

# Exploring the Phonetic Context Effects on the Production of Portuguese /r/ by L1-Mandarin Learners

Yidan Cao, Chao Zhou & Gueorgui Hristovsky

Center of Linguistics, School of Arts and Humanities, University of Lisbon  
 madalena970603@163.com, zhouchao@edu.ulisboa.pt, ghristovsky@campus.ul.pt

## ABSTRACT

In this study, we examined whether L1-Mandarin learners' production of European Portuguese tap in onset cluster /CrV/ is conditioned by the preceding consonant and the following vowel. 14 native Mandarin speakers first learned to associate 12 Portuguese pseudo-words with /CrV/ cluster to 12 pictures and then performed a picture-naming task. L2 production data were judged by 25 Portuguese native raters and analysed with a logistic mixed-effects regression.

Results indicate that the preceding stop affects L2 production of /r/ as a function of place of articulation (/tr/ > /kr/ > /pr/). We argue that this may be accounted for by different degrees of articulatory effort, rather than by phonological markedness or lexical frequency. An effect of the following vowel was also found (/i/ facilitates the rhotic production more than /a/ does), which might also be articulatory-based. The implication of these results on the theoretical development of L2 speech research is discussed.

**Keywords:** Phonetic context, Onset cluster, European Portuguese, Mandarin Chinese, L2 speech.

## 1. INTRODUCTION

The European Portuguese (EP) tap /r/ is notoriously difficult for L1-Mandarin learners. It is very often confused with an L1 sound category, /l/, both in speech perception and production (see [32] for an overview). Recent empirical studies even show that the complexity in acquiring this novel sound goes beyond segmental confusability ([29, 31, 32]). For instance, Zhou [31] observed that the EP tap of an onset cluster was produced by L1-Mandarin learners more accurately when preceded by an alveolar stop (e.g. /dr/ or /tr/) than by a bilabial one (e.g. /pr/). Zhou [31] reasoned that this result may stem from the fact that the most prevalent repair strategy employed by L1-Mandarin learners (/r/ → [l]) is blocked when the preceding consonant is an alveolar stop. This is because such replacement would give rise to a cross-linguistically marked structure ([dl] or [tl]) (e.g. [11, 12]). Examining L2 perception of EP /r/ by Mandarin-speaking learners, Vale [29] found that, in

intervocalic onset position, the adjacent vowel /a/ helps L1-Mandarin participants identify /r/ (/ara/), compared to /i/ (e.g. /iri/). Zhou [32] further reported a similar facilitating effect of adjacent vowel /a/ over /i/ in a perceptual discrimination task, where L1-Mandarin learners were asked to discern the difference between /r/ and its confusable category /l/. According to Zhou [32], this vowel effect might be due to the fact that [a] leads to more salient F2 transition slopes than [i] does, thus empathizing the acoustic difference between [l]-[r]. Taken together, these results suggest that L2 acquisition of EP /r/ does not proceed equally in all contexts, i.e. it is subject to its surrounding phonetic contexts.

In spontaneous speech, the perception [26] and articulation [20] of a given speech sound are considerably shaped by its adjacent context. Since an L2 sound is normally acquired within a word or larger prosodic unit, it is not surprising that contextual effects on L2 segmental learning have long been documented (e.g. [16, 17, 22]). Prior research on such effect either targets the preceding or the following context, while, in the case of EP tap in onset cluster (/CrV/), the acquisition of /r/ may be prone to contextual influences from both directions. In this study, expanding on previous studies, we aim to assess the contextual effects on the L2 acquisition of EP /r/ by L1-Mandarin learners, by exploring both the role of the preceding consonant and also that of the following vowel.

## 2. PHONETICS AND PHONOLOGY OF EP /r/

The EP /r/ is most frequently realised as a tap. Depending on the adjacent segment and syllable position, other phonetic variants have also been reported [25]. In coda position, when preceding a stop (e.g., *carta*, “letter”), it is very often produced with an epenthetic vowel, while a realization as fricative is more common, if followed by another fricative (e.g., *curva*, “curve”). In word-final position, /r/ is often produced as a voiceless fricative [14] and can even be omitted, especially when the following word starts with a consonant (e.g. [19, 21]).

