

How to model the influence of orthography on L2 representations with BiPhon Neural Networks

Silke Hamann, University of Amsterdam
& Chao Zhou, University of Lisbon

Many studies have shown that written forms influence the acquisition of a second language. This influence can be helpful, as is the case of the English /æ/-/ɛ/ contrast that is notoriously difficult for Dutch learners but where the written form can aid in the creation of the distinction (e.g., Weber & Cutler 2004; Escudero & Wanrooij 2010). But orthography can also cause the creation of so-called ghost contrasts, which do not exist in the L2, as is the case with the intervocalic singleton/geminate contrast in the L2 English of Italian speakers (e.g., Bassetti 2017; Hamann 2018).

In this talk, we illustrate how such orthographic influences on the creation of L2 representations can be formalized, by this yielding theoretical predictions that can be tested again in experimental studies. Our formalization is performed with a symbolic neural network based on the Bidirectional Phonetics-Phonology model (Boersma 2007) and its extension by a reading grammar (Hamann & Colombo 2017).

Our main data comes from an experimental study on Mandarin (Zhou & Hamann 2020): 23 L1-Mandarin speakers with no prior knowledge of EP (naïve listeners), representing the initial stage of L2 acquisition, performed a delayed-imitation task. They were presented with EP nonce words containing /r/ in intervocalic onset (e.g., *parafa*) or word-internal coda (e.g., *parfa*), first auditorily, and then with accompanying orthography. Our results show 1) that participants only produced L1 [ɹ] when exposed to orthography, confirming that the use of Mandarin rhotic in L2 speech is orthographically driven; and 2) that even at the initial stage the substitution with Mandarin [ɹ] occurs almost exclusively in coda position, reminiscent of L2 learners (Zhou 2017; Liu 2018).

Bassetti, Bene (2017). Orthography affects second language speech: double letters and geminate production in English. *Journal of Experimental Psychology: Learning, Memory and Cognition*, DOI: 10.1037/xlm0000417.

Boersma, Paul (2007). Some listener-oriented accounts of h-aspiré in French. *Lingua* 117: 1989–2054.

Boersma, Paul, Benders, Titia, & Klaas Seinhorst (2020). Neural network models for phonology and phonetics. *Journal of Language Modelling*, 8(1), 103–177. <https://doi.org/10.15398/jlm.v8i1.224>

Escudero, Paola & Karin Wanrooij (2010). The effect of L1 orthography on non-native vowel perception. *Language and Speech* 53: 343–365.

Hamann, Silke (2018). “Ghost phonemes in second languages – How orthography can create contrasts without perceptual correlates.” Presentation at the *15th Old World Conference in Phonology*, London.

Hamann, Silke & Ilaria Colombo (2017). A formal account of the interaction of orthography and perception: English intervocalic consonants borrowed into Italian. *Natural Language & Linguistic Theory* 35: 683–714.

Liu, Wen. (2018). *Aquisição da Vibrante Simples [r] pelos Alunos Chineses Aprendentes de Português como Língua Estrangeira*. Unpublished MA thesis, University of Macau, Macau, China.

Zhou, Chao (2017). *Contributo para o estudo da aquisição das consoantes líquidas do português europeu por aprendentes chineses*. Unpublished MA thesis, University of Lisbon, Lisbon, Portugal.

Zhou, Chao & Silke Hamann (2020). Cross-linguistic interaction between phonological categorization and orthography predicts prosodic effects in the acquisition of Portuguese liquids by L1-Mandarin learners. *Proceedings of Interspeech 2020*, Hong Kong, 4486–4490. doi: 10.21437/Interspeech.2020-2689