

THE LONG AND SHORT OF IT: L2 PRODUCTION OF FINNISH GEMINATES

Jeanne McGill, Indiana University Bloomington

The acquisition of Finnish by L1 speakers of American English is relatively understudied, although the language pairing is of linguistic interest, especially because Finnish contains frequent phonemic geminates (e.g. a /k/ vs. /kk/ contrast), while in English, spelled doubled consonants are not lengthened. Additionally, a mismatch exists between English prosody and Finnish geminates, which appear in contexts where English speakers are not listening for consonant clusters or used to producing them. This work presents a pilot case study of an L1 English speaker with one year of Finnish study, who was recorded reading a word list containing distractors and minimal pairs that had the contrast intervocalically. The learner's production and a native speaker's (NS) production of the same minimal pairs were measured using Praat. Results show that in 7/8 pairs her geminates were longer than her singletons. However, when geminate vs. singleton durational ratios were calculated for both speakers, only 2/7 approximated the NS's production. In 6/7 cases her singletons were longer than the NS's, at times approaching NS's geminate length, which could cause communication breakdowns in real speech. Further research needs to be done to determine an acquisitional path; however, results suggest that instructors should teach that geminates are lengthened but focus on shortening singletons for better clarity in communication.

INTRODUCTION

The study of Finnish as a foreign language is increasingly popular in North America: the language learning company Duolingo said that Finnish was their most requested course ever at the time of its release in June 2020 (Smith, A.O., Helsinki Times, 29 June 2020). At least 7,500 people located in the US began the course within four days of its release. It is noteworthy that many people learning Finnish do so via English even if not native English speakers. Additionally, Finnish has thriving programs at multiple North American universities (e.g., University of Toronto, University of Minnesota). However, the acquisition of the singleton/geminate contrast in Finnish could pose a challenge. English does not lengthen spelled doubled consonants, while in Finnish they must be lengthened because the contrast can be phonemic, like *palo/pallo* (fire/ball), *kuka/kukka* (who/flower), and *tapaa/tappaa* (meets/kills). Finnish geminates also appear in contexts where native speakers of North American English are not listening for them nor used to producing consonant clusters.

Cross-Linguistic Realization of Geminates

Geminates pattern differently across languages (Davis, 2011). For stop consonants, geminates can be realized by holding the closure longer, while for other types of consonants like fricatives or liquids, the sound can be held longer. While other differences exist in the phonetic realization of geminates (Doty et al., 2007), length is the main indicator (Dmitrieva, 2012, p. 7).

According to Maddieson (1984), one database with 317 languages contained only 11 with a contrast between short and long consonants, yet Finnish has contrastive length for both consonants and vowels. Dmitrieva (2012) found that geminates typically occur “in intervocalic positions, often after stressed vowels, but are avoided in adjacency to other consonants and on word boundaries, more so word-initially than word-finally; sonorant geminates are more infrequent than obstruent geminates” (p. iv). She argues that geminates are easier to perceive in these contexts; in other words, it is not difficult articulatory production that drives where and how geminate consonants are realized, but perception.

While Finnish follows some of these tendencies, such as the restriction on geminates in word boundaries, it differs in at least two ways. First, Finnish allows post-consonantal geminates in consonant clusters such as /rkk/, a phenomenon so rare that it is not even tested in Dmitrieva’s large-scale study of geminate production and perception. Secondly, Finnish does not follow the tendency to have more obstruent geminates than sonorants. These differences maximize the frequency because there are fewer restrictions.

How frequently do geminates appear in Finnish in relation to other geminate languages? Aoyama (2001) examined two small corpora of non-academic speech (books for juveniles) to determine geminate frequency in Japanese and Finnish, since both languages have a length contrast with vowels and consonants. By counting all consonants and vowels and marking them as short/single or long/geminate, Aoyama found that the ratio in Finnish was 9.9 singleton vowels for every geminate one, and 9.2 singleton consonants for every geminate. In Japanese, geminate vowels only occurred at a rate of 15 to 1 and geminate consonants at a ratio of 19.6 to 1 (p. 25). These data suggested that Finnish has about twice as many geminates as Japanese, making this contrast unavoidable in Finnish.