Regarding its distribution, /r/ may occupy all syllabic and word-level positions, except word-initially [18]. It may form an onset cluster with both

stops (/b, p, d, t, g, k/) and labial-dental fricatives (/f, v/), although clusters /C<sub>fricative</sub>t/ have a more restricted distribution [18]. The vocalic element of /CrV/ cluster can be any of EP phonological vowels /a, ε, e, i, o, ə, u/.

### 3. CURRENT STUDY

In order to explore whether and how the preceding and following contexts influence the L2 acquisition of EP /t/, we assessed the production of /CrV/ cluster by L1-Mandarin learners via a picture-naming task. We deem this task more suitable for the present study than other production tasks, because picture-naming does not require the activation of orthographic representations (as in the case of a reading task) and it avoids misperception effect (as in the case of an imitation task), for instance. Both orthographic and perceptual influences have been previously reported on L1-Mandarin learners' acquisition of EP /t/ [32].

On the basis of Zhou [31]'s results, we predicted that a preceding alveolar stop will promote the production of /t/, namely /t/ > /pt, kt/. Following Vale [29] and Zhou [32], we further expected to observe that the production accuracy of /t/ is higher when followed by /a/ than by /i/.

#### 3.1. Participants

Fifteen female Mandarin native speakers (mean age: 20 years old), all second-year Portuguese majoring students at the XXXX University, were recruited and tested in China. The production data of one participant was excluded from the analysis, because her L1 Mongolian has a trill /r/ in its inventory [27], which may promote the acquisition of EP /t/. None of the participants has fluency in or regular use of another language than English (college-level). None has reported hearing or speech impairment.

#### 3.2. Stimuli

Twelve dissyllabic pseudowords were created with onset cluster /CrV/ in penultimate stressed syllable. For assessing the role of the preceding consonant, 9 items contain clusters starting with stops varying in place of articulation (labial, alveolar and velar), namely 3 with /pra/, 3 with /tra/ and 3 with /kra/. Only voiceless initial stops were chosen as the voicing feature is quite problematic for L1-Mandarin learners (e.g. [23, 30]). In order to explore the effect of the neighboring vowel, another 3 items with onset cluster /pri/ were added to contrast with those with /pra/. Here, the bilabial stop /p/ was chosen as the pre-rhotic consonant to minimize the carry-over coarticulation effect between the initial consonant and /t/. Finally, 3 items with onset cluster /pla/ were included as fillers.

Fifteen coloured pictures representing different kinds of objects (e.g. 3 different types of knives; 3 different types of boats) were selected and were randomly associated with 15 test items with /CrV/ cluster. By doing so, we assume that L2 learners can hardly know the name of all subcategories of a given object, so that it is possible to adopt novel word learning paradigm. See the supplementary material for a complete list of pictures and items ([osf.io/4fuec/?view\\_only=82fdd9ffaf8d4f39abf9b2aa5ab33e](https://osf.io/4fuec/?view_only=82fdd9ffaf8d4f39abf9b2aa5ab33e)).

#### 3.3. Task and procedure

The entire experimental session consists of two parts. The first is a novel word learning phase, during which the participants were given the full list of test items with associated pictures. To guarantee the lexicalisation of these items, they were given a week to memorize them and were told that they can review the list as many times as they want before the testing phase. This time interval was defined, based on the necessary time for lexical consolidation reported in the literature (e.g. [7, 13]).

During the testing phase, each participant performed a picture naming task, administrated in Microsoft PowerPoint. The testing material contains a total of 45 slides, 3 repetitions of each stimulus. The presentation order was totally randomised. The task on average took about 5 minutes. Recordings were made in a quiet room in the library of Beijing Jiaotong University to a Tascam DR100mkIII recorder and were digitized at an audio sampling rate of 44.1 kHz. All recorded sound files were adjusted to the average intensity of 70 dB in Praat [3].

