficulties processing connected speech, in comparison with the test group. Nevertheless, these results reveal that some parsing problems may arise, mostly due to listening and writing at the same type.

4.2. Results of the production data

In order to understand the impact of all these phenomena in students’ acquisition trajectory it was interesting not only to understand how word identification is affected, but also to know if students could produce vowel reduction and other phonetic processes in spontaneous speech.

Results show that students do not produce vowel reduction or any other sandhi process. Figure 1 and Figure 2 illustrate examples showing that vowel reduction is missing, for instance, the deletion of [u] (“concord[u]” in Figure 1) and [i] (“apri[s]entar” in Figure 2) in unstressed positions. In addition, there is no vowel reduction in vowel encounters, for instance, the deletion of the final unstressed sound and the consequent articulation process with the first sound of the following word (Est[õ]pinhão, in Figure 1 and pod[õ]resentar in Figure 2), as expected in fluent EP [19]. Our results with spontaneous speech are consistent with the results described in [22], for reading tasks. However, in our study listening and segmenting words is heavily impacted by the presence of vowel reduction, the speech rate, and even the degree of vowel reduction and other sandhi phenomena. In a classroom context, the students of an intermediate level are not able to fully understand or produce several sandhi processes, mostly vowel reduction ones. This task is even more constrained due to spontaneous speech and didactic materials extracted from real contexts, an usual practice in classroom environments. It is interesting, however, to notice that some students apply re-syllabification with [y] in fixed expressions such as “por[y] exemplo” (“for instance”) or “por[y] isso” (so that”). This means students learned that specific structure as a fixed lexical expression and they do not apply the process in other contexts. In the production data it is also worth mentioning the very distinctive prosodic and intonational patterns. The students, even the highest scored student, utter every stress word as a prosodic unit, with frequent pauses and stressing words which are unstressed, as if each word received a tone [20].

5. Conclusions

In this paper we investigated the impact of vowel reduction and other phonetic-phonological phenomena in Chinese L2 learners of Portuguese at the intermediate level (B1). The aim of this study was to understand the influence of these phenomena in word identification in continuous speech. From a didactic point of view, our study also aims to understand whether the acquisition of vowel reduction and connected speech processes takes place and how listening skills can be trained. The majority of EP L2 courses do not address these phonetic processes. Thus, we applied our study in a classroom setting with 12 native speakers of Mandarin and 12 EP native speakers, our control group. We designed two experiments concerning four word identification tasks involving these phenomena in increasing degrees of difficulty. Results show that the rate of correct answers for each task (in percentage) matched our initial expectations: decreasing from 94% to 65%, when single words were produced with vowel reduction, and substantially decreasing across word boundaries (31% and 16%). In contrast, the group of native speakers had higher rates: 100% in single word identification tasks similar rates in tasks testing connected speech (93% for task 3 and 98% for task 4). The results highlight that word recognition can be a challenge for L2 learners, due to the occurrence of vowel reduction (even within a word) and connected speech processes [8], [9], [10], [2]. Students are transferring properties from their L1 to their L2 and this is particularly true for lower proficiency levels [1], [22].

Production analysis reveals that learners do not produce vowel reduction nor any other phonetic-phonological phenomena across word boundaries, except within fixed expressions such as “por exemplo” (for instance), in which they apply re-syllabification with [y]. The phonetic-phonological component in second language teaching should include some training concerning segmental and suprasegmental processes using authentic speech materials [9], [25]. The exposure to all these processes will improve their listening skills [8], [2]. The didactic outcomes of the experiments may contribute to the development of teaching materials for EP as L2, focusing on these processes.

Future work will tackle the comparison with other levels, in order to understand the degree of fluency when considering vowel reduction and other sandhi phenomena, always in a classroom environment and using the continuum from read-to spontaneous speech.

6. Acknowledgements

This work was supported by national funds in Portugal through Fundação para a Ciência e a Tecnologia (FCT) with reference UIDB/50021/2020.
7. References


