JYX JYVÄSKYLÄN YLIOPISTO UNIVERSITY OF JYVÄSKYLÄ

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Zetterholm, Elisabeth; Kuronen, Mikko; Nyyssönen, Valtteri

Title: The impact of a short auditory training on L2 pronunciation in languages with different orthographic depth

Year: 2021

Version: Published version

Copyright: © ISCA 2021

Rights: In Copyright

Rights url: http://rightsstatements.org/page/InC/1.0/?language=en

Please cite the original version:

Zetterholm, E., Kuronen, M., & Nyyssönen, V. (2021). The impact of a short auditory training on L2 pronunciation in languages with different orthographic depth. In ISAPh 2021 : Proceedings of the 3rd International Symposium on Applied Phonetics (pp. 71-75). ISCA. https://doi.org/10.21437/ISAPh.2021-13

The impact of a short auditory training on L2 pronunciation in languages with different orthographic depth

Elisabeth Zetterholm[1], Mikko Kuronen[2], Valtteri Nyyssönen[2]

[1]Linköping University, Sweden, [2]Jyväskylä University, Finland,
[1]*elisabeth.zetterholm@liu.se*, [2]*mikko.j.kuronen@jyu.fi*, [2]*valtteri.j.nyyssonen@gmail.com*

Abstract

It is sometimes assumed that the pronunciation of an L2 is more predictable and thus easier to learn if its orthography is transparent. This study aims to find out whether this assumption holds true in the first stages of L2 learning in languages with different orthographic depth. The study also examines the effect that a short auditory training (supported by simultaneous orthographic input) has on L2 pronunciation. A central finding was that the pronunciation of an L2 with a transparent orthography was not easier to learn for a naïve learner when compared to an L2 with an opaque orthography. A second finding was that even a short period of auditory training can introduce a significant improvement of a naïve L2 learner's speech. Further, the results show that mimicking one specific native speaker could be an effective strategy for pronunciation learning. This method should be studied in more detail in future research.

Keywords: Orthography, orthographic depth, L2 pronunciation, auditory training, mimicking.

1. Introduction

Languages differ with respect to the degree of orthographic transparency, also called orthographic depth [1, 2, 3]. Orthographic depth refers to the phoneme-to-grapheme (or letter-to-phoneme) correspondence [4, 5]. A language with a transparent (shallow) orthography has a relatively consistent correspondence between orthography and pronunciation, whereas in a language with an opaque (deep) orthography the correspondence is relatively inconsistent [4]. For example, in Finnish the phoneme-to-grapheme correspondence is transparent in the sense that graphemes often relate to phonemes in a way which is intuitively close and consistent [2, 6]. Thus, the grapheme <o> is pronounced in Finnish as [o], <oo> as [o:], <u> as [u], <uu> as [u:], <ee> as [e:], <i> as [ɪ] etc. almost without exceptions. In a language with an opaque orthography – for example Swedish – the phoneme-to-grapheme correspondence is more inconsistent in the way that there are often several ways to pronounce one letter (feedforward, i.e., letter-to-phoneme inconsistency) and several ways to spell one sound (feedback, i.e., phoneme-to-letter inconsistency). Problems arise especially when a grapheme represents some other sound that its regular correspondent; e.g., in Swedish, /ʃ/ can be spelled <skön>, <stjärna>, <skjuta>, <choklad>, <giraff>, <jalusi>, <shopping> and <garage> (Eng. lovely, star, to shoot, chocolate, giraffe, jalousie, shopping and garage). Thus, a Finnish second language (L2) learner of Swedish must learn many new grapheme-to-phoneme and phoneme-to-grapheme correspondences, and not to use some in his/her first language (L1), according to what learning of new orthography-pronunciation correspondences in L2 means [7].

When acquiring their L1, children learn pronunciation primarily from auditory input and oral communication [8]. In turn, adult L2 learners are often simultaneously exposed to both auditory and orthographic input, and both types of input affect pronunciation learning. Orthographic input can affect both the perception and production of L2 speech [9]. The impact of orthographic input on pronunciation has been proved in several studies on L2 acquisition. In a series of experiments, Bassetti and Atkinson showed that orthographic forms can affect even experienced learners' pronunciation of known words in L2: 85% of Italian learners of English added a phone to the target word (<walk> was pronounced with /l/ etc.) in reading aloud tasks, while 56% of the participants added a phone to the target word in word repetition [10]. Words and syllables that follow the most frequent grapheme-to-phoneme correspondences (e.g., <hit> /hɪt/ in English) are considered regular, and they are usually not problematic for L2 learners. In this case, orthographic input can be a visual support to auditory input and help the learner to both perceive and produce the word correctly [3, 6]. For example, Japanese learners of English have been argued to be able to pronounce [l] and [r] correctly also when they do not perceive the distinction, if they know whether the word is spelled with <l> or <r> [9]. Irregular words do not

10.21437/ISAPh.2021-13

follow the usual correspondences, which is why it is harder to learn their correct pronunciation.