#### 3.4. Data preparation and analysis

A total of 630 tokens (3 items × 5 conditions × 14 speakers × 3 repetitions) were collected, among which 25 tokens were excluded (3.96%) from analysis, due to segmental changes either in first syllable (target form: /'traʃo/; learner form: ['praʃo]) or the second syllable (target form: /'pɹita/; learner form: ['pɹifi]). The occurrence of segmental replacement led us to consider that the learners have not lexicalised these items.

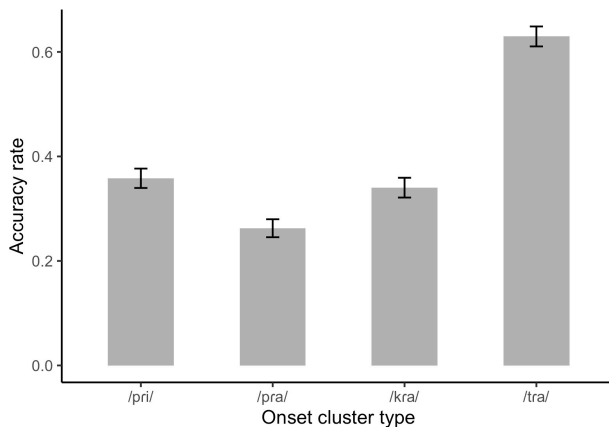
In order to determine whether the tap was produced accurately by L1-Mandarin participants, 25 native speakers of EP were recruited and were asked to identify the target segment in each production. The remaining 605 tokens were divided into 5 subtests. Each subtest contains the production by 3 informants and was judged by 5 native raters. During the judgment task, native raters listened to one auditory token each time and was asked to choose from three available response options representing the target tap

("r"), the most frequent replacement ("l") and other production deviants ("neither r nor l"). The learners' production of /r/ was regarded as correct, if a rater's response for that production is "r".

A practice session with five test items produced by a native EP speaker was included before the main session to help native judges understand the procedure. Another five items produced by the native speaker were placed randomly in the main session serving as control items. For these 5 control items, if a listener corrected on at least 4, her responses would be considered valid, which was the case for all 25 native judges.

#### 4. RESULTS

The L2 production accuracy of the EP /CrV/ clusters judged by 25 native raters were visualised in Figure 1.



**Figure 1:** L1-Mandarin learners' production accuracy as a function of EP onset cluster type.

A visual inspection suggests that both the preceding consonant (the comparison between "pra" and "cra" and "tra") and the following vowel (the comparison between "pri" and "pra") seem to have an influence on the L1-Mandarin learners' production of /r/. To explore these effects, we fitted a mixed-effects logistic regression model, which has *Context* as predictor (four levels: "pri", "pra", "cra" and "tra"). The predictor was coded using Backward Difference Coding, which allows us to compare each level with its adjacent one. Both random intercepts and slopes by participant and random intercepts by test item were included in the model. The model's results are summarised in Table 1.

Contrast	<i>b</i>	<i>SE</i>	<i>p</i> .value	95% CI
pri - pra	-1.66	0.66	0.012	[-2.95, -0.37]

pra - cra	1.85	0.68	<0.01	[0.51, 3.2]
cra - tra	4.32	0.73	<0.001	[2.89, 5.75]

**Table 1:** Results of the mixed-effects logistic regression model.

The statistical analysis confirms the visual inspection. In contrast to previous studies (e.g. [29, 32]), the results of this study reveal that L2 production of EP /r/ is promoted when the tap is followed by /i/ than by /a/. Regarding the preceding consonantal context, the alveolar stop /t/ indeed helped L2 learners produce more target-like tap, in line with Zhou [31]; moreover, the accuracy rate of producing /kr/ is higher than that of /pr/ clusters, implying the following acquisition order /tr/ > /kr/ > /pr/.