Acquisition of the Finnish Contrast

The relative frequency of geminates in Finnish could account for Finnish children acquiring them more quickly than Japanese children do in Japanese. Finnish infants can distinguish between short and long consonants by six months (Richardson, 1998) and can produce them by the end of the one-word period (Kunnari et al., 2001), whereas Japanese children cannot (Savinainen-Makkonen, 2007). The input children hear is clear: Engstrand and Krull (1994) found that, even in lively conversational speech, although some reduction was observed, Finnish adults speaking to each other maintain a consistent contrast between singletons and geminates. This clarity is necessary for comprehension, because many combinations are possible: i.e., *muta* (mud), *muuta* (other), *mutta* (but), and *muuttaa* (to move).

L2 acquisition problems are likely because of the mismatch between English and Finnish consonantal prosody. Dmitrieva (2012) reviewed previous literature about American English and concluded that “in English consonant shortening applies in clusters, while consonant lengthening happens in the onsets of stressed syllables and at the edges of words” (p. 91). As noted above, however, in Finnish geminates can appear post-consonantally in these consonant clusters: /rkk/, /rpp/, /rtt/, /ntt/, /nss/, /nkk/, /lkk/, /ltt/, and /mpp/. While Dmitrieva did test the pre-consonantal position for geminates, the post-consonantal position was not considered. If English consonant clusters are shortened, the mismatch between that tendency and geminate consonant clusters is

apparent. Considering the other direction, the conditions under which consonant lengthening occurs in English do not correspond to Finnish. Geminate consonants are not allowed in the onset of stressed syllables, because in Finnish, the first syllable is always stressed, and geminate consonants do not occur word-initially or word-finally.

Undoubtedly, it is vitally important for learners of Finnish to acquire the contrast to understand and communicate clearly, yet little research has been done to understand the learning task or the developmental pathway for English speakers. Further research is needed, but what we do know about acquiring the contrast will be outlined here.

Considering the starting point, naive listeners of Finnish who were monolingual native speakers of American English ($n = 26$) did not seem sensitive to consonant length (Daidone et al., 2023). Another study (Porretta & Tucker, 2014) examined if instruction on the contrast boosted ability to perceive it by dividing naive listeners of Finnish, all L1 speakers of Canadian English, into a *No-Instruction* group and an *Instruction* group. Both groups categorized non-words as having a long or short intervocalic consonant. However, the *Instruction* group was informed that they would hear Finnish words that could have this contrast and were given a written example, while the *No-Instruction* were given no information about the language they heard. Results showed that their ability to categorize geminate consonants was “significantly enhanced by the simple knowledge that a consonant length distinction exists in the language” (p. 250).

While these results were promising, showing that instruction can enhance participants’ noticing, as in Schmidt (1990), we must remember that these participants were not actually learning Finnish, only listening for length. There was no communicative burden, search for meaning, or need to respond, all of which would impact classroom acquisition. For this reason, more research should be done involving participants who are actual learners of Finnish, such as Ylinen et al. (2005). This study examined L1 Russian speakers living in Finland, finding that most of those who had been there more than 5 years had developed at least some sensitivity to phonological quantity.