Sometimes the effect of orthography is so strong that it overrides an L2 learner's ability to hear the actual pronunciation of the word, thus preventing him/her from learning pronunciation accurately [10, 11, 12]. Further, it seems that exposure to orthographic input alone leads more often to non-target like pronunciation than exposure to auditory input. This is probably one reason for length of residence (LOR) in an L2 environment seeming more beneficial for learning L2 pronunciation than formal instruction, which is often based on or supported by written forms of language [13, 14, 15]. Adult multilingual learners can use previous knowledge from other languages to find out links between orthography and pronunciation [16]. However, it is difficult to teach pronunciation from spelling alone.

In the present paper, we study the impact of L2 orthography on L2 pronunciation in the very first stages of L2 learning, more specifically when the learner has no prior knowledge of the target language; a naïve learner. Since earlier studies on the effect of orthographic input on L2 pronunciation have focused on L2 learners with different experience of and exposure to the target language, the naïve learner's perspective is a novelty of our study. The learner's L1 is Swedish, and the target languages are Finnish and Portuguese. The target languages were chosen based on their degree of orthographic depth. In Finnish, the orthography is nearly transparent [2, 6], while Portuguese has clearly a more opaque orthography [2].

Further, we study what kind of an effect a short auditory training (supported by simultaneous orthographic input) has on L2 pronunciation in languages with different orthographic depth.

1.1. Aim and research questions

The first aim of the study is to augment our understanding of the impact of orthography on pronunciation in the first stages of L2 learning. Secondly, we study what kind of an effect a short auditory training has on L2 pronunciation in languages with different orthographic depth. We address the following research questions:

1. Is an L2 with a transparent orthography easier to pronounce for a naïve learner than an L2 with an opaque orthography?
2. What kind of effect does a short auditory training have (supported by simultaneous orthographic input) on comprehensibility and accuracy of L2 speech in languages with different orthographic depth?

2. Method and Material

A professional speech impersonator (a middle-aged male) with Swedish as his L1 read a short text in Finnish and Portuguese. Both languages were previously unfamiliar to him and had different degrees of orthographic depth. First, the impersonator read aloud a passage (ca. 150 words) from a novel without any training or instructions how to pronounce the text. Next, he listened to the same text read by a native speaker of the language. He was requested to listen to, mimic and train, and – when he thought he was ready for the task – read aloud the same texts again. Thus, he also had the text available during the auditory training and the second reading. The training session before the second reading was about two hours in both languages. All recordings were done in the impersonator's own audio studio, sent to the researchers by a secure internet link, stored on external disks, and only available for the researchers.

We assume that the impersonator's L1 Swedish affects his pronunciation especially in the first recordings because he has no knowledge of the target languages (although he might recognize them). Thus, he has to rely at least to some degree on his L1. In the second recording he has gained some knowledge of the orthography-pronunciation correspondences and discrepancies in the target languages. These hypotheses indicate that comprehensibility should be better and the pronunciation more accurate in the second recordings.

2.1. Listener test

A listener test with pre- and post-training recordings (hereafter Test 1 and Test 2) was constructed. Native speakers of Finnish and Portuguese rated comprehensibility of the reading on a scale from 1 to 6. The scale was described with the following wordings: 1 = I understand nothing, 2 = I understand a couple of words, 3 = I understand quite little, 4 = I understand quite much, 5 = I understand almost everything, 6 = I understand everything. Further, the listeners graded the accuracy of pronunciation in the recordings by answering (1) whether Test 1 or Test 2 sounded better ('better' in the sense 'closer to the target language'; a forced choice with 4 alternatives), and (2) what pronunciation features caused the possible difference between the tests (an open question). The listener test was done with an online survey tool. The listeners were able to listen to the speech samples as many times as they wanted and advance in the test at their own pace. All in all, the test took ca. 15 minutes to complete. The listeners had Finnish (n = 28) and Portuguese (n = 30) as their L1. The listener selection was based on two criteria: self-reported L1 and a minimum age of 18 years; i.e., the

listeners consisted of a group of adult L1 listeners of the language.