#### 5. DISCUSSION

In the current study, we explored phonetic context effects on L2 segmental learning, by studying how the preceding consonant and the following vowel affect L2 production of EP /r/ in onset cluster. 14 native Mandarin speakers first learned to associate 12 EP pseudo-words containing /CrV/ clusters to different pictures and then performed a picture-naming task. The L2 production data, judged by 25 EP native raters, indicate that both the preceding and the following phonetic contexts have an impact on the L2 production of EP /r/.

Regarding the preceding consonant, we found an effect of its place of articulation (/tr/ > /kr/ > /pr/). The fact that the L2 production of /r/ is promoted when preceded by an alveolar stop is in line with our prediction. According to Zhou [31], L1-Mandarin learners tend to use [l] when they fail to produce /r/ in onset cluster (/pr/ as [pl]), but they rarely do so when the preceding stop is alveolar (/d/ or /t/). This is because the replacement with [l] would yield a cross-linguistically marked structure [dl] or [tl]. The phonological markedness may force L2 learners to rely less on such repair strategy, which may in turn lead to fast acquisition of /r/. However, it should be noted that the phonological markedness does not account for the difference with respect to the production accuracy between /kr/ and /pr/, since the resulted repair forms [kl] and [pl] are both allowed in the target language EP.

Another plausible explanation for the observed preceding consonant effect pertains to lexical frequency. Although all test items produced by L1-Mandarin speakers are newly learned words, prior research suggests that, as long as these new words become engaged in the mental lexicon, the processing of these novel lexical items can be influenced by

previously learned words [28]. Given that the Chinese participants were colleague students that mainly received Portuguese input in the classroom setting, the frequency of EP onset clusters in the L2 lexicon was estimated on the basis of their coursebook [5], instead of a database of the native Portuguese lexicon. It turns out that the estimated frequency (/pt/ > /tr/ > /kr/) does not really resemble the observed acquisition order.

Instead of phonological markedness and lexical frequency, we argue that the preceding consonant effect on L2 production of /t/ may stem from the articulatory representation of the EP onset clusters. Following the articulatory-based featural representation for onset clusters proposed by Duanmu [6], one can assume that the EP /tr/ cluster is represented by one articulator [\*Coronal], while two articulators are needed for both /kr/ ([\*Dorsal, Coronal]) and /pr/ [\*Labial, Coronal]. Consequently, the simpler representation for /tr/ may imply less articulatory effort in its production (see also Espiga [8]), in comparison to other clusters. The articulation-based view also offers a plausible account for the order /kr/ > /pr/. Although, to our best knowledge, no articulatory evidence exists for the EP onset clusters, a recent study on the Spanish ones shows that the latency from the end of the target gestural plateau for the stop to the onset of that for /t/ is longer in /kr/ than in /pr/ [10]. This difference might imply an increased articulatory effort for L2 learners in the production of /pr/ clusters, where they have to realise the complex and novel gestural coordination of /t/ (a tongue back-lowering followed by a tongue top raising) in a more limited temporal space.

Turning to the effect of the following vowel (/pri/ > /pra/), we speculate that the nature of this vocalic effect might be articulatory as well. The articulation of [r] has been shown to be highly influenced by its surrounding vocalic context [4]. For instance, when surrounded by /i/, the tongue tip of [r] is more raised and advanced toward the alveolar region than in any other vocalic context (e.g. /a/). Accordingly, the vocalic context /i/ might have helped L2 learners realise this tongue tip gesture, which is crucial for the articulation of a tap.

The vowel effect observed in this study goes in the opposite direction to what has been reported previously (e.g. [29, 32]). Two methodological factors may underlie this asymmetry, namely the stimulus type and the examined speech modality. In this study, the vocalic context can only exert backward effect on the production of /t/, as the EP rhotic only occurs in onset cluster (/CrV/), while both forward and backward influences are expected in Vale [29] and Zhou [32], where the EP /t/ was assessed in intervocalic onset position (/V.rV/).

Moreover, the current study targets the contextual effects in L2 speech production, while Vale [29] and Zhou [32] examined L2 speech perception. Many (psico)linguistic models assume explicitly that speech perception and production involve different representational levels and paralinguistic processes [2], which seems to be supported by an increasing number of studies showing that L2 learners' production may not mirror their perceptual performance (e.g. [15, 24]). Future studies targeting both speech modalities are needed, in order to gain a better understanding of the surrounding vowel effect on L2 acquisition of EP /t/.

In our view, the current findings regarding the phonetic context effects on L2 segmental learning have important implications for the development of L2 speech theories. Some major L2 speech models (e.g. [1, 9]) are mainly concerned with the acquisition of a phonetic or phonological category in isolation. It is unclear how the phonetic context effects can be accounted for by these models. Nevertheless, L2 phonetic context effects can be nicely integrated, if one assumes an explicit distinction between phonetics and phonology, as in the *Bidirectional Phonology and Phonetics Model* [2]. In particular, the phonetic form necessarily involved in speech production is a continuous representation, consisting of articulatory gestures, e.g. tongue and lip movements, jaw depression. This phonetic representation is subject to a set of articulatory constraints, which may reflect the articulatory effort during speech production. For instance, an OT-like constraint ranking instantiating articulatory effort \*LIPS<sub>ART</sub> > \*TONGUE TIP<sub>ART</sub>, borrowed from Boersma (2011), may account for the acquisition order /tr/ > /pr/, because the articulatory form of /pr/, which involves both tongue lip and tongue tip gestures violates both constraints, while the articulatory form of /tr/ is only penalised by the lower-ranked constraint that militates against tongue tip gesture.

## 6. ACKNOWLEDGEMENTS

This work was partially funded by grant UIDB/00214/2020 from the Portuguese Foundation of Science and Technology (FCT).

## 7. REFERENCES

- [1] Best, C. T., Tyler, M. D. 2007. Nonnative and second-language speech. *Language experience in second language speech learning: In honor of James Emil Flege*, 13-34.
- [2] Boersma, P. 2011. A programme for bidirectional phonology and phonetics and their acquisition and evolution. In *Bidirectional Optimality Theory*, eds.



- Anton Benz and Jason Mattausch, 33–72. Ams- terdam: Benjamins.
- [3] Boersma, P., Weenink, D. 2019. Praat. Computer program, version 6.0.48. Retrieved from [www.praat.org](http://www.praat.org).
- [4] Celata, C., Vietti, A., Spreafico, L. 2018. An articulatory account of rhotic variation in Tuscan Italian: Synchronized UTI and EPG data, in Mark Gibson, and Juana Gil (eds), *Romance Phonetics and Phonology*. Oxford Academic.
- [5] Coimbra I., Coimbra, O. M. 2018. *Novo Português sem Fronteiras*. LIDEL.
- [6] Duanmu, S. 2002. Two theories of onset clusters. *Chinese Phonology*. 11, 97-120.
- [7] Dumay, N., Gaskell, M. G. 2007. Sleep-associated changes in the mental representation of spoken words. *Psychological Science*. 18(1), 35-39.
- [8] Espiga, J. W. D. R. 1997. *Influência do espanhol na variação da lateral posvocálica no português de fronteira*. MA dissertation, Universidade Católica de Pelotas.
- [9] Flege, J. E. 1995. Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research*. 92, 233-277.
- [10] Gibson, M., Sotiropoulou, S., Tobin, S., Gafos, A. I. 2019. Temporal aspects of word initial single consonants and consonants in clusters in Spanish. *Phonetica*. 76(6), 448-478.
- [11] Hallé, P. A., Segui, J., Frauenfelder, U., Meunier, C. 1998. Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance*, 24(2), 592–608.
- [12] Hallé, P. A., Best, C. T. 2007. Dental-to-velar perceptual assimilation: A cross-linguistic study of the perception of dental stop+/l/ clusters. *The Journal of the Acoustical Society of America*, 121(5), 2899–2914.
- [13] Hayes-Harb, R., Masuda, K. 2008. Development of the ability to lexically encode novel second language phonemic contrasts. *Second Language Research*. 24(1), 5-33.
- [14] Jesus, L. M., Shadle, C. H. 2005. Acoustic analysis of European Portuguese uvular [χ, ʁ] and voiceless tapped alveolar [t̪] fricatives. *Journal of the International Phonetic Association*. 35(1), 27-44.
- [15] de Leeuw, E., Stockall, L., Lazaridou-Chatzigoga, D., Gorba Masip, C. 2019. Illusory vowels in Spanish–English sequential bilinguals: Evidence that accurate L2 perception is neither necessary nor sufficient for accurate L2 production. *Second Language Research*, 37(4), 587–618.
- [16] Levy, E. S. 2004. *Effects of language experience and consonantal context on perception of French front rounded vowels by adult American English learners of French*. PhD dissertation, City University of New York, USA.
- [17] Levy, E. S., Strange, W. 2008. Perception of French vowels by American English adults with and without French language experience. *Journal of phonetics*. 36(1), 141-157.
- [18] Mateus, M. H.M., d'Andrade, E. 2000. *The phonology of Portuguese*. OUP Oxford.
- [19] Mateus, M.H.M., Rodrigues, C. 2003. A vibrante em coda no português. *Teoria Lingüística. Fonologia e outros temas*, 181-199.
- [20] Recasens, D. 2022. Acoustic characteristics and placement within vowel space of full schwa in the world's languages: A survey. *Journal of the International Phonetic Association*, 52(1), 59-94.
- [21] Rodrigues, C. 2003. *Lisboa e Braga: Fonologia e Variação*. Ph.D. Dissertation, University of Lisbon, Portugal.
- [22] Ogden, D. C. 2015. Phonetic context effects in second-language phonetic category discrimination. *The Journal of the Acoustical Society of America*. 137(4), 2381–2381.
- [23] Oliveira, D. 2021. *Auditory selective attention and performance in high variability phonetic training: The perception of Portuguese stops by Chinese L2 learners*. PhD thesis, University of Minho, Portugal.
- [24] Sakai, M. 2016. (Dis)Connecting perception and production: Training native speakers of Spanish on the English /i/-/I/ distinction. PhD dissertation). Georgetown University, USA.
- [25] Silva, A. 2014. *Análise Acústica da Vibrante Simples do Português Europeu*. Dissertação de Mestrado, Universidade de Aveiro.
- [26] Stilp, C. 2020. Acoustic context effects in speech perception. *Wiley Interdisciplinary Reviews: Cognitive Science*. 11(1), e1517.
- [27] Svantesson, J. O., Tsendina, A., Karlsson, A., Franzén, V. 2005. *The phonology of Mongolian*. OUP Oxford.
- [28] Tamminen, J., Gaskell, M. G. 2008. Newly Learned Spoken Words Show Long-Term Lexical Competition Effects. *Quarterly Journal of Experimental Psychology*, 61(3), 361–371.
- [29] Vale, A. A. 2020. *Perceção das Consoantes líquidas /r/ e /l/ do Português Europeu sob Influência do Mandarim L1*. MA dissertation, University of Minho, Braga, Portugal.
- [30] Yang, S. 2014. *Perceção das consoantes oclusivas de Português L2 sob a influência de Mandarim L1*. MA dissertation, University of Minho, Braga, Portugal.
- [31] Zhou, C. 2017. *Contributo para o estudo da aquisição das consoantes líquidas do português europeu por aprendentes chineses*. MA dissertation, University of Lisbon, Lisbon, Portugal.
- [32] Zhou, C. 2021. L2 speech learning of European Portuguese /l/ and /r/ by L1-Mandarin learners: Experimental evidence and theoretical modelling. *Language Acquisition*. DOI: 10.1080/10489223.2021.1952591