CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

Finnish inserted vowels: a case of phonologized excrescence

Robin Karlin

University of Wisconsin-Madison, Waisman Center, Madison, WI, 53705, USA
Email for correspondence: rkarlin@wisc.edu

Abstract

In this paper, I examine a case of vowel insertion found in Savo and Pohjanmaa dialects of Finnish that is typically called "epenthesis", but which demonstrates characteristics of both phonetic excrescence and phonological epenthesis. Based on a phonological analysis paired with an acoustic corpus study, I argue that Finnish vowel insertion is the mixed result of phonetic excrescence and the phonologization of these vowels, and is related to second-mora lengthening, another dialectal phenomenon. I propose a gestural model of second-mora lengthening that would generate vowel insertion in its original phonetic state. The link to second-mora lengthening provides a unified account that addresses both the dialectal and phonological distribution of the phenomenon, which have not been linked in previous literature.

1. Introduction

In this paper, I examine a case of vowel insertion found in Savo and Pohjanmaa dialects of Finnish that has typically been analyzed as a phonological repair, but which demonstrates characteristics of both phonetic excrescence and phonological epenthesis. Using both acoustic data and a phonological analysis of the distribution, I argue that Finnish vowel insertion originated as a phonetic intrusion, but then became phonologized over time. I follow Hall (2006) in assuming that excrescent vowels are the result of gestural underlap, and argue that the original gestural underlap was caused by second-mora lengthening, another phenomenon present in these dialects. Finally, I describe a gestural model of second-mora lengthening that predicts the original excrescent state of this vowel insertion, and discuss possible influences on the path to phonologization. This paper adds to the debate on these inserted vowels by combining phonological analysis with phonetic analysis, using acoustic data from a publicly available corpus, and is the first to couch the analysis in Articulatory Phonology.
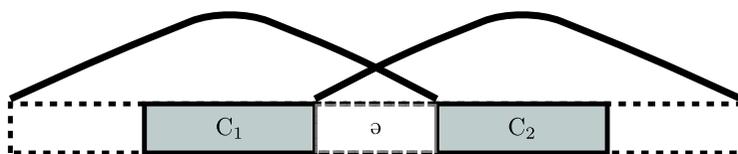
CrossMark

Figure 1. Schematic representations of gestural underlap between two consonant gestures, producing an excrescent vowel. Each gesture is represented by one arc and one rectangle; dotted lines on the rectangle indicate that the gesture is active but not producing sufficient closure to produce a consonant. Time is on the horizontal axis; for simplicity, the full target achievement of each gesture is at the peak of each gestural trajectory. Note the gap between peaks.

1.1 Inserted vowels: epenthesis vs. excrescence

There are two types of inserted (i.e., not underlying) vowels: those that are produced as the result of phonetic intrusion (excrescent vowels), and those that are inserted as part of a phonological repair (epenthetic vowels) (Hall 2006; Fougeron and Ridouane 2008; Hall 2011; Plug et al. 2019). These two types of vowels have different phonological and phonetic characteristics: for example, on the phonology side, excrescent vowels are frequently invisible to phonological constraints and processes such as word minimality and stress assignment, while epenthetic vowels participate in these same processes; on the phonetic side, excrescent vowels are more likely to be schwa than epenthetic vowels, and are also more variably present.

Previous attempts to account for these differences have frequently hypothesized that the differential phonological behavior of these two vowel types is due to insertion at different stages of the derivation (Levin 1987; Vaux 2003; Hall 2006). However, invoking derivational stages does not predict both the phonological behavior and the phonetic characteristics. Hall (2006) thus proposes that excrescent vowels are the result of gestural underlap—i.e., the vowel occurs simply because there is not sufficient constriction in the mouth to form a consonant, not because there was a deliberate gesture meant to produce this vowel (illustrated in Figure 1). The lack of deliberate gesture accounts for excrescent vowels' invisibility to phonological processes such as word minimality constraints, stress assignment, and syllabification: there is no actual vowel gesture, so it cannot be targeted or recruited. In addition, the lack of deliberate vowel gesture correctly predicts that these vowels are variably inserted, are of shorter duration, and more likely to be schwa and/or influenced by surrounding (consonant) gestures.

In contrast, vowels inserted via phonological epenthesis are associated with a full gesture of their own. They are not necessarily part of the underlying representation, but are predictably inserted as part of the phonological grammar. As such, they are composed of a deliberate gesture that can be recruited by phonological processes. These vowels are thus frequently inserted as the solution to a phonotactic violation: for example, a vowel may be inserted to satisfy word minimality, or to repair marked consonant sequences. The existence of a deliberate gesture also correctly predicts that this type of inserted vowel is consistently inserted, is comparable in duration to other phonologically present vowels, and is more likely to have a set vowel quality.

Table 1. Diagnostic criteria, as summarized by Hall (2006)

	Excrescent vowels	Epenthetic vowels
Environment	Between heterorganic consonants only	May or may not be between heterorganic consonants
Markedness	Not a repair	Repairs marked structures that are avoided elsewhere in the language
Quality I	Often schwa; influenced by neighboring consonants	Not necessarily schwa, may be a fixed vowel; not influenced by consonant
Quality II	If copy of nearby vowel, only across sonorants or gutturals	No restrictions on copying
Optionality	Optional; influenced by speech rate	Obligatory; presence not dependent on speech rate
Duration	Variable duration	Consistent duration

Hall's (2006) gestural account thus provides a single source for the phonetic and distributional differences between these two types of inserted vowels. She thus proposed a set of diagnostic criteria to apply when analyzing a language, including the environments in which insertion occurs, the quality of the inserted vowel, and the consistency with which the vowel appears. These diagnostics are summarized in Table 1 (adapted from Hall 2006).

It should be noted, however, that vowel insertion is not a static phenomenon; in particular, as was noted by Hall (2006) and as I show in this paper for Finnish Vowel Insertion (FVI), excrescent vowels may phonologize over time, and thus acquire apparent mixed characteristics. Thus, any analysis of vowel insertion phenomena must address multiple characteristics rather than relying on a single criterion as a litmus test. In this paper, I accomplish this by combining a phonological analysis of vowel distribution (addressing the "Environment" and "Markedness" criteria), the phonetic analysis of acoustic corpus data (the remaining criteria), and an additional consideration of the inserted vowels' interactions with morphological and phonological processes. In the section that follows, I will present the previously reported characteristics of FVI, including triggering environment and quality of the inserted vowel. Based on the diagnostics of Hall (2006), I will argue that the distributional characteristics of inserted vowels align largely with excrescent (phonetic) vowels, not epenthetic (phonological) vowels.

1.2 Distribution diagnostics

This paper focuses on vowel insertion as exhibited in two major dialect groups: the Savo dialects, which are the major dialect group of the Eastern branch of Finnish (*itämurteet*), and the Keski- and Pohjois-Pohjanmaa dialects (henceforth Pohjanmaa[1]), which are spoken in the region adjacent to the Savo dialects but form a subgroup of the Western branch of Finnish (*länsimurteet*) (Palander 2011). In these dialects, words that have the form $(C)VC_2.C_3V(X)$ in Standard Finnish are realized as $(C)V.C_2V_i.C_3V(X)$, where the subscript i indicates insertion (Kettunen 1940; Harms 1976; Suomi 1990; Harrikari 1999). For example, the word

silmä 'eye' is realized as [sil**i**mä],[2] and the word *halpa* 'cheap' is realized as [hal**a**pa]. This insertion is always between a coda consonant and a syllable onset—native Finnish words do not have onset consonant clusters; all words with initial clusters are relatively recent borrowings. More specifically, insertion occurs when C_2 is in second mora position, a fact whose relevance will come to light in due course. Thus, the apparent single exception to this coda-onset split, where C_3 is a geminate (as in *helppo* 'easy' > [hel**e**ppo]), indeed still follows the second mora position generalization.

Previous analyses have treated this inserted vowel as a phonological repair. For example, Harrikari (1999) argued that codas are overall marked in Finnish, and the epenthetic vowel serves to repair this marked structure; Suomi (1990) more specifically argued that only sequences that are not present in $(C)VVC_2.C_3V(X)$ words are repaired. However, the pattern of triggering vs. non-triggering sequences in fact aligns closely with Hall's (2006) characterization of excrescent vowels. The first two non-triggering sequence types are listed and exemplified below:

(1) **Heterorganicity requirement**: FVI only occurs in heterorganic sequences, as in (1a). This requirement excludes geminates (1b), as well as other homorganic sequences (1c).

  a. silmä   > silimä   *eye*
  b. malli   > *mal$V_i$li   *model*
  c. ilta   > *il$V_i$ta   *evening*

(2) **Position requirement**: The triggering sequence must be between the first and second syllable, as in (2a). That is, the same consonant sequence later in the word does not trigger insertion, as in (2b).

  a. silmä   > silimä   *eye*
  b. kuvitelma > *kuvitel$V_i$ma *fantasy*

According to Hall (2006), excrescent vowels tend to occur in heterorganic consonant sequences and do not necessarily occur in more marked structures (i.e., structures that are cross-linguistically rare); in contrast, epenthetic vowels repair marked structures that are avoided elsewhere in the language, and do not occur more frequently in heterorganic sequences. The requirements for vowel insertion in Finnish show that these vowels clearly aligns with excrescence: only heterorganic sequences trigger insertion, and the same sequence later in the word is not 'repaired.' Harrikari (1999) noted this exception in the word *vadelma*, which does not insert to become *[vadelema]. However, she attributes this exception to metrical structure, arguing that inserted vowels cannot be the head of a foot (i.e., *(va.de)(le$_i$.ma)). In a word like *kuvitelma*, this would not be an issue, as the inserted vowel would be in the unstressed syllable of the foot (i.e., (ku.vi)(te.le$_i$)ma). For this paper I have consulted with nine native speakers of inserting dialects, who all agree that insertion would not happen in *kuvitelma* or other similar words where footing would not be a barrier to insertion. Thus, vowel insertion is only triggered after the first syllable (i.e., when C_2 is the second mora in the word), and is not triggered by a general markedness condition, e.g., *lm.

Although Hall ([2006](#)) does not include voicing restrictions explicitly in the diagnostic criteria, she does discuss examples of excrescence where the underlap in voiceless sequences are transcribed as aspiration, while the underlap in voiced sequences is transcribed as a schwa (see Gafos [2002](#) for Sierra Popoluca). FVI shows similar patterning:

(3) **Voicing requirement**: C_2 must be voiced, as in (3a) (compare 3c); C_3 can but need not be voiced for insertion to occur. The voicing can be phonetic, rather than underlyingly, as in (3b), where /h/ can receive voicing between sonorants (i.e., after a vowel, before a sonorant consonant; compare 3d).

    a.   alkaa   > alakaa   *begin*.3SG[3]
    b.   ahma   > ahama   *wolverine*
    c.   pitkä   > *pit$\mathbf{V}_i$kä   *long*
    d.   ahkera   > *ah$\mathbf{V}_i$kera *hard-working*

This restriction is evidence for excrescence because it indicates that the inserted vowel is not fully specified or deliberate—it is simply an interval without sufficient constriction. When C_2 is voiced, the acoustic result of this underlap is some schwa-like vowel; however, in the cases where C_2 is voiceless, the underlapped portion would be voiceless as well, as there is no deliberate set of vocalic gestures to create voicing (i.e., there is no glottal adduction gesture). Note that in Finnish this appears to be a phonetic constraint, rather than a phonological constraint, as demonstrated by the apparent exception for /hC/ sequences—only /hC/ sequences where the second C is a sonorant exhibit insertion.