Little, if any, research has examined L1 American English speakers’ production of Finnish phonology generally or geminate consonants specifically. Bassetti (2017) found that native English speakers did not produce a length contrast when a word was spelled with a double letter. While English does not have true geminate consonants, there are words spelled with double consonants. Bassetti and Atkinson (2015) found that even experienced L1 Italian learners of English had difficulty ignoring orthography and producing a single short consonant. They lengthened consonants like the /tt/ in *kitty* 50% more than the /t/ in *city*. It is likely that English speakers beginning to learn Finnish, used to ignoring double letters in their L1, would have trouble producing them when required by Finnish orthography, causing communication breakdowns because this contrast is frequent and phonemic in Finnish. A replication of Porretta and Tucker (2014) with production by naïve learners would show if pointing out the short/long also assists there. In Kabak et al. (2011), L1 German speakers with no experience learning Italian were able to produce a length contrast when repeating Italian-like nonce words, suggesting that training by repeating words could be helpful for learners.

Presumably, North American learners of Finnish have been instructed to make geminates longer, but are they able to? With this language pairing, we have a situation in which American English speakers acquiring Finnish may have significant trouble both perceiving and producing geminates because they appear in contexts where they are not listening for them or used to producing them, but this area is vastly understudied, with little idea of how development might proceed. Al-Deaibes and Jarrah (2023) found that L1 American English beginning learners of Arabic, a language that has geminates like Finnish, were unable to produce the contrast, with an average ratio of 0.9:1. Advanced learners of Arabic, with three years of study instead of one like the beginning group, produced an average ratio of 1.4:1, while native speakers of Arabic had an average ratio of slightly more than double the singletons. Like Ylinen et al. (2005), whose participants had been in Finland more than five years, this result is suggestive of a long acquisitional path.

Research Question

Does the intervocalic singleton and geminate production of an L1 American English learner, who has completed one year of Finnish, show sensitivity to the geminate contrast?

I hypothesize that she will produce a contrast; however, considering the length of residence in Finland in Ylinen et al. (2005), it will not match the native speaker's contrast in terms of duration.

METHODS

Participant

This project was a case study, the first step in a larger research study, whose participant, whom we will call Aulikki, is a learner in the United States with Finnish heritage. Despite her background, she did not grow up hearing Finnish; but knew it was a family language, which sparked her desire to study it. Her L1 is English, and she has acquired near-native Spanish, at times living in a Spanish-speaking country. She knows some French. At the time of data collection, she had completed one academic year of university Finnish study, with class attendance five days per week, followed by three weeks at a summer immersion program in Finland. She was tested a few months after returning from Finland, and she was not at the time enrolled in another Finnish course. A learner with one year of Finnish was chosen for the study because, based on the literature review, I did not expect learners with less than one year of study to be able to produce the contrast.

Materials

Although the site no longer appears to exist, at the time of data collection, there was a webpage called *Helsinki Phonetics* that had examples of Finnish minimal pairs with and without intervocalic geminates. For example, one minimal pair was *suma* (jam) and *summa* (sum). All the tested minimal pairs can be seen in Table 1. Most of these words (14/16) had an accompanying voice file spoken by a female described as a native speaker on the site so that one

could hear how the minimal pairs sounded. Each word was accompanied by the English translation.

For this study, eight minimal pairs were embedded into a word list with 32 distractors for a 2-to-1 ratio of distractors to target items, for a total of 48 words. The word list was divided into three pages, with 16 words on each, and the target items randomly spread throughout (Appendix 1).

Procedure

Aulikki visited the soundproof recording booth and was asked to read the word list one word at a time at a normal pace. She was not aware of the focus of the study and did not mention noticing the minimal pairs embedded in the word list. Her production was recorded. Afterwards, the length of her intervocalic consonants was measured using Praat software. The intervocalic consonants from the audio files on the Helsinki Phonetics website were also measured, and Aulikki's production compared to that of the NS where possible. Two words with no NS audio file could not be measured and are marked on Table 1 with *n/a*.

RESULTS

The research question asked if the intervocalic geminate production of an L1 English learner with one year of Finnish study showed sensitivity to the geminate contrast. Table 1 shows the actual measurements in milliseconds by both Aulikki and the NS. The singleton measurements are on top and the geminates underneath.

We can see from the table that in 7 of 8 attempts, her geminates are longer than her singletons. After one year of study, she can produce geminates longer than singletons for the most part. However, we can see that differences exist between Aulikki's production and NS's. In 6/7 tokens, Aulikki's singletons are longer than NS's. In 2/7 tokens, NS's production is less than half of Aulikki's.

Table 1

Native and Non-native Singleton and Geminate Production in Milliseconds

Singleton	nupi	suma	kato	kisa	peni	palo	varas	pako
NS	84	171	95	100	n/a	56	46	83
Aulikki	101	141	181	152	97	124	152	143
Geminate	nuppi	summa	katto	kissa	penni	pallo	varras	pakko
NS	298	298	255	250	225	175	n/a	261
Aulikki	315	250	269	276	278	263	113	204

Table 2 considers the relationship between the two data points for each speaker. In other words, how much longer is the geminate than the singleton? This number was calculated by taking the geminate number of milliseconds and dividing it by the singleton number. For example, Aulikki produced an intervocalic geminate for the word *nuppi* with a length of 315 milliseconds. Divided by her singleton production for *nupi*, we get a result of 3.12, meaning that her geminate was just over three times as long as her singleton, while the NS produced a geminate that was 3.5 times as long. Expressing the results this way accounts for differences in speech rate, allowing us to directly compare Aulikki and NS. We can also calculate an overall ratio average, which for the NS is 2.79 and for Aulikki is 1.92.

Table 2

Geminate Length Expressed as a Multiple of Singleton Length

Sound	/pp/	/mm/	/tt/	/ss/	/nn/	/ll/	/rr/	/kk/
NS	3.55	1.74	2.68	2.50	n/a	3.13	n/a	3.14
Aulikki	3.12	1.77	1.49	1.82	2.87	2.12	0.74	1.43

We can see that NS's and Aulikki's contrasts for /mm/ and /pp/ are similar. However, Aulikki's difference in length for most contrasts is smaller than NS. For /k/ vs. /kk/, her number is less than half of what NS produced. Looking back at Table 1, we can see that NS produces a singleton that is much shorter than Aulikki's singleton and a geminate that is much longer, accounting for the difference in Table 2.

DISCUSSION

The results show that, as in Porretta and Tucker (2014) with the *Instruction* group, Aulikki knows that geminates are longer than singletons and mostly produces them that way. Because her singletons are longer than NS's, they could be misunderstood as geminates by native speakers. In particular, the contrasts for /l/ and /k/ could cause confusion, since her singletons are approaching the NS's geminate length. Aulikki's difficulty with /r/ vs. /rr/ might be due to interference from Spanish, another language she speaks.

This project has several limitations, the biggest being that the NS data is incomplete. For both, these data are just one snapshot of production. For example, Aulikki may not have trilled the /r/ in a different moment. Averaging production over several trials would yield more complete results. We do not know what she would produce in conversation, nor was her perception tested.

These data suggest that training in shortening singletons could be helpful for L2 learners of Finnish. Personal observations of first-year Finnish classes have shown that instructors tend to emphasize lengthening geminates by having students say them as long as possible in an exaggerated way, but rarely, if ever, instruct students to shorten singletons. This emphasis may cause learners to overcompensate by producing longer singletons too.

This study was a step towards beginning to understand the acquisition task faced by speakers of English learning Finnish. More work needs to be done to better understand the developmental path for acquisition of geminates and how to improve instruction, since both Ylinen (2005) and Al-Deaibes and Jarrah (2023) found that 3-5 years of study or residence in country were not enough to approximate a NS contrast. We need to determine if the contrast learners have acquired is functional (Darcy & Krüger, 2012). In other words, will NSs perceive the correct word produced without native-like ratios?