Statistical significances were calculated by (1) Wilcoxon Signed-Rank Test for the difference between Test 1 and Test 2 within the same language, (2) Mann-Whitney U for the difference between the languages in both Test 1 and Test 2, respectively, and (3) paired samples t-test for the difference between Test 1 and Test 2 in all data.

3. Results

First, we present comprehensibility as rated by the L1 listeners, and thereafter the accuracy of pronunciation as rated by the L1 listeners.

3.1. Comprehensibility as rated by L1 listeners

On average, the listeners understood "quite little" (3) or "quite much" (4) (on a scale from 1 to 6) in both languages in Test 1. The Finnish listeners rated the comprehensibility as somewhat lower than the Portuguese listeners. The mean score was 3.50 in Finnish and 3.93 in Portuguese in Test 1 (a non-significant difference, *p=.070*, cf. Table 1). The results show that listeners understood something – in many cases quite much – of both languages in Test 1. The finding is in line with human listeners' ability to interpret even fairly distorted (in the sense of not target language like) speech signals. In Test 2, the difference between the languages was smaller than in Test 1; the mean score was 4.21 in Finnish and 4.43 in Portuguese (a non-significant difference, *p=.376*, cf. Table 1). Thus, a short auditory training (supported by simultaneous orthographic input) improved comprehensibility of the L2 pronunciation evidently in both languages (Table 2).

Table 1: Pre-training (1) and post-training (2) mean, standard deviation (SD) for comprehensibility in the two languages and significance differences (Wilcoxon Signed-Rank Test) between the languages are given in the table.

	Test 1		Test 2	
	Portuguese	Finnish	Portuguese	Finnish
Mean	3.93	3.50	4.43	4.21
SD	0.73	0.92	0.89	0.99
Sig.	.0.70		.376	

Table 2: Pre-training (1) and post-training (2) mean for comprehensibility in the two languages. Significance differences (Mann-Whitney U) between Test 1 and Test 2 are also given in the table.

	Test 1 and Test 2		Test 1 and Test 2	
	Finnish		Portuguese	
Mean	3.50	4.21	3.93	4.43
Sig.	.001		.008	

As regards comprehensibility, we can conclude that (1) on average, the listeners understand *quite little* or *quite much* of both languages in Test 1, (2) the languages were equally easy or difficult to pronounce for the naïve L2 speaker, (3) both languages got significantly higher comprehensibility ratings in Test 2 than in Test 1, and (4) the effect of a short auditory training (supported by simultaneous orthographic input) was similar and evident in both languages.

3.2. The accuracy of pronunciation as rated by L1 listeners

The listeners compared the accuracy of pronunciation between Test 1 and Test 2 in both languages. They were asked to answer to a forced choice question with four alternatives: *Test 1 sounded better*, *the tests were equal* (= they sounded equally good or bad), *Test 2 sounded better* or *Test 2 sounded much better*. In this context, the word 'better' meant 'closer to the target language', hence the term accuracy. In addition, the listeners were asked (not forced) to comment on what pronunciation features caused the possible difference between the tests. The answers were given a numeric value in the analysis: negative development (= 0); no changes, the tests were equal (= 1); Test 2 was better (= 2); Test 2 was much better (= 3). The results reveal that Test 2 was considered *better* or *much better* in both languages. Portuguese underwent the most positive development (cf. Table 3).

Table 3: Development in accuracy of pronunciation as rated by the listeners (negative development = 0, no development = 1, Test 2 was better = 2, Test 2 was much better = 3).

	Finnish	Portuguese
0	1	0
1	4	5
2	21	18
3	2	7
Mean	1.85	2.06

Listeners' comments on what pronunciation features the development was caused by are presented in Figure 1.