This exception is likely due to the realization of voicing on /h/: intervocalically, it is typically a fully voiced [ɦ] (Suomi et al. [2008](#)), and /h/ can receive similar full voicing in other sonorant contexts, such as before a sonorant consonant (see example from a production of *rihlapyssyseppä* 'smith for rifled guns' in Figure [2](#)). This kind of passive voicing does not occur on other fricatives or stops in Finnish, which excludes tokens like *ritva* 'bough' from insertion. This exceptional behavior of /h/ likely comes from its affiliation with the glides: /h, v, j/ are the only segments in Finnish that do not underlyingly occur as geminates.[4] As such, it is likely that /h/ does not have a deliberate glottal spread gesture and instead is influenced by surrounding segments' voicing gestures, which can be timed early relative to their respective stricture gestures. If the glottal spreading associated with /h/ ends while some voicing gesture is active (whether from the preceding vowel or from the following sonorant consonant), the result would be a modal vowel—though note also that a voiced [ɦ] is itself extremely similar to a vowel, without oral constriction and with formant structure. Thus, exception 1.2 also falls neatly into the excrescence account.

Finally, /rC/ sequences are an exceptional case, which at first blush seem to be an exception to the generalizations of heterorganicity and voicing.

(4) **/rC/ sequences**: If C_2 is an /r/, insertion may or may not be triggered.

    a.   sormi   > ʔsor**o**mi   *finger*
    b.   kirkolla > ʔkir**i**kolla *church*.ADESS
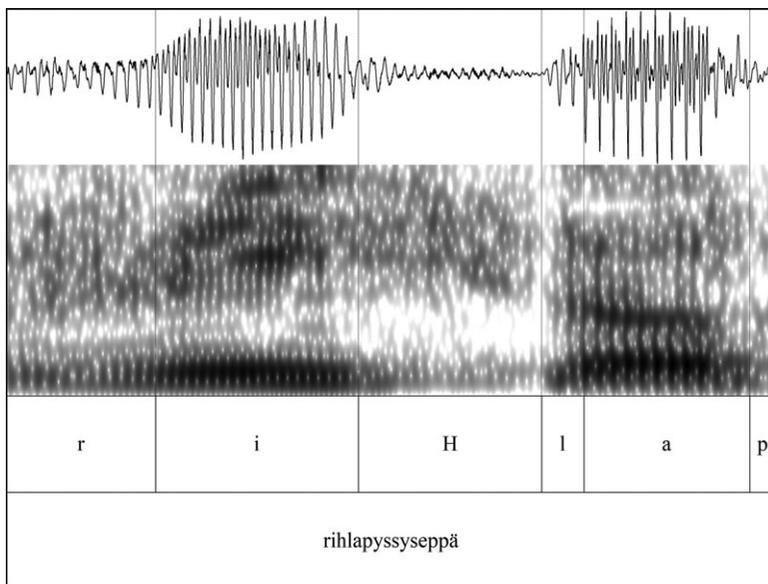   c.   kirja   > ʔkir**i**ja   *book*

Figure 2. A production of *rihlapyssyseppä* 'smith for rifled guns' from a speaker from Varpaisjärvi (Savo), focusing on *rihla*. Note that there is only a very faint voicing bar in the spectrogram, but periodicity can be seen throughout the [ɦ] (labeled H), indicating full voicing in this environment. Note also that this speaker tends to not insert in /hC/ environments and has not in this particular token.

These sequences have been debated in the previous literature. Suomi (1990) argued that insertion occurs in sequences that do not exist when V_1 is a long vowel; since there are words like *käärme* 'worm', insertion must *not* occur in /rC/ sequences. In contrast, Harrikari (2003) argued that all codas were marked unless they were homorganic with the next consonant, and thus insertion *must* occur in heterorganic /rC/ sequences. As their analyses are both based on FVI being a phonological repair, rather than a phonetic artefact, /rC/ sequences are a pivotal case for each.

Treating these vowels as excrescent (as indicated by the other distributional characteristics) would resolve this debate. In coda position, the Finnish /r/ is a trill (Suomi et al. 2008). Trills are inherently somewhat variable; changes in airflow can lead to early termination of the trill, in which case there would be an excrescent vowel as the tongue stops vibrating and instead remains in approximant- or vowel-like position. The variable nature of the trilled /r/ can thus explain the debate between Suomi (2000) and Harrikari (1999): sometimes there is a gap between consonant constrictions, and sometimes there is not. According to Hall (2006), variable appearance is a characteristic of excrescent vowels, rather than epenthetic vowels.

An excrescent account would also address Suomi's (2000) argument that speakers of these dialects do not report hearing these vowels in /rC/ contexts—if the vowels are purely phonetic artefacts of gestural underlap, one might expect speakers to not be aware of them. In fact, there is varied evidence that Finnish speakers tend to not perceive these vowels at all. For example, Harms (1976: 74) describes these vowels as "purely transitional in nature" and notes that speakers do not perceive them as forming syllable nuclei:

[mɛləkein] (melkein) 'almost' has essentially the same vowel qualities ([ɛ, ə, ei]) and relative durations as the English verb delegate—[dɛləgeit]. From a descriptive phonetic point of view, the Finnish [intrusive] schwa and the English reduced-vowel schwa represent very nearly identical classes of vowel sounds; i.e., they vary over a wide central area, with their range of variation conditioned by the preceding and following segments. But here the similarity ends. The schwa in the above Finnish forms is purely transitional in nature. *Speakers perceive these forms as containing only two syllables, not three* [emphasis added].

Wiik (1965: 28) also notes that many Finnish speakers produce this inserted vowel without being aware of doing so, and also carry an expectation of some sort of schwa-like vowel in this prosodic position with them when learning English:

Many Finnish pronounce a short schwa-vocoid between /l/ and /p/ as in /kalpa/ = "sword"... without being aware of the existence of this vocoid.... When these Finns hear the words [skæləpɪŋ] = "scalloping" and [skælpɪŋ] = "scalping" in English, they segment both utterances into 7 segments, and thus they do not hear the difference between the utterances.

Thus, while there may be some speakers that are aware of this vowel insertion, self-identification of vowel production is not a good heuristic per se for determining whether or not these vowels are truly present.

1.3 Quality diagnostics

Hall (2006) also discusses vowel quality characteristics in her diagnostics for excrescent vs. epenthetic interpretations. In this case, too, FVI behaves as phonetic excrescence. In his dialect atlas, Kettunen (1940) distinguished between two qualities of FVI. In the "default" case of insertion (exhibited in Pohjanmaa and northern Savo dialects), the inserted V_i is a copy of V_1. This is exemplified in (5).

(5) a.   silmä   > sil**i**mä   *eye*
     b.   kulma   > kul**u**ma   *corner*
     c.   jalka   > jal**a**ka   *leg*

Hall (2006) shows that excrescent vowels can be copies of an adjacent vowel, as long as the intervening consonant is a sonorant or a guttural. Recall that, in order for FVI to occur, C_2 must be voiced. In Finnish, there are no voiced obstruents; as such, if C_2 is voiced, it is a sonorant. The one minor exception to this is /h/, which triggers insertion when next to a voiced C_3. In this case, Hall's (2006) requirements for excrescent copying still apply, as /h/ is a guttural (and, as previously mentioned, perhaps phonologically a glide). Thus, the first quality diagnostic indicates that FVI is excrescent.

Other regional variants of FVI also point to an excrescent origin. In southern Savo dialects, the vowel is described as having an "intermediate" quality, or a quality between the two surrounding vowels (Kettunen 1940). This is illustrated in (6).

(6)   a.   silmä   > sil**e**mä   *eye*
    b.   kulma   > kul**o**ma   *corner*
    c.   kylmä   > kyl**ö**mä   *cold*

These "intermediate" vowels fall under the excrescent category. According to Hall (2006), excrescent vowels are frequently schwa-like or of "intermediate quality", while epenthetic vowels are a fixed vowel (either schwa or not). The "intermediate quality" of excrescent vowels is once again due to the nature of gestural underlap: the gap in consonantal stricture occurs while the tongue is moving from the preceding vowel to the following vowel, and resulting in an intermediately colored vowel. Thus, one might expect the mid front vowel [e] when moving from the high front vowel [i] to the low front vowel [ä], or the mid back vowel [o] when moving from the high back vowel [u] to the low back vowel [a].

Hall's (2006) final vowel quality diagnostic addresses the influence of surrounding consonants. Specifically, she notes that excrescent vowels may be affected by surrounding consonants, but epenthetic vowels are fixed and unaffected. In most cases of FVI, the vowel is not influenced by the surrounding consonants; however, when C_3 is a /j/, the inserted vowel is [i], regardless of the quality of the two adjacent vowels. This is illustrated in (7) (note the non-high vowels both preceding and following).

(7).   a.   paljon   > pal**i**jon   *a lot*
     b.   kalja   > kal**i**ja   *beer*

Once again, this quality diagnostic points towards an excrescent origin, rather than phonological epenthesis. In these words, the vowel quality is influenced by the upcoming /j/, which is articulatorily very similar to /i/. Referring again to the gestural underlap account, a not-quite-/j/ would resemble [i]. Thus, all diagnostics for vowel quality indicate that FVI is an excrescent phenomenon. Further details on the vowel space of inserted vowels relative to the vowel space of underlying vowels will be provided in Section 2.

1.4 Duration diagnostics

The last set of characteristics discussed by Hall (2006) examines the duration of the inserted vowels. Excrescent vowels are variably produced, and influenced by parameters such as speech rate; their duration is also inconsistent, with a tendency to be quite short. In contrast, epenthetic vowels are consistently produced, regardless of speech rate, with very little variability in duration. In addition to the full copied and intermediate quality vowels, Kettunen (1940) indicates a third type of insertion in some northern Pohjanmaa dialects, where the inserted vowel is written in parenthetical superscript, e.g. [jal[(a)]ka] for *jalka*; this suggests that in that particular pocket of Pohjanmaa, the vowel is short and optional. Conversely, this also suggests that FVI in other regions is consistent and of comparable duration to underlying vowels. However, little explicit description of either the consistency or the duration of these inserted vowels exists in the literature. For this reason, in Section 2, I present

the results of an acoustic corpus study that examines the phonetic characteristics of the inserted vowels, focusing on the characteristics detailed above.

2. Dialect corpus: acoustic study

The current study addresses the quantifiable aspects of FVI as relevant to determining the phonological status of the inserted vowels: frequency of occurrence, duration, and quality. In this section, I first describe the data corpus and methods of analysis, and then present the quantitative results, showing that in both dialects, there is some degree of phonologization of these vowels, but with cross-dialect and cross-sequence variation.

2.1 Data corpus

The majority of the acoustic data for this study comes from the *Suomen Murrekirja* ('dialect book') (Lyytikäinen et al. 2013), an online repository of Finnish recordings. The Murrekirja is a digital corpus of about 500 recordings that are tagged with the city or town of origin.[5] Each of these locations is also labeled with a dialect area. The recordings consist of one person talking for between two and five minutes, usually telling some kind of story—for example, about how life was when they were young, or about a bear hunter, or a wedding. The result is casual and dialectal speech, which is optimal for finding vowel insertion, a non-standard feature. The birth years of the speakers included in this study ranged from 1874 to 1905, with age at the time of recording ranging from 60 to 93. Gender was not indicated in the corpus, but was discernible from speaker names and the recordings (see Table 1 for more detailed information).