Future work will collect cross-sectional data, testing both perception and production with students enrolled in the first, second, and third years of university Finnish courses in a foreign language context, which would allow us to begin determining an L2 acquisitional path and attempt to answer many questions, such as: When can they hear and say the difference? Do perception and production develop in tandem, or do they develop autonomously, as Hattori and Iverson (2010) found for L1 Japanese perception and production of /r/ and /l/ in L2 English? Do the different geminates develop at the same pace? In other words, are some easier to produce and perceive than others?

Ideally, this work could also be longitudinal by testing participants at the beginning and end of a semester or year. In addition to this cross-sectional work, having a group of Finnish native speakers listen to the learners' production to see if they understood what the learners were trying to say would show if the acquired contrast was functional or not. In other words, ideally L2 production data could be used to see if Finnish native speakers understand what the learners said, and native speaker production data to test learners' perception.

Additionally, because it is so unusual cross-linguistically, research needs to be done on post-consonantal geminates like /rkk/. What is the native speaker range? How long do learners need before they can produce a post-consonantal geminate? Are there minimal pairs for this that affect communication?

Finnish consonant gradation means that some geminates become singletons under certain conditions. For example, when expressing possession, the Finnish name *Mikko* becomes *Miko-* to which the possessive suffix *-n* is added to form *Mikon*. What effect does this change have on acquisition? It may be more difficult for learners to produce the contrast when it can change back and forth.

Comparing Finnish acquisition to Japanese acquisition will help clarify if frequency helps learners, as it seems to with L1 acquisition. Since Finnish geminates appear twice as frequently, and Finnish infants can distinguish them faster, does that hold for L2 acquisition?

In conclusion, one year of Finnish instruction plus three weeks in an immersion setting were enough for Aulikki to produce a contrast in most pairs, but not a ratio similar to NS. Perception studies by naive learners showed that a lot of difficulty identifying consonantal geminates, as the mismatch discussed above implied they would. Additionally, the "short" exposure group in Ylinen et al. (2005) had lived in Finland less than 5 years, indicating that acquiring the contrast could take a long time even in a second language situation.

This case study suggests that instructors of languages with geminates should focus on training students to shorten singletons. The learners' task is not just acquiring certain sounds but a whole system of contrasts. Eventually, the goal is to understand this different phonological system on an intuitive level, putting sound and meaning together in a new way.

ACKNOWLEDGMENTS

Thank you to Aulikki for donating her time. Thank you to Kelly Berkson, whose feedback and support were invaluable. Thank you also to Amanda Huensch and to Isabelle Darcy for their helpful comments. Finally, thank you to the IU Second Language Studies Department for a Harry Gradman Fall 2023 Travel Award which funded my attendance at PSLLT.

ABOUT THE AUTHOR

Jeanne McGill is a Doctoral Candidate in Second Language Studies at Indiana University Bloomington. Her dissertation is entitled *Do Words Matter? How Lexical Input Influences German/English Bilinguals' Syntax in Beginning Swedish*, which more deeply investigates leading models of third language acquisition. She has doctoral minors in Linguistics and Finnish and an MA in German. Her main research interests are L2/L3 syntax and phonology, language policy and revitalization, and German and Nordic literature. She has been fascinated by Finnish for nearly ten years now.

REFERENCES

- Al-Deaibes, M., & Jarrah, M. (2023). Production of Arabic geminates by English speakers. *Journal of Psycholinguistic Research*, 52(6), 2877-2902.
- Aoyama, K. (2001). *A psycholinguistic perspective on Finnish and Japanese prosody: Perception, production and child acquisition of consonantal quantity distinctions*. Springer US.
- Bassetti, B. (2017). Orthography affects second language speech: Double letters and geminate production in English. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(11), 1835.
- Bassetti, B., & Atkinson, N. (2015). Effects of orthographic forms on pronunciation in experienced instructed second language learners. *Applied Psycholinguistics*, 36(01), 67-91.
- Daidone, D., Lidster, R., & Kruger, F. (2023). Free classification as a method for investigating the perception of nonnative sounds. *Studies in Second Language Acquisition*, 45(4), 1104-1130. doi:10.1017/S0272263123000050
- Darcy, I., & Krüger, F. (2012). Vowel perception and production in Turkish children acquiring L2 German. *Journal of Phonetics*, 40(4), 568-581.
- Davis, S. (2011). "[Geminates](#)". In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell Companion to Phonology*, (pp. 837-859). Wiley Blackwell.
- Dmitrieva, O. (2012). *Geminate typology and the perception of consonant duration*. (Doctoral dissertation, Stanford University).

- Doty, C. S., Idemaru, K., & Guion, S. G. (2007). Singleton and geminate stops in Finnish-acoustic correlates. *INTERSPEECH* (pp. 2737-2740).
- Engstrand, O., & Krull, D. (1994). Durational correlates of quantity in Swedish, Finnish and Estonian: Cross-language evidence for a theory of adaptive dispersion. *Phonetica*, 51 (1-3), 80-91.
- Hattori, K., & Iverson, P. (2010). Examination of the relationship between L2 perception and production: An investigation of English /r/-/l/ perception and production by adult Japanese speakers. *Proceedings of Second language studies: Acquisition, learning, education and technology*. Retrieved from <https://www.researchgate.net/profile/Kota-Hattori/publication/267810044>
- Helsinki Phonetics. Link now defunct. Accessed at: http://www.helsinki.fi/puhetieteet/projektit/Finnish_Phonetics/konsonanttiesim_eng.htm
- Kabak, B., Reckziegel, T., & Braun, B. (2011). Timing of second language singletons and geminates. *Proceedings of International Congress of Phonetic Sciences*. Retrieved from <https://kops.uni-konstanz.de/server/api/core/bitstreams/d3ee2762-377a-43a0-bbc0-22b421716795/content>
- Kunnari, S., Nakai, S., & Vihman, M. M. (2001). Cross-linguistic evidence for the acquisition of geminates. *Psychology of Language and Communication*, 5(2), 13-24.
- Maddieson, I. (1984). *Patterns of sounds*. Cambridge University Press.
- Porretta, V. J., & Tucker, B. V. (2014). Perception of non-native consonant length contrast: The role of attention in phonetic processing. *Second Language Research*, 31, 239–265.
- Richardson, U. (1998). *Familial dyslexia and sound duration in the quantity distinctions of Finnish infants and adults*. (Doctoral dissertation, University of Jyväskylä). Jyväskylä: Studia Philologica Jyväskyläensia.
- Savinainen-Makkonen, T. (2007). Geminate template: A model for first Finnish words. *First Language*, 27(4), 347-359.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics* 11, 129–58.
- Smith, A.O. (2020). 50,000 people sign up for Duolingo's Finnish course within first four days. *Helsinki Times*. Accessed at: <https://www.helsinkitimes.fi/finland/news-in-brief/17812-50-000-people-sign-up-for-duolingo-s-finnish-course-within-first-four-days.html>
- Ylinen, S., Shestakova, A., Alku, P., & Huotilainen, M. (2005). The perception of phonological quantity based on durational cues by native speakers, second-language users and nonspeakers of Finnish. *Language and Speech*, 48(3), 313-338.

APPENDIX 1

Word List

1. koe
2. sorsa
3. lasku
4. lukio
5. pako
6. mainos
7. vyö
8. pää
9. puu

10. nupi
11. peni
12. tosin
13. talo
14. pensus
15. pallo
16. huivi
17. joki
18. pakko
19. tukka
20. maa
21. katto
22. käsi
23. vasta
24. olo
25. paha
26. varras
27. otsa
28. juoksu
29. nuppi
30. sormi
31. varas
32. lomake
33. kissa
34. lento
35. suma
36. kaula
37. palo
38. retki
39. hame
40. passi
41. kisa
42. penni
43. tammi
44. orava
45. luento
46. kato
47. summa
48. apu