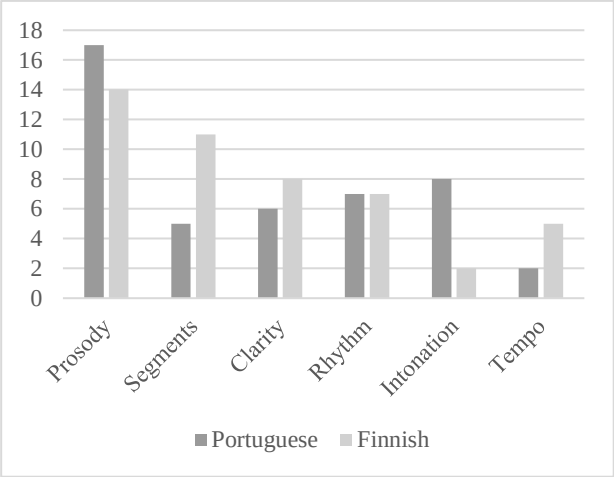

Figure 1: Listeners' comments (n = 58) on what pronunciation features the development in Test 2 concerned. The number of individual comments is given on the y-axis, the commented features on the x-axis. Regarding Prosody, the prosodic features Rhythm, Intonation and Tempo[1] are given separately in the figure. Further, Clarity is given as own feature, because the comments on Clarity concerned overall clarity of the speech, not prosody or segments.

Even though Finnish and Portuguese have considerable phonetic differences, the pattern of development in the accuracy of pronunciation was similar in the two languages: Prosodic features, especially rhythm (the term was used by the listeners; we don't know exactly what they meant by it), were highly valued by the listeners in both languages (cf. Figure 1). Quite many Finnish listeners also mentioned a better pronunciation of segments in Test 2, while Portuguese listeners mentioned intonation as a feature that was better in Test 2 (cf. Figure 1).

Prosodic features, especially rhythm, have been shown to correlate most with comprehensibility in earlier studies on L2 pronunciation in many different languages [17, 18, 19, 20]. Thus, it is perhaps not surprising that the development mainly concerned prosodic features also in the present study. The listeners might have paid more attention to segmental features if the speaker had been more advanced and his speech easier to understand.

Concerning the accuracy of pronunciation, we can conclude that there was a significant development between the two tests in both languages. We also found a clear connection between the accuracy of pronunciation and comprehensibility: both underwent an evident improvement from Test 1 to Test 2.

4. Summary

A first finding of our study is that an L2 with a transparent orthography is not easier to pronounce for a naïve L2 learner than an L2 with an opaque orthography. This seems to be the case at least in the very first stages of language learning when the L2 learner has not yet gained knowledge of the orthography-pronunciation correspondences and discrepancies in the L2. This result is an answer to our first research question.

A second finding of our study, answering the second research question, is that even a short period of auditory training (supported by simultaneous orthographic input) can induce a significant improvement in both comprehensibility and accuracy of L2 speech. This improvement was similar in the languages despite the difference in orthographic depth between them.

Previous studies have suggested that transparent orthography is beneficial for L2 pronunciation learning [3, 10]. This is probably true in later stages of L2 learning than was the case in the present study. Yet, we found it questionable how positive an effect orthographic input can have on L2 pronunciation learning – especially the accuracy of pronunciation (= the accentedness) – in any language, because the phonetic realization differs from the orthography in tens of small but important details also in languages with a transparent orthography. For example, in Finnish important prosodic features as sentence stress, intonation, proportional syllable durations within words and phrases and reduction are not visible in the orthography.

The second finding of our study – that even a short period of auditory training can induce improvement in both the comprehensibility and accuracy of L2 speech – seems to be the case at least when the L2 speaker is skillful in mimicking an L1 speaker. In our study, the impersonator could concentrate on just mimicking without the cognitive burden of producing grammatical and pragmatically functional speech [21], and he had one specific L1 speaker as his pronunciation model. Even with ordinary L2 speakers, mimicking one specific native speaker could be an effective strategy for pronunciation learning. This method should be studied in more detail in future research. Further, the results suggest that mimicking an L1 speaker could be effective especially when learning the important prosodic features. Prosodic features can be seen as the larger, more general characteristics of a language than segments, which is why they might be easier to mimic than segmental features. Prosodic features may also be more salient for the naïve listener for the same reason.

5. Acknowledgements

A special thanks to the speech impersonator. This study is part of the research project *Fokus på uttalsinlärningen med svenska som mål- och*

källspråk (FOKUS, see www.jyu.fi/fokus), which was funded by the Society of Swedish Literature in Finland (SLS) 2015–2019.

6. References

[1] van den Bosch, A., Content, A., Daelemans, W., de Gelder, B. (1994). Analysing orthographic depth of different languages using data-oriented algorithms. *Proceedings of the 2nd International Conference on Quantitative Linguistics, Moscow, September 1994.* 1–9. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.48.9845.

[2] Seymour, P. H. K., Aro, M., Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94(2), 143–174. https://doi.org/10.1348/000712603321661859.

[3] Erdener, V. D., Burnham, D. K. (2005). The role of audiovisual speech and orthographic information in nonnative speech production. *Language Learning*, 55(2), 191–228. https://doi.org/10.1111/j.0023-8333.2005.00303.x.

[4] Katz, L., Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. *Advances in Psychology*, 94(2), 67–84. https://doi.org/10.1016/S0166-4115(08)62789-2.

[5] Schmalz, X., Marinus, E., Coltheart, M., Castles, A. (2015). Getting to the bottom of orthographic depth. *Psychonomic Bulletin & Review*, 22(6), 1614–1629. doi.10.3758/s13423-015-0835-2.

[6] Iivonen, A., Tella, S. (2009). Vieraan kielen ääntämisen ja kuulemisen opetus ja harjoittelu [Teaching and practicing of second language pronunciation and listening comprehension]. In O. Aaltonen, R. Aulanko, A. Iivonen, A. Klippi, M. Vainio (Eds.), *Puhuva ihminen – Puhetieteiden perusteet* (pp. 269–281). Helsinki: Otava.

[7] Hayes-Harb, R., Nicol, J., Barker, J. (2010). Learning the phonological forms of new words: Effects of orthographic and auditory input. *Language & Speech*, 53(3), 367–381. https://doi.org/10.1177/0023830910371460.

[8] Clark, E.V. (2009). *First language acquisition* (2nd ed.) Cambridge: Cambridge University Press.

[9] Basetti, B. (2008). Orthographic Input and Second Language Phonology. In T. Piske, M. Young-Scholten (Eds.), *Input Matters in SLA* (pp. 191–206). Clevedon: Multilingual Matters. https://doi.org/10.21832/9781847691118-013.

[10] Bassetti, B., Atkinson, N. (2015). Effects of orthographic forms on pronunciation in experienced instructed second language learners. *Applied Psycholinguistics*, 36(1), 67–91. https://doi:10.1017/S0142716414000435.

[11] Jarvis, S., Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. New York, NY: Routledge.

[12] Peltola, K. U., Tamminen, H., Alku, P., Peltola, M. S. (2015). Non-native production training with an acoustic model and orthographic or transcription cues. The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow: University of Glasgow. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0236.pdf

[13] Suter, R.W. (1976). Predictors of pronunciation accuracy in second language learning. *Language Learning*, 26(2), 233–253. https://doi.org/10.1111/j.1467-1770.1976.tb00275.x.

[14] Purcell, E., Suter, R.W. (1980). Predictors of pronunciation accuracy: A reexamination. *Language Learning*, 30(2), 271–287. https://doi.org/10.1111/j.1467-1770.1980.tb00319.x.

[15] Piske, T., MacKay, I.R.A., Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2. A review. *Journal of Phonetics*, 29(2), 191–215. https://doi.org/10.1006/jpho.2001.0134.

[16] Bartolotti, J., Viorica, M. (2018). Learning and processing of orthography-to-phonology mappings in a third language. *International Journal of Multilingualism*, 16(4), 377–397. https://doi.org/10.1080/14790718.2017.1423073.

[17] Saito, K., Trofimovich, P., Isaacs, T. (2016). Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. *Applied Psycholinguistics*, 37(2),217–240. https://doi.org/10.1017/S0142716414000502.

[18] Anderson-Hsieh, J., Johnson, R., Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning* 42(4), 529–555. https://doi.org/10.1111/j.1467-1770.1992.tb01043.x.

[19] Kang, O., Rubin, D., Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of English language learner proficiency in oral English. *Modern Language Journal* 94, 554–566. https://doi.org/10.1111/j.1540-4781.2010.01091.x.

[20] Heinonen, H. (2020). *Uttal och dess begriplighet i finskspråkiga gymnasisters L2-svenska*. JYU Dissertations 296. Jyväskylä: University of Jyväskylä. http://urn.fi/URN:ISBN:978-951-39-8325-3.

[21] Ullakonoja, R. (2011). *Da. Eto vopros! Prosodic development of Finnish students' read-aloud Russian during study in Russia.* [Doctoral dissertation.] Jyväskylä: University of Jyväskylä. http://urn.fi/URN:ISBN:978-951-39-4209-0.

[1] Tempo could also be classified as a component of fluency but is regarded in Figure 1 as a prosodic feature.