Data was collected from the locations listed in Table 1, where the database only includes one recording per location. The locations are marked in Figure 3, where Savo locations are marked with a circle and Pohjanmaa locations are marked with a diamond. Locations were selected from the list based on multiple criteria, including audio quality and presence of $CVC_2.C_3V$ words; I also made an effort to include locations from different regions of the dialect area, so as to not effectively be analyzing one small geographic region's tendencies. Data from seven locations from each region was included, in total 14 recordings from the dialect corpus.

In addition to this corpus, a native speaker of the Oulu dialect provided me with a recording of them reading a text that they wrote to showcase Oulu inserted vowels (i.e., with the deliberate inclusion of words that show insertion). The text and recording were created without my input and before the conceptualization of this paper, and as such were not influenced by the hypotheses of this study. With this recording, the Pohjanmaa dialect has eight recordings, for a total of 15 recordings included in the dataset.

Two types of words were labeled for use in the quantitative analysis. The first, which I will refer to as triggering tokens, included words with triggering sequences in the correct prosodic position, such as *lehmä* 'cow' and *kylmä* 'cold'. Triggering consonant sequences beyond the second mora were only included if they were in the second mora position of the second element of a compound, such as *nykypolvelle* [nyky+polvelle] 'today's generation.ALL'—i.e., a word like *vadelma* 'raspberry' would be excluded. In addition, /rC/ sequences were excluded from this category
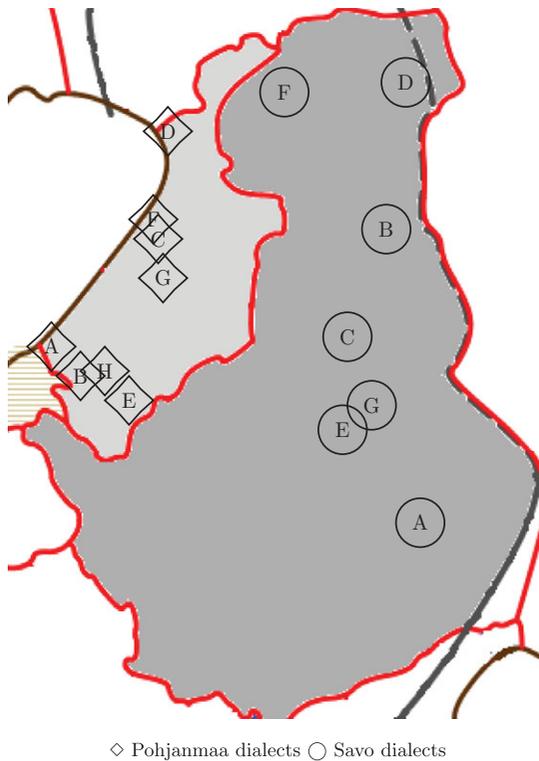
◇ Pohjanmaa dialects ◯ Savo dialects

Figure 3. Map of Finnish dialects used in this study, adapted from Lyytikäinen et al. (2013) and cut to focus on relevant dialect regions (Pohjanmaa dialects in light gray region; Savo dialects in dark gray region). Location tags correspond to the letters in Table 1.

due to their controversial status, and instead will be discussed qualitatively. Although the exact consonant sequences could not be controlled for in the corpus, the samples are relatively balanced across dialect. Both dialects had similar proportions of different C_2C_3 sequences: /hC/ (Savo 18, Pohjanmaa 15); /lC/ (Savo 42, Pohjanmaa 36); and /nh/ (Savo 4, Pohjanmaa 7).

The second type of words (henceforth referred to as baseline tokens) were of the shape $CV_1CV_2(X)$. These were used to set a baseline of the expected qualities and durations of uncontroversially phonological vowels in V_1 and V_2 position. As V_1 and V_2 of a single word are directly compared, words with /i/ were only included if both V_1 and V_2 were /i/, such as in the word *niminen* 'so-named'. This exclusion was made in order to avoid effects of intrinsic vowel duration muddying the data, as Wiik (1965) noted that /i/ was significantly shorter than even other high vowels.

The full summary of the resulting corpus is presented in Table 2, with details on the number of tokens and lemmas in each location. Lemmas could be split based on trigger vs. baseline status; for example the word *ulos* 'outside.LAT' was counted as a lemma for the baseline tokens, while a related form *ulkona* 'outside.ESS' was counted as a separate lemma for the triggering tokens. Additional forms of *ulko-* were subsumed under the appropriate lemma—e.g. *ulkopuolella* 'on the outside' was counted as a member of the *ulkona* lemma. In all, the Savo dialect had 64 tokens of

Table 2. Locations used in the acoustic corpus study, with the number of tokens (T) and lemmas (L) for the Triggering (TR) and Baseline (BL) words. Gender, year of birth (YOB), and age at time of recording are given for each speaker. The asterisk indicates the recording that was provided independently of the corpus

	○ **Savo**				TR		BL	
	Location	Gender	YOB	Age	T	L	T	L
A	Heinävesi	F	1894	82	13	7	19	13
B	Hyrynsalmi	M	1886	79	6	4	3	2
C	Kajaani	M	1890	75	13	9	10	10
D	Kuusamo	M	1892	68	5	5	11	10
E	Maaninka	F	1896	80	10	6	2	3
F	Ranua	M	1890	71	8	5	6	6
G	Varpaisjärvi	F	1887	77	9	6	4	4
Total					64	29	55	42
	◇ **Pohjanmaa**				TR		BL	
	Location	Gender	YOB	Age	T	L	T	L
A	Himanka	M	1885	78	3	3	2	2
B	Kannus	F	1905	77	6	3	8	4
C	Kempele	F	1896	86	12	8	11	8
D	Kuivaniemi	M	1888	76	7	4	11	8
E	Lestijärvi	F	1898	80	6	6	18	11
F	Oulu*	F	1988	23	11	3	2	2
G	Rantsila	M	1902	60	7	4	8	4
H	Sievi	M	1874	93	8	7	5	5
Total					60	27	65	33

29 triggering lemmas and 48 tokens of 42 baseline lemmas; the Pohjanmaa dialect had 60 tokens of 27 triggering lemmas and 65 tokens of 33 baseline lemmas.[6] In total, the corpus had 124 tokens of 46 triggering lemmas, and 120 tokens of 67 baseline lemmas. Thus, while the corpus is small and may present some problems of generalizability, it does contain a fairly wide range of different words as well as some repetitions of those words to account for variability between productions and forms.

2.2 Labeling and processing

As the Murrekirja does not include transcripts of the sound files, the help of native speakers of Finnish was enlisted to transcribe the recordings. The resulting transcripts were used to verify the presence of triggering and baseline tokens. Based on these transcripts, sound files were segmented using Praat (Boersma and

Weenink 2017); all triggering tokens were included unless sound quality issues prevented quantitative analysis (e.g., phrase-final devoicing of an entire word, low signal to noise ratio making segmentation impossible). Triggering tokens were judged to have a vowel or not with both auditory and visual (referencing both the waveform and the spectrogram) confirmation. If present, inserted vowels were marked in an interval separate from C_2; if no vowel was judged to be present, only the consonant was marked. In order to not bias the marking of the length of the inserted vowel, segmentation of each sound file was done without knowing from which region the recording originated.

The resulting TextGrids were then processed in R (R Core Team 2019), using a script that gathered data on the identity and duration of each segment, using the package rPraat (Bořil and Skarnitzl 2016). This data structure was then fed to a Praat script written by the author to obtain the vowel quality of V_1 and V_i/V_2. Data analysis was carried out in R. Linear mixed effects models were implemented with the package lme4 (Bates et al. 2014), with speaker as a random effect. These models were built incrementally and compared with likelihood ratio tests.

2.3 Measures and hypotheses

As in Section 1, in presenting the results of the acoustic study I will focus on the diagnostic criteria for inserted vowels provided by Hall (2006). Three of the diagnostics are quantifiable in this data:

1. **Frequency of insertion in triggering sequences**. Is a vowel inserted consistently (indicating phonological status), or does its existence vary (indicating phonetic status)?
2. **Inserted vowel duration**. For this analysis, I compare the duration ratio between V_i and V_1 (i.e., divide V_i by V_1) in triggering tokens to the ratio between V_2 and V_1 (i.e., divide V_2 by V_1) in baseline tokens. Are inserted vowels as long as underlying V_2 (indicating phonological status), or are they shorter and/or more variable (indicating phonetic status)?
3. **Inserted vowel quality**. For this analysis, I again compare V_i to V_1 in triggering tokens, and V_2 to V_1 in baseline tokens that have matching vowel qualities. Is the Euclidean distance between V_i and V_1 on the one hand and V_2 and V_1 on the other comparable (indicating phonological status), or is the quality of V_i less predictable (indicating phonetic status)?

2.4 Consistency of insertion: mixed results

In this section, I examine the rates of insertion in triggering tokens, where any vocalic element is considered positive for insertion, regardless of duration. There is cross-dialect variation in the consistency of insertion, which indicates different degrees of phonologization in each dialect. Insertion is much more consistent in Pohjanmaa dialects (59 of 60 triggering tokens showed insertion) than in Savo dialects (47 out of 64 tokens showed insertion). This difference is statistically significant (Pearson's $\chi^2 = 12.14$, p $= 0.0002$). Insertion rates for each dialect are illustrated in Figure 4.
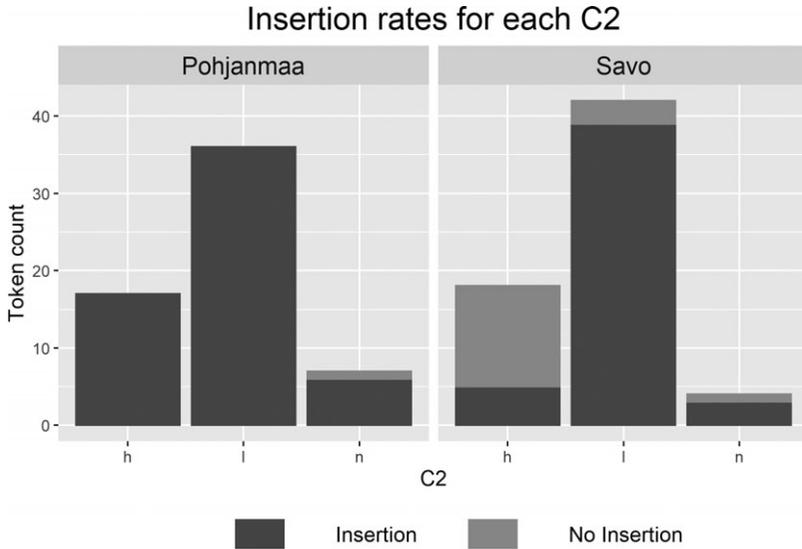
Figure 4. Summary of triggering C_2C_3 sequences in the data, and the frequency of vowel insertion in each. C_2C_3 sequences are collapsed into three categories: /hC/, /lC/, and /nC/ (all /nh/).

In the Pohjanmaa dialect, the single token that failed to insert was a token of the word *vanha* 'old', produced by the speaker from Lestijärvi. Although Kettunen (1940) specifically separates the word *vanha* (and other /nh/ sequences) from the main vowel insertion map, I included it in the "triggering" types of tokens because the area indicated on the map was roughly coextensive with the more general types of vowel insertion, and the /nh/ sequence follows the phonological requirements for insertion. In the corpus data, there are an additional four tokens of *vanha* (or some derived form); all exhibited insertion. As the Lestijärvi speaker did not produce this word or consonant sequence again, it is impossible to tell if this is an issue of variability within one speaker, or if this sequence simply never triggers insertion for them.

In contrast, almost all Savo locations had at least one triggering sequence fail to insert. The major locus of variability is in the /hC/ sequences: out of 18 /hC/ tokens, 13 (72.2%) did not show insertion. Five of these tokens were produced by the Varpaisjärvi speaker, who appears to never insert in /hC/ sequences—out of five words with an /hC/ sequence, none showed insertion. However, it does not seem to be the case that /hC/ sequences just fail to trigger insertion in some Savo dialects, as the speaker from Kajaani sometimes inserted in /hC/ sequences and sometimes did not. This variability was present even within the same word, illustrated in Figure 5. In this figure are three instances of the word *kahvi* 'coffee': the first (Figure 5a), has a vocalic portion with strong formant structure between the /h/ and /v/; the second (Figure 5b) has a vocalic portion between the /h/ and /v/ with less clear formant structure; and finally, the third (Figure 5c) has no such vocalic element, instead going straight from the /h/ to the /v/.

There was also some variability in /lC/ sequences, though all tokens that failed to insert were from the speaker from Ranua. In this case, there was some variability
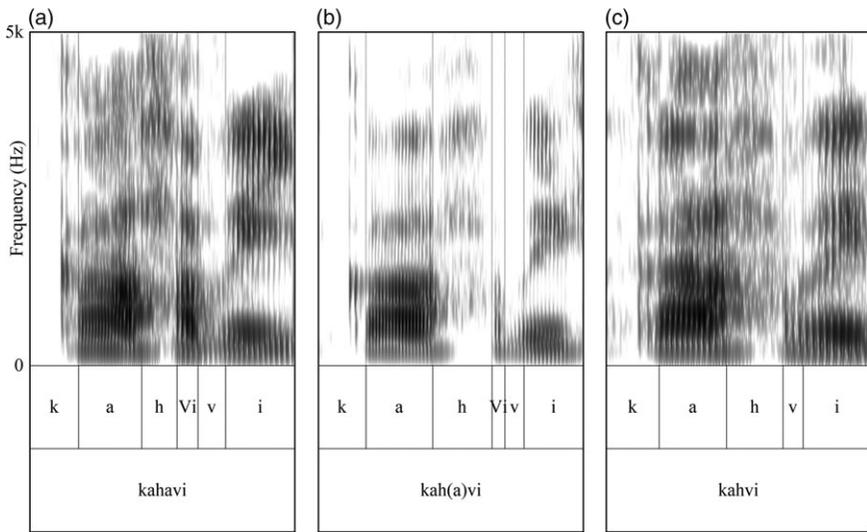
Figure 5. Three examples of *kahvi* 'coffee' by Kajaani (Savo) speaker. Inserted vowels marked with Vi.

in /lk/ sequences; some tokens of the same words showed insertion, while others did not. This is illustrated in Figure 6. In this figure, there are two instances of the word *ulkona* 'outside.ESS', accompanied by another word with the same consonant sequence, *jalkaa* 'leg.PART' (Figure 6a). In the first instance of *ulkona* (Figure 6b), there is a clear—if relatively short—vocalic element after the /l/; in the second (Figure 6c), there is no inserted vowel. The example of *jalkaa* (labeled "jalakaa") provides a good comparison for the previous two tokens: the inserted vowel is clearly present and quite long in comparison to V_1. Thus, although there is some variability in /lC/ clusters as well, it is limited to one speaker and clearly a case of token-to-token variability, rather than overall failure to trigger.

Of all C_2C_3 sequences, those involving /h/ would be the most likely to lag behind in phonologization. As described in Section 1, FVI only occurs in C_2C_3 sequences where C_2 has phonetic voicing. In sequences like /hm/ and /hv/, the /h/ frequently receives some voicing by being surrounded by sonorants, which provides an environment for FVI. However, a voiced [ɦ] is very vowel-like; Ladefoged and Maddieson (1996) noted that it has been described as a breathy-voiced version of a vowel, and Keating (1988) observed that the vocal tract during a glottal fricative simply takes the shape of whatever segments are around it. Thus, even if an excrescent vowel occurs after the voiced [ɦ], the articulatory and acoustic similarity between the [ɦ] and vowel would make it more difficult to perceive the vowel as a separate segment. This, in turn, would impede phonologization: unlike in, for example, /lC/ sequences, where insertion would have been irregular but perception consistent, /hC/ sequences would have shown variability both in insertion and in perception.

In sum, the two dialects show variation in the consistency of insertion, with an interaction between dialect and C_2. Pohjanmaa shows nearly 100% insertion, which is a trait of phonological epenthesis. Savo insertion is more variable, indicating
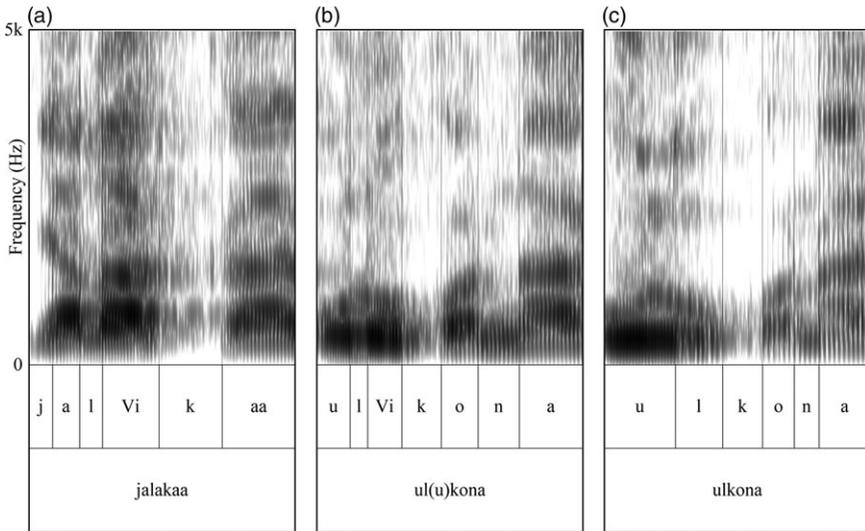
Figure 6. Three /lk/ sequences by a Ranua (Savo) speaker. (a) jalakaa (long inserted vowel); (b) ul(u)kona (relatively short insertion); (c) ulkona (no insertion).

phonetic excrescence, but the variability is centered on the /hC/ sequences. This suggests that the two dialects are at different stages of phonologization: Pohjanmaa has fully phonologized the inserted vowel, while Savo is lagging behind, specifically in the /hC/ sequences.

2.5 Duration of inserted vowels indicates phonological status

In this section, I compare the duration of inserted vowels in triggering tokens to the duration of the underlying V_2 in baseline tokens. Since inserted vowels are typically copies of V_1, it was possible to normalize duration across speakers, speech rate, and phrase position by dividing the duration of V_2 or V_i by the duration of V_1 (in all cases an underlying vowel), producing the ratio of V_2 (or V_i) to V_1 in a particular token. In total, there were 65 baseline words for the Pohjanmaa dialect, and 55 baseline words for the Savo dialect. This analysis only includes triggering tokens where insertion was observed; thus, the triggering tokens where insertion was predicted but did not occur are excluded.

Overall, the duration of inserted vowels patterns the same as underlying V_2, despite small statistically significant differences. When including all tokens with insertion, the addition of word type significantly improves the model ($\chi^2(1) = 6.17, \mathrm{p} = 0.01$), where the ratio between inserted V_i and V_1 ($\beta = 1.44$, SE $= 0.06$) is significantly smaller than the ratio between underlying V_2 and V_1 ($\beta = 1.61$, SE $= 0.07$). The difference in ratio between word types is quite small, which does not suggest a difference in phonological status between inserted and underlying vowels. In addition, the ratios for both word types is approximately 1.5, as illustrated in Figure 7. This is in fact expected in these dialects, which exhibit so-called "second-mora lengthening" (Suomi and Ylitalo 2004). In these dialects, V_2 of a $CV_1CV_2(X)$ word is approximately 1.5 times as long as V_1 (further detail
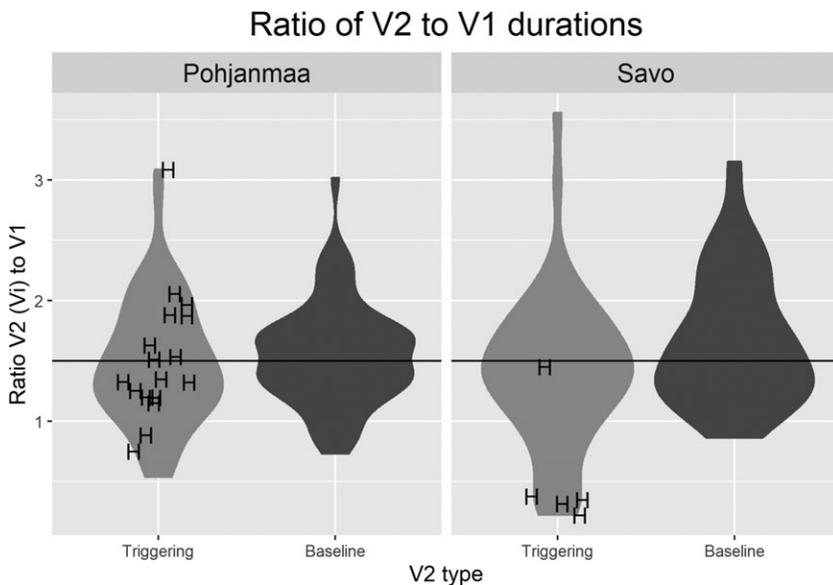
Figure 7. A violin plot showing the ratio of V_2 to V_1 durations, divided by dialect. For triggering tokens (CVC$_2$C$_3$VX), V_2 is the inserted vowel; for baseline tokens (CVCVX), V_2 is an underlying V_2. A horizontal line is at 1.5, the typical ratio from second-mora lengthening. The location of /hC/ tokens is marked with H.

on second-mora lengthening will be provided in Section 3). The presence of this lengthening points again to phonological status: not only is the inserted vowel as long as an underlying V_2—it is that long because it is treated as a second mora and lengthened accordingly.

The major exception to this appears to be the /hC/ tokens that do show insertion in Savo dialects, and these tokens drive the difference in ratio between V_2/V_1 and V_i/V_1. In general, the vowels in Savo /hC/ sequences are quite short, even compared to vowels in Pohjanmaa /hC/ sequences. This is illustrated in Figure 7, where almost all of the tokens in the lower tail of the Savo inserted vowel distrubtion are /hC/ tokens, marked with an "H": of the five /hC/ tokens that did trigger insertion, four are at the very bottom of the distribution. When /hC/ tokens from Savo speakers are removed from the dataset, the addition of word type as a factor does not significantly improve the model ($\chi^2(1) = 3.53, p = 0.06$). The addition of dialect as a factor does not significantly improve either model ($\chi^2(1) = 0.0001, p = 0.99$ with all tokens included; $\chi^2(1) = 0.14, p = 0.71$ with Savo /hC/ tokens removed), nor does the interaction between dialect and word type ($\chi^2(1) = 2.16, p = 0.14$ with all tokens included; $\chi^2(1) = 0.51, p = 0.78$ with Savo /hC/ tokens removed). Thus, for Pohjanmaa dialects, both consistency of insertion and the duration of the inserted vowel indicate that FVI is a phonological phenomenon. For Savo dialects, the evidence is mixed: /lC/ sequences very consistently produce long vowels, while /hC/ sequences inconsistently show insertion, and the inserted vowels that exist are generally much shorter than underlying vowels.

2.6 Vowel quality: copied vowels in both dialects

In this section, I compare the Euclidean distance (in Hz) between V_1 and inserted V_2 (in triggering tokens) and V_1 and underlying V_2 (in baseline tokens). Formant measures were taken from the middle 20 ms of each vowel using Praat's Get Formant functions. Euclidean distances were calculated using the F1 and F2 of each vowel. In order to compare directly between inserting words (where V_i is a copy of V_1) and baseline words, only baseline tokens with the same vowel quality for both V_1 and V_2 were included in these models. As there are very few tokens per person, and not all peripheral vowel qualities could be extracted for every speaker, it was impossible to do a study-wide normalization; however, as for the duration measures, the comparison of vowels of the same quality within one word effectively normalizes the vowel space across the study.

In both dialects, inserted vowels are no further from V_1 than an underlying V_2 of the same quality. The addition of word type does not significantly improve the fit of the model ($\chi^2(1) = 0.007, p = 0.93$), and the inclusion of dialect as a second fixed effect also does not significantly improve the model ($\chi^2(1) = 0.50, p = 0.48$). The interaction between dialect and word type also does not significantly improve the model ($\chi^2(1) = 0.04, p = 0.83$). This indicates that in both dialects, inserted vowels are typically copies of V_1, and not either schwa or some intermediate vowel. This is as reported by Kettunen (1940)—the only region with intermediate quality vowels is the southernmost region of Savonia, and no speakers in this sample are from that particular area.

As Euclidean distances convey only absolute value and not direction, it is also prudent to consider the peripherality of inserted vs. underlying vowels. Although a statistical analysis is not possible, the plots in Figure 8 illustrate that inserted vowels are not, on the whole, more centralized than the underlying V_1. If anything, Savo inserted vowels appear to be largely more peripheralized, and Pohjanmaa vowels appear to be shifted slightly back. These two facts combined indicate that the inserted vowels have undergone a high degree of phonologization.

The one exception to inserted vowels being a copy of V_1 is the vowel inserted in /lj/ sequences. As described in Section 1, inserted vowels are reported to be [i] when C_3 is /j/, rather than a copy of V_1. In the dataset overall, there are 7 tokens with insertion where C_3 is /j/; the mean Euclidean distance between V_i and V_1 is 507.09 Hz when C_3 is /j/, compared to a mean of 147.99 Hz for the other 119 tokens. The movement is overall towards the /i/ quality, as pictured in Figure 9. As Hall (2006) argues that only excrescent vowels are affected by surrounding consonants, this last pattern of [i] in /Cj/ sequences indicates, minimally, an excrescent origin, if not enduring phonetic status.

2.7 Qualitative analysis

In this section, I present a qualitative analysis of the insertion phenomenon. First, I discuss some tokens of "blocking" sequences that appear to have some sort of vocalic interval, which indicates excrescence; this is followed by a brief discussion of the participation of the inserted vowels in phonological processes.
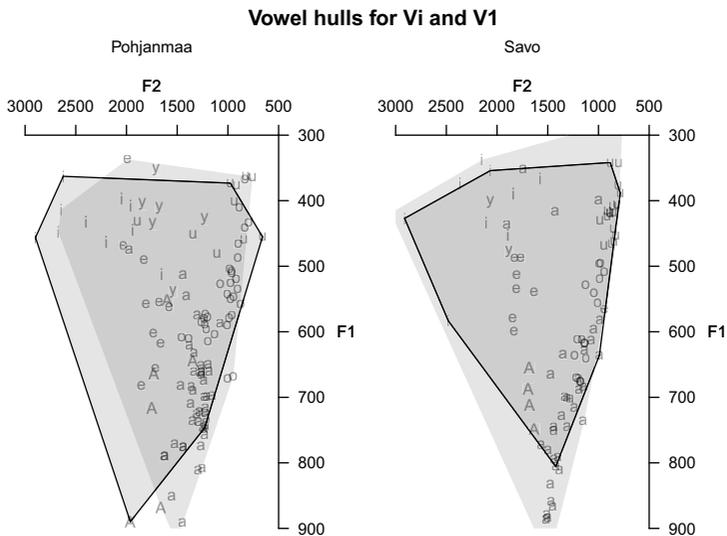
Vowel hulls for Vi and V1

Figure 8. Vowel spaces for triggering tokens in Pohjanmaa and Savo dialects, represented as convex polygons that encompass all tokens in the respective datasets. The polygons of V$_i$ have no outline and the polygons of V$_1$ have a solid outline. Individual vowel values are labeled with their quality; A stands in place of *ä* (/æ/).
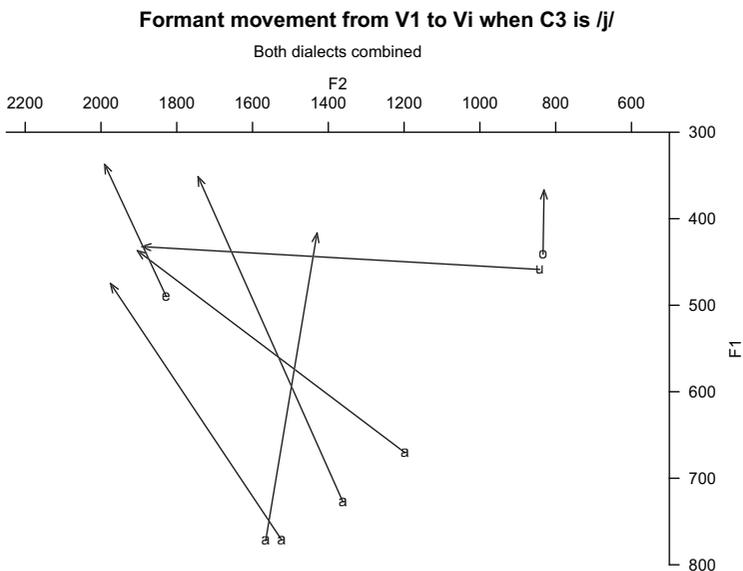
Formant movement from V1 to Vi when C3 is /j/

Figure 9. Vowel space showing the average movement from V$_1$ to V$_i$ when C$_3$ is /j/ (both dialects combined).
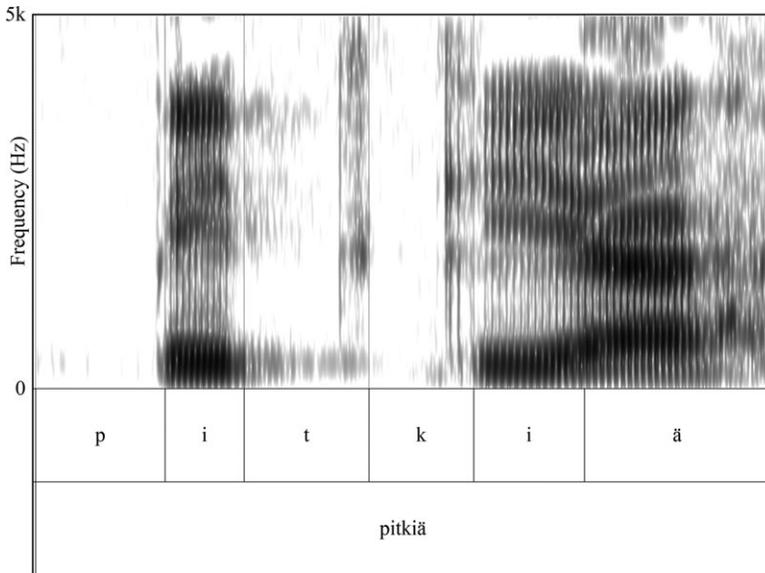
Figure 10. The word *pitkiä* 'long-PART.PL' with a 35 ms aspirated interval of C_2, by a Kuivaniemi (Pohjanmaa) speaker.

2.7.1 Evidence from blocking sequences

In this section, I present some evidence from "blocking" tokens (i.e., /rC/ and voiceless C_2 sequences) that contrast with the apparently phonologized inserted vowels, and thus provide support for an excrescent origin. Both blocking types present their own impediments to phonologization. First, heterorganic voiceless sequences could potentially have an excrescent vowel, but due to the lack of a voicing gesture in the consonants, the gap would be voiceless. This is in fact similar to Dep(f), the constraint proposed by Harrikari (1999) to address the same gap; for Harrikari, the insertion of the feature [+voice] was prohibited in [-voice][-voice] consonant sequences, and thus a vowel could only be inserted if there is at least one voiced consonant. This voicelessness would discourage reinterpretation as a vowel, particularly in a language (such as Finnish) that does not have word-internal devoiced vowels. However, while the excrescent vowel may not be voiced, there still appears to be a fairly long release in at least some sequences of voiceless consonants, as illustrated for the word *pitkiä* in Figure 10. In this word, there is a 35 ms space of aspiration-like noise after the release burst for the /t/, reminiscent of the aspiration described for Sierra Popoluca (Gafos 2002).

Yet another barrier to phonologization could be the attributability of the vocalic interval to C_2, for example when C_2 is an /r/, which is typically accompanied by short vocalic intervals even without additional gestural underlap. As noted in Section 1, there is a debate in the literature over the insertion status of /rC/ sequences, though the proponents of each side both believe in a fully phonological origin of the inserted vowels. For this paper, I will take the stance that vowel insertion in heterorganic /rC/ sequences is present, but the vowel is still fully excrescent and thus inconsistent, and resists phonologization due to factors I will discuss below.
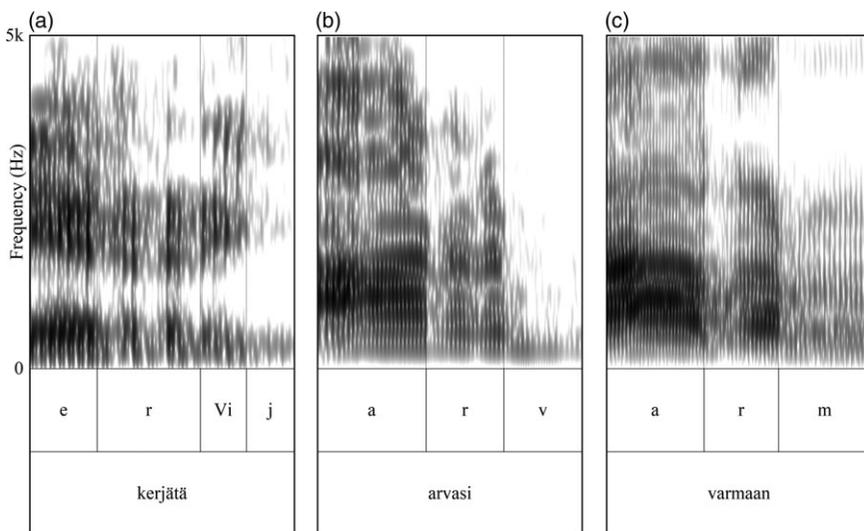
Figure 11. Three examples of /rC/ sequences in the corpus. (a) The word *kerjätä* 'to beg' with an excrescent vocalic portion following a trilled /r/, by a Kajaani (Savo) speaker; (b) the word *arvasi* 'guess.3SG.PAST', by a Lestijärvi (Pohjanmaa) speaker; (c) the word *varmaan* 'surely', by a Maaninka (Savo) speaker.

The first piece of evidence comes from consultation with native speakers of inserting dialects. Although Suomi (2000) reported that his Oulu consultants categorically denied insertion, the six native inserting speakers that I consulted on this question have no such consensus. One respondent felt that it was not possible to insert a vowel in any /rC/ sequences at all; another thought that insertion in /rC/ sequences was perhaps a bit strange, but not impossible. Interestingly, three respondents thought that insertion was possible only in /rj/ sequences (i.e., not /rk/ or /rv/ sequences or other sequences with similar places of articulation in C_3), but that the vowel is "weak" (example given: *kirja* 'book' > [kiria], without the [j]); a fourth thought that there would probably be no insertion in /rk/ or /rv/ sequences except possibly in fast speech (example given: *surkeaa* 'pitiful.PART' > [sur(u)kiaa[7]), but that there definitely was an inserted [i] in /rj/ sequences (examples given: *korjata* 'to fix' > [kor**i**jata], *kirja* 'book' > [kir**i**ja]). This disagreement among speakers suggests that there may sometimes be vowel-like intervals in /rC/ sequences, but they are certainly not deliberately produced.

Accordingly, there are some /rC/ tokens in the corpus that appear to have a longer-than-expected open interval between the closed portion of /r/ and the next consonant. As previously described, /r/ in coda position is canonically realized as a trill (Suomi et al. 2008). In some tokens in the corpus, the last open portion of the trill cycle lasts long enough such that it may be considered vowel-like. An example of this is provided in Figure 11a. In this token, a Kajaani (Savo) speaker produces the word *kerjätä* 'to beg' with a full trill (three closures), and the last open portion before the glide is 40 ms long, which is short compared to V_1 (82 ms) but long compared to the other open portions in the trill (both less than 20 ms long).

It should be noted that the example provided in Figure 11a is an /rj/ sequence, which is the sequence that four of the six native speakers thought was likely to have a

vowel. This apparent "exception" may be the result of a conflict between the articulatory postures necessary for /r/ and /j/ that makes gestural underlap more likely: the tongue body has to be braced in a certain position so that the tongue tip is free to vibrate, and this posture is incompatible with the high front constriction necessary for /j/. Thus, there is almost necessarily an interval of time where the tongue posture is no longer producing a trill, but also has not fully achieved the /j/ constriction; this interval could be interpreted either as part of the /r/ or as a separate vowel segment. In contrast, there would be much less conflict between the postures for /r/ and /k/, for example, and no conflict at all between /r/ and any labial consonant. Another example of a trill, this time with a bilabial C_3 and no long vocalic portion, is illustrated in Figure 11b. In this token, a Lestijärvi (Pohjanmaa) speaker produces the word *arvasi* 'guess.3SG.PAST' with a two-cycle trill; in contrast with the token in Figure 11a, the final open portion is approximately equal in duration to the first open portion.

There are, however, some sequences other than /rj/ that seem to have vowel-like intervals between the consonants, particularly when the /r/ is produced as a tap, rather than a trill. It is unclear if the tokens of /rC/ words with taps are idiosyncratic to the speaker or to a particular production of a word, but in these types of /rC/ tokens especially there seems to be a tendency to have a vocalic interval following. Taps are very short and previous work has described a vocalic portion on either side of the tap, as the body of the tongue must brace for the tap motion (Gibson et al. 2017); any delay of the following consonant constriction would simply extend this unconstricted interval, creating an excrescent vowel. One example of this type vocalic interval is provided in Figure 11c (from Maaninka, Savo dialect). Here, the word *varmaan* 'surely' has a 47 ms vocalic portion after the single closure for /r/. This vocalic portion is comparable in duration to some of the excrescent /hC/ tokens in Savo dialects, but is much shorter than would be expected from a vowel in second mora position.

Overall, /rC/ sequences exhibit inconsistent insertion, and tokens that do have insertion uniformly have short vocalic intervals. The variability of insertion seems to have impeded the phonologization of the vowel, similarly to /hC/ sequences in the Savo dialect. Furthermore, as /r/ inherently comes with short open intervals (whether tapped or trilled), phonologization may be further inhibited by the possibility of attributing these excrescent vowels to the consonant itself, which is not available for consonants like /l/ and /n/. These characteristics address the debate in the literature: vowels sometimes appear (as Harrikari 1999 argues), and sometimes do not (as Suomi 1990 argues); as they are not phonological and appear as the result of gestural underlap, they are less likely to be perceived by speakers. A failure to be perceived by speakers does not necessarily mean that the vowels do not occur, but it does indicate that they are not phonologized.

2.7.2 Participation in phonological processes

Finally, Finnish inserted vowels have an interesting dual patterning in other phonological and morphophonological processes. In her 1999 analysis, Harrikari treated FVI as a phonological repair, indicating phonological status. However, she also noted that the inserted vowel appears to be invisible to allomorphy that is sensitive to the number

of syllables. In Finnish, there is variability in the plural cases, where for example the trisyllabic word *lakana* 'sheet' is realized as [lakanoita] in the partitive plural (Anttila 2002, 2010). Disyllables, on the other hand, are realized without the /t/: the word *kala* 'fish' is realized as [kaloja] in the plural partitive and *[kaloita] is ungrammatical. Words that are disyllables without insertion and trisyllables with insertion function as disyllables in this respect: e.g. the plural partitive of *ahma* 'wolverine' ([ah**a**ma] in the inserted form) can only be realized as [ah**a**moja], not *[ah**a**moita]. The impossibility of the trisyllabic forms has been confirmed by the panel of native inserting speakers consulted for this paper—and unlike the possibility of inserting vowels in /rC/ forms, they are all in agreement on this fact.

On the other hand, although the inserted vowels are never stressed, they do appear to count towards footing. In Finnish, primary stress is fixed on the first syllable, and secondary stress falls on every other syllable thereafter modulo an attraction to weight (Kiparsky 2003; Karvonen 2005). Final syllables that follow an unstressed syllable are also unstressed unless they are heavy, in which case they optionally receive secondary stress. However, when comparing the footing rules described by Karvonen (2005) to Pohjanmaa judgments (provided by a subset of the panel of native speakers that have an intuition for secondary stress), it is clear that the inserted vowel is visible to footing:

(8) **Standard form**   **FVI form**   **Gloss**
    a. ˈta.pah.(ˌ)tuu   ˈta.pah.tuu   *happen.*3SG
    b. ˈmuut.taa   ˈmuut.taa   *change.*3SG
    c. ˈal.ko.vi   ˈa.la.ˌko.vi   *alcove*
    d. ˈhal.val.la   ˈha.la.ˌval.la   *easy.*ADESS

The relevant forms here are (8c) and (8d), which both have four syllables in the inserted form and have two trochees. There are two things to notice in this data: first, that the consulted speakers explicitly included the syllable boundaries that indicated that the inserted vowel can form a nucleus, and second, that these inserted segments count for the footing process. Thus, at least for these speakers, the vowels are phonologically active in stress assignment. The apparent contradictory patterning of the inserted vowel between morphophonological processes (where the vowel is not treated as a nucleus) and stress assignment (where the vowel is treated as a nucleus) supports an excrescent origin with phonologization: the plural partitive patterns for words with vowel insertion were fossilized prior to the phonologization of vowel insertion, but post-lexical stress processes treat the vowel as a true phonological unit.

3. Origins of FVI and a gestural model

Thus far, I have shown that FVI shows characteristics of both phonetic excrescence and phonological epenthesis: the phonological distribution of triggers strongly suggests an excrescent nature, while the consistency and duration of the vowels indicates phonological status (see Table 3 for summary). Based on the joint phonological and phonetic data, I have argued that FVI has an excrescent origin, which accounts for the phonological distribution, but over time has become phonologized, which accounts for the consistency and duration. I've also proposed that

Table 3. A summary of the results of the diagnostic tests. Letters indicate if the characteristic is true of S(avo) dialects and/or P(ohjanmaa) dialects

		Excrescent vowels		Epenthetic vowels
Environment	S, P	Between heterorganic consonants only		May or may not be between heterorganic consonants
Markedness	S, P	Not a repair		Repairs marked structures that are avoided elsewhere in the language
Optionality	S	Optional (**/hC/ tokens only**)	P, S	Obligatory; presence not dependent on speech rate
Duration	S	Variable duration (**/hC/ tokens only**)	P, S	Consistent duration
Quality	S, P	Often schwa; **influenced by neighboring consonants**		Not necessarily schwa, may be a fixed vowel; not influenced by consonant
Quality II	S, P	If copy of nearby vowel, only across sonorants or gutturals		No restrictions on copying

(at least at the time of corpus collection), the Savo and Pohjanmaa dialects are at two different stages of phonologization.

What still remains unexplained is the motivation for excrescence in the first place. As was described in Section 1, insertion is limited to specific positions in the word: the sequence /lm/ triggers insertion, but only when the /l/ is the second mora of the word. Thus, a gestural account of this phenomenon must predict excrescence only in this prosodic position, and not assume that all heterorganic sequences exhibit underlap.

In this section, I argue that FVI is crucially linked to another dialectal phenomenon, Second Mora Lengthening, which was briefly referred to in Section 2. I first situate SML in a broader context, with two major components: a discussion of the phonological and prosodic systems of Finnic languages, as are relevant to SML and FVI; and a description of the phenomenon in Finnish. I then show that a connection between FVI and SML accounts for both the prosodic and the dialectal distribution of FVI, predicting gestural underlap precisely when the second mora is a consonant, and only in dialects with SML. I then sketch the necessary components of a model that would produce this gestural underlap as the result of SML.

3.1 SML and the dialectal distribution of FVI

Second-mora lengthening (SML) is precisely as the name implies: in dialects that have SML, the segment in second mora position is lengthened. Hints of SML are visible in other Finnic languages in the so-called "half-long vowel", which describes the extended duration[8] of V_2 in $CV_1(C)V_2(X)$ words (see Prince 1980; Asu and Teras 2009 for Estonian (est), Kiparsky 2016 for Livonian (liv), and Gordon 2009 for Ingrian (izh)). This half-long vowel has also been documented as a feature of central and northern Finnish dialects since at least the late 19th century:

"... the vowel in an unstressed syllable following a light stressed syllable 'is stretched.' If the word receives sentence accent, the single vowel in question corresponds, according to Setälä, durationally to a double vowel, but in other instances the vowel has only one and a half times the duration of single vowels in other positions." (Setälä 1882, cited in Ylitalo 2009: 35)

More recently, a number of phonetic studies (Suomi and Ylitalo 2004; Suomi 2009; Ylitalo 2009) have shown that the half-long vowel is still robust in central and northern dialects, and is preserved by university-educated female speakers of Northern Finnish even in laboratory-elicited speech (Ylitalo 2004). Moreover, these studies have also shown that the half-long vowel phenomenon has been generalized to the second-mora position, thus generating the name "second-mora lengthening" (Suomi and Ylitalo 2004; Suomi 2009; Ylitalo 2009). This phenomenon affects consonants in second mora position as well as vowels—that is, precisely the consonant that triggers vowel insertion. The effects of SML are summarized and exemplified below:

(9)   Affects V_2 in initial CVV syllables, in both long vowels and diphthongs, as in (9a) and (9b);
    a.   ka**a**to   > kaaꞏto   *fall*
    b.   ma**i**to   > maiꞏto   *milk*
(10) Affects V_2 in words with initial CV.CV, as in (10a);
    a.   kat**o**   > katoꞏ   *deficiency*
(11) Affects C_2 in initial CVC syllables, as in (11a) and (11b);
    a.   ka**t**to   > katꞏto   *roof*
    b.   sy**n**tyy   > synꞏtyy   *be born.*3SG
(12) Is not iterative—i.e., does not affect the second mora of every foot, as in (12a).
    a.   kuv**i**telma   > kuviꞏtelma   *fantasy*
                    * > kuviꞏtelꞏma

Linking FVI explicitly to SML neatly accounts for the dialectal distribution of FVI: crucially, the dialects of Finnish that do not exhibit SML also do not have FVI (Kettunen 1940; Wiik and Lehiste 1968; Spahr 2012). If FVI were viewed as an entirely separate innovation, it would be one that developed independently across the major East-West Finnish dialect split, as Savo dialects are in the Eastern group and Pohjanmaa dialects are in the Western group. In addition, at least some of the Western Finnish dialects spoken in the Lapland region (the "Far North" dialects) also exhibit FVI (Kettunen 1940) as well as SML (Spahr 2012); however, due to the lack of experimental and corpus data available from these regions, they have been excluded from the present paper.

If we instead view SML as a necessary precursor to FVI, the major innovation is limited to the hämäläismurteet, which are the subset of Western dialects spoken by over two million people in the major regions of the south of Finland. SML is not currently documented in these dialects (see e.g. Ylitalo 2009 for the dialect spoken in Tampere, as well as Wiik and Lehiste 1968 for a more general survey), and thus must have been lost minimally before the phonologization of the inserted vowel occurred. An investigation of why some dialects with SML develop FVI but not

others is beyond the scope of this paper; however, in some cases SML may be produced by a different timing structure that would not generate FVI. For example, the dialect region of Turku has been documented to have the half-long vowel (Wiik and Lehiste 1968; Ylitalo 2009) but not vowel insertion. However, Ylitalo (2009) argues that SML in Turku dialects is not the same as SML in the Oulu (Pohjanmaa) dialect: while the Oulu dialect achieves SML through lengthening the second mora, the Turku dialect achieves a SML through shortening the first mora. Thus, the lengthening process would not be targeting C_2, and accordingly would not generate vowel insertion.

3.2 SML and the prosodic distribution of FVI

As Hall's (2006) model of excrescent vowels is rooted in articulatory gestures in the Articulatory Phonology tradition, a gestural model of SML and FVI is necessary to reap the full benefits—to create gestural underlap, one must first have a gestural model. A fully formalized and computationally implemented model is beyond the scope of this paper; however, I discuss here the components of a gestural model that I believe would generate both SML and FVI. A link between SML and FVI addresses the positional limitation, which has previously not been satisfactorily accounted for. The relevant environment is $CVC_2.C_3V(X)$, and not later C_2C_3 sequences—this is precisely when C_2 would be targeted by SML. This makes a slightly more complicated argument for gestural underlap. Typically, excrescence via gestural underlap occurs whenever that particular sequence of consonants occurs, or is limited by position within a syllable (e.g., excrescence occurs only in onsets, or only in codas, or heterosyllabically); in these dialects of Finnish, underlap is conditioned by second mora status.

In order to account for the very particular environments, I propose the existence of an oscillator in Finnish that corresponds to a bimoraic foot, which is coupled both to segmental gestures as well as to a boundary π gesture. This configuration is illustrated in Figure 12. In Articulatory Phonology, an oscillator is simply "any process that tends to repeat itself regularly" (O'Dell and Nieminen 2009: 2), after harmonic oscillators in physics. Articulatory gestures are typically analyzed as oscillators in Articulatory Phonology, as in repetitive speech (such as [mamama]) there would be an oscillatory motion of the articulators. In addition, oscillators for larger prosodic units have also been proposed (O'Dell and Nieminen 2009), such as for the mora, the syllable, and the foot.

The "bimoraic foot" descriptor of this proposal is simply a gestural interpretation of Suomi's (2009) bimoraic "locus of accentual lengthening". In his study, Suomi found that segments in the first two moras of a word were lengthened relative to comparable segments in later positions; the second mora, in particular, was greatly affected, producing SML. Karlin (2015) also argued for a bimoraic foot at the beginning of a word in Central and Northern dialects of Finnish, citing both SML as an instantiation as domain-final lengthening, as well as another dialectal repair that effectively prevents a bimoraic foot from dividing a syllable (i.e., CVCV: > CVC:V:).

The production of SML itself relies on an additional π gesture (Byrd and Saltzman 2003), at the end of the bimoraic oscillator. As the alternative name for these gestures, "clock-slowing gesture", implies, a π gesture at prosodic boundaries slows

(a)

CVCV /kala/ and CVC /kal/ with a bimoraic oscillator with no π gesture.

(b)

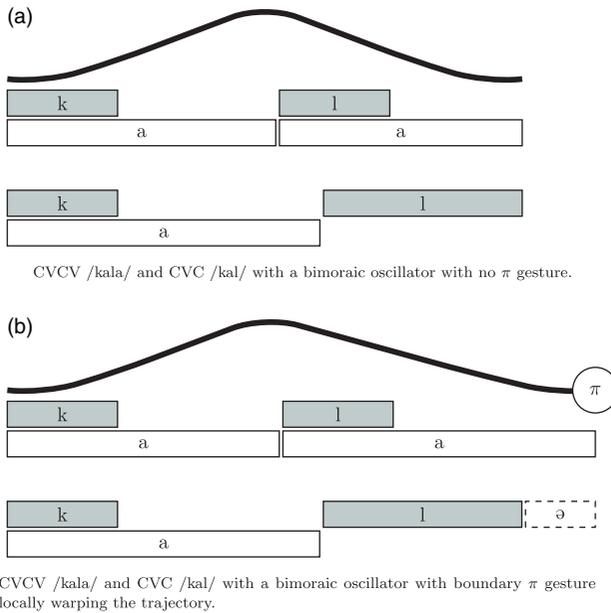CVCV /kala/ and CVC /kal/ with a bimoraic oscillator with boundary π gesture locally warping the trajectory.

Figure 12. Schematics for the bimoraic oscillator (black trajectory) coupled to CVCV(X) and CVC(X) gestural scores. Note that in both cases the coda /l/ is stretched due to its position in the coda, but the π gesture warps the bimoraic oscillator further than the stretched coda can go (figures left-aligned for ease of comparison).

the progress of some oscillator, effectively locally lengthening the gestures. These gestures have been invoked to describe processes such as phrase-final lengthening and the warping of gestural timing relations near prosodic boundaries (Byrd and Saltzman 2003; Katsika et al. 2014). In dialects with SML, there appears to be an active prosodic boundary at the end of the bimoraic foot, which comes with boundary-associated lengthening. A clock-slowing gesture here would locally slow the progress of the bimoraic oscillator, producing SML.

An oscillator may also be coupled to another oscillator, either of the same type (e.g., segment-to-segment coupling) or different (e.g., foot-to-segment). In this model, the bimoraic oscillator is hierarchically coupled to the gestures that make up the segmental content. An oscillator has a particular preferred frequency, but its actual timing may be influenced by being coupled to another oscillator. For example, O'Dell and Nieminen (1999) proposed that a driving force of so-called "stress-timed" and "syllable-timed" languages is in fact the dominance of one type of oscillator over another: in stress-timed languages, an oscillator corresponding to stress groups dominates over a syllable oscillator, while in syllable-timed languages, the dominance relationship is reversed.

Dominance may additionally be affected by other properties of oscillators, such as mass or stiffness. The remaining question for this proposal is why consonant sequences show underlap and resultant excrescence, while vowels simply stretch to accommodate this newly-created "half-long" spot. I attribute this failure to stretch to high gestural stiffness, which effectively describes an oscillator's resistance

to deformation. In Articulatory Phonology, stiffness has been proposed as one of the defining parameters of gestures (Browman and Goldstein 1989, 1990); gestures with high stiffness have a shorter cycle and higher amplitude (such as consonants with high degrees of closure), while gestures with low stiffness have a longer cycle and lower displacement amplitude (such as vowels; see e.g. Fuchs et al. 2011 for kinematic data and mathematical models). In this particular situation, the bimoraic oscillator is coupled with sufficient dominance to the segmental gestures to force vowel gestures to deform, as they have low resistance to displacement, but not consonant gestures, as they are highly resistant to deformation. Thus, the bimoraic foot oscillator is deformed by the π gesture, but cannot sufficiently deform the consonant gesture, effectively producing a gap where the bimoraic oscillator is still active without full articulatory closure.

Finally, although for this paper I have focused on $CVC_2.C_3V(X)$ forms, it should be noted that insertion is also present when C_3 is a geminate, as in the word *helppo* 'easy' > [hel**e**p.po]. The gestural proposal still works for these words—the π gesture is still targeting the second mora, but in this case, the second mora simply has two associated closure gestures. The same effects thus apply: the closure gesture for C_2 attempts to stretch but fails, producing a gap between that gesture and the following C_3; the lengthening of the second gesture may also fail but the homorganic closure that follows closes the gap. This type of "non-local" lengthening has been documented in phrase-final lengthening, where, for example, lengthening effects are found in penultimate syllables, rather than just final segment or even syllable (Turk and Shattuck-Hufnagel 2007); as the π gesture has been proposed as the major driver of phrase-final lengthening, the current application fits with existing data.

3.3 The phonologization of phonetic detail

Finally, there is still work to be done regarding the process of phonologization of the excrescent vowels. Hall (2006) cites several cases of fully phonologized excrescent vowels, including related cases such as Lule Saami (Engstrand 1987) and Lapua (southern Pohjanmaa) Finnish (Harms 1976). Ohala (1993) has argued for language change due to processes of hypocorrection, where phonetic effects are reinterpreted as intentional, as in tonogenesis where the raised F0 that accompanies voiceless consonants is reinterpreted as the primary cue. In this case, phonologization essentially entails a process of turning nothing into something—i.e., a gap with no active constriction is reanalyzed as a full segment with an intentional gesture. Hall (2006: 424) also notes that phonologization is neither "an inevitable fate for intrusive vowels, nor does it happen automatically upon their reaching some threshold of phonetic duration". In this section, I turn to the process of phonologization of FVI in Central and Northern dialects, and discuss possible influences on phonologization as derived from the coupled oscillator model, and as evident from the acoustic corpus.

The two end stages of this process are illustrated in the schemata in Figure 13, using a gestural score that correlates to the previously described coupled oscillator model. It is worth noting that this model does not necessarily predict *no* lengthening of the consonants, simply insufficient lengthening to form sufficient constriction for the entire "half-long" time warping of the bimoraic oscillator (a slight lengthening of
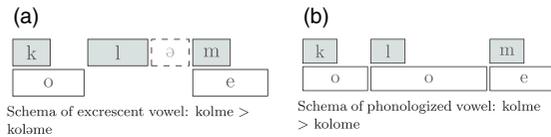
(a)

```
k        l    ⌐ə⌐   m
     o              e
```
Schema of excrescent vowel: kolme > kolᵊme

(b)

```
k        l           m
     o        o      e
```
Schema of phonologized vowel: kolme > kolome

Figure 13. Schemata of a simplified gestural score corresponding to (a) an excrescent vowel and (b) a fully phonologized epenthetic vowel in the word *kolme*.

/l/ is depicted in Figure 13a). More specifically, the model predicts that stiffer consonants would be most likely to create gestural gaps, as they would be the most resistant to deformation; less stiff gestures (such as the fricative /h/) would be less resistant and thus more able to follow the stretching of the foot-level oscillator. This aligns with the /lC/ and /nC/ exhibiting full phonologization in both dialects, as they both involve full closure at the tongue tip and are thus very stiff. It also aligns with the /hC/ sequences lagging behind in the Savo dialect, as /h/ is more open, less stiff, and thus more prone to deformation.

These stiffness differences would have a twofold effect on the phonologization process. First, one factor that one would expect to influence phonologization is the *consistency* with which insertion occurs. If /h/ is sufficiently flexible to only occasionally create gaps, then those gaps would be less likely to be interpreted as deliberate. Similarly, the behavior of /rC/ sequences also suggests that /r/ is less stiff than /l/ and /n/; this may be literally true in the stiffness of the tongue tissue (which has also been suggested to be linked to gestural stiffness, as in Fuchs et al. 2011) despite the full closures. However, the /r/ case is even more complex, as it is influenced by articulatory and aerodynamic constraints, as well as inherent open intervals, none of which affect either /l/ or /n/. Second, although Hall (2006) noted that duration by itself is not sufficient to predict phonological status (as evidenced by Scots Gaelic), longer gaps are more likely to be interpreted as intentional. As /l/ and /n/ are less easily deformed, the small amount that they do deform would leave a bigger gap than the more flexible /h/. Thus, this model predicts some cross-segment variability and thus does not rule out different rates of phonologization. Although the corpus data overall reflects a fairly late stage of phonologization, the Savo dialect in particular hints at stiffness-correlated differences in phonologization. A future study may look for dialects with true excrescence across the board and examine the duration of the inserted vowel with different C_2.

4. Conclusion

In this paper, I have presented evidence that FVI is the mixed result of phonology and phonetics: the inserted vowels originated as excrescence, but over time have been phonologized, whether completely (as in the Pohjanmaa dialect) or partially (as in the Savo dialect). I have also argued that a link between FVI and SML satisfactorily explains both phonological and dialectal distribution, which has not previously been shown. Finally, I proposed a gestural model using coupled oscillators that describes this link between FVI and SML.

There are still many avenues to explore in this approach. First, a formalization of the proposed gestural model and a computational implementation to test the predicted effects are in order. Second, although I have argued that SML is caused by boundary lengthening of a bimoraic foot, it is still a rather curious phenomenon, as it frequently produces length on an unstressed syllable (i.e., in CVCV words). An additional curiosity is the ability of Central and Northern Finnish speakers to eliminate inserted vowels from their speech when speaking standard Finnish (as elicited in lab speech), while still maintaining SML (Ylitalo 2009; Suomi 2009). The dialects included in Suomi and Ylitalo's work work are from Pohjanmaa, and as such their vowels are fully phonological and thus more easily removed than a sub-phonemic durational effect. However, my current model of SML and FVI predicts that some vowel excrescence should occur in this situation, and no such vowels were reported by either Suomi or Ylitalo. It is possible that there are additional shifts in the articulatory grammar of such speakers when speaking Standard Finnish; perhaps the active suppression of vowels in this position creates some gestural crowding between C_2 and C_3. Further work on Northern dialect-influenced standard Finnish is needed.

In addition, although I have argued that the vowels have phonologized (to some extent) in both Savo and Pohjanmaa dialects, it is also unclear what the underlying form of words that show insertion is. The linguistic situation in Finland is fairly unique, as there is a state of triglossia: first, the dialectal Finnish that children would acquire and use at home, such as the Savo and Pohjanmaa dialects; second, standard spoken Finnish (*puhekieli*, 'spoken language'), the standardized form with local "colorings" (Suomi 2009) that is used in the media and in schools; and third, the written standard (*kirjakieli*, 'book language'), which is used in educational contexts throughout Finland. Both puhekieli and kirjakieli would provide learners with evidence that the vowel insertion is not "real". However, the extent to which all these versions of Finnish are linked in their representation is unclear.

Notes

1. This is a slightly inaccurate name, since I have excluded southern Pohjanmaa dialects as they are in a separate subgrouping.

2. As Finnish orthography is largely phonemic, I will be using the orthography rather than the IPA throughout. *ä* represents /æ/, *ö* represents /œ/, and *v* represents the sonorant /ʋ/.

3. Abbreviations used in this paper: 3SG third person singular, PART partitive case, ALL allative case, ADESS adessive case, ESS essive case, LAT lative case.

4. With the singular exception of *hihhuli* 'religious fanatic', which has onomatopoetic origins.

5. It is unclear if the location tagged is location of recording or the place where the speaker grew up; however, as most of the recordings are from the 1960s to the 1970s, it is likely that the speakers did not move far from their hometown.

6. Note that the T(oken) columns in Table 1 will add up but not the L(emma) columns, as some lemmas were repeated in different locations.

7. The change of /e/ to /i/ is an additional feature of the Oulu (northern Pohjanmaa) dialect that is unrelated to vowel insertion.

8. Note that all of the aforementioned Finnic languages (including Finnish) are "full-fledged" quantity languages, distinguishing length in both consonants and vowels; the term "half-long" describes a lengthening that does not approach the categorical boundary (Ylinen et al. 2005).

References

Anttila, A. (2002). **Variation and phonological theory. In Chambers, J. K. and Schilling, N.**, editors, *The Handbook of Language Variation and Change*, chapter 8. Malden, MA and Oxford, UK: Blackwell.

Anttila, A. (2010). Word stress in Finnish. Handout from colloquium talk given at Yale University, February 1, 2010.

Asu, E. L. and Teras, P. (2009). Estonian. *Journal of the International Phonetic Association*, **39**(3):367–372.

Bates, D., Maechler, M., Bolker, B., Walker, S., et al. (2014). lme4: Linear mixed-effects models using Eigen and s4. *R package version*, **1**(7):1–23.

Boersma, P. and Weenink, D. (2017). *Praat: doing phonetics by computer.* http://www.fon.hum.uva.nl/praat/.

Bořil, T. and Skarnitzl, R. (2016). **Tools rPraat and mPraat. In Sojka, P., Horák, A., Kopeček, I., and Pala, K.**, editors, *Text, Speech, and Dialogue: 19th International Conference, TSD 2016, Brno, Czech Republic, September 12-16, 2016, Proceedings*, pages 367–374, Cham. Springer International Publishing.

Browman, C. P. and Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology*, **6**(2):201–251.

Browman, C. P. and Goldstein, L. (1990). Gestural specification using dynamically-defined articulatory structures. *Status Report on Speech Research*, pages 95–110.

Byrd, D. and Saltzman, E. (2003). The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics*, **31**(2):149–180.

Engstrand, O. (1987). Preaspiration and the voicing contrast in Lule Sami. *Phonetica*, **44**(2):103–116.

Fougeron, C. and Ridouane, R. (2008). On the nature of schwa-like vocalic elements within some Berber clusters. In *Proceedings of the 8th International Seminar on Speech Production, Strasbourg, France*, pages 441–444.

Fuchs, S., Perrier, P., and Hartinger, M. (2011). A critical evaluation of gestural stiffness estimations in speech production based on a linear second-order model. *Journal of Speech, Language, and Hearing Research*, **54**(4):1067–1076.

Gafos, A. I. (2002). A grammar of gestural coordination. *Natural Language & Linguistic Theory*, **20**(2): 269–337.

Gibson, M., Sotiropoulou, S., Tobin, S., and Gafos, A. (2017). On some temporal properties of Spanish consonant-liquid and consonant-rhotic clusters. In *Proceedings of the Conference on Phonetics and Phonology in German-speaking Countries, Berlin, Germany*, pages 73–76.

Gordon, M. (2009). Prominence and gemination in Ingrian. *Linguistica Uralica*, **45**(2):81–100.

Hall, N. (2006). Cross-linguistic patterns of vowel intrusion. *Phonology*, **23**(3):387.

Hall, N. (2011). **Vowel epenthesis. In van Oostendorp, M., Ewen, C., Hume, E., and Rice, K.**, editors, *The Blackwell Companion to Phonology*, chapter 67, pages 1576–1596. Wiley Online Library.

Harms, R. T. (1976). The segmentalization of Finnish 'nonrules'. *Texas Linguistic Forum*, **5**:73–88.

Harrikari, H. (1999). Epenthesis, geminates, and the OCP in Finnish. *Nordic Journal of Linguistics*, **22**(1):3–26.

Harrikari, H. (2003). From phonology to morphology: A synchronic account of dialectal gemination in Finnish. *MIT Working Papers in Linguistics*, **45**:67–82.

Karlin, R. (2015). Second mora phenomena in Central and Northern dialects of Finnish. *Cornell Working Papers in Phonetics and Phonology*, pages 121–138.

Karvonen, D. H. (2005). *Word prosody in Finnish*. PhD thesis, University of California Santa Cruz.

Katsika, A., Krivokapić, J., Mooshammer, C., Tiede, M., and Goldstein, L. (2014). The coordination of boundary tones and its interaction with prominence. *Journal of Phonetics*, **44**:62–82.

Keating, P. A. (1988). Underspecification in phonetics. *Phonology*, **5**(2):275–292.

Kettunen, L. (1940). *Suomen murteet: murrekartasto*. Suomalaisen kirjallisuuden seura.

Kiparsky, P. (2003). Finnish noun inflection. In Mannine, S. and Nelson, D., editors, *Generative approaches to Finnic and Saami linguistics*, pages 109–161. CSLI publications, Stanford, California.

Kiparsky, P. (2016). Livonian stød. In Kehrein, W., Köhnlein, B., Boersma, P., and van Oostendorp, M., editors, *Segmental Structure and Tone*, chapter 7, pages 195–210. Walter de Gruyter, Berlin.

Ladefoged, P. and Maddieson, I. (1996). *The Sounds of the World's Languages*. Blackwell Publishers, Oxford, OX, UK.

Levin, J. (1987). Between epenthetic and excrescent vowels. In *Proceedings of the 6th West Coast Conference on Formal Linguistics, Tucson, AZ*, volume 6, pages 187–202.

Lyytikäinen, E., Rekunen, J., and Yli-Paavola, J. (2013). *Suomen murrekirja*. Gaudeamus.

O'Dell, M. and Nieminen, T. (1999). Coupled oscillator model of speech rhythm. In *Proceedings of the XIVth International Congress of Phonetic Sciences, San Francisco, CA*, volume 2, pages 1075–1078. University of California Berkeley.

O'Dell, M. and Nieminen, T. (2009). Coupled oscillator model for speech timing: Overview and examples. In *Prosody: Proc. 10th conference, Helsinki, Finland*, pages 179–190.

Ohala, J. J. (1993). The phonetics of sound change. In Jones, C., editor, *Historical Linguistics: Problems and Perspectives*, chapter 9, pages 237–278. Longman, London, UK.

Palander, M. (2011). *Itä- ja eteläsuomalaisten murrekäsitykset*. Suomalaisen Kirjallisuuden Seura.

Plug, L., Shitaw, A., and Heselwood, B. (2019). Inter-consonantal intervals in Tripolitanian Libyan Arabic: Accounting for variable epenthesis. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, **10**(1). http://doi.org/10.5334/labphon.122.

Prince, A. S. (1980). A metrical theory for Estonian quantity. *Linguistic Inquiry*, **11**(3):511–562.

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Spahr, C. E. (2012). Dialectal gemination in Finnish: Phonetics/phonology interplay in moraic theory. Master's thesis, University of Toronto.

Suomi, K. (1990). Huomioita yleiskielen konsonanttien yhdistelyrajoituksista ja pohjalaismurteiden epenteettisestä vokaalista. *Virittäjä*, **94**(2):139–160.

Suomi, K. (2000). Nearly a million Ostrobothnians cannot be wrong: Comments on an Optimality Theoretic analysis of "Ostrobothnian". *Nordic Journal of Linguistics*, **23**:89–93.

Suomi, K. (2009). Durational elasticity for accentual purposes in Northern Finnish. *Journal of Phonetics*, **37**(4):397–416.

Suomi, K., Toivanen, J., and Ylitalo, R. (2008). *Finnish Sound Structure: Phonetics, Phonology, Phonotactics, and Prosody*. Number 9 in Studia Humaniora Ouluensia. Oulu University Press, Oulu, Finland.

Suomi, K. and Ylitalo, R. (2004). On durational correlates of word stress in Finnish. *Journal of Phonetics*, **32**(1):35–63.

Turk, A. E. and Shattuck-Hufnagel, S. (2007). Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics*, **35**(4):445–472.

Vaux, B. (2003). Syllabification in Armenian, Universal Grammar, and the lexicon. *Linguistic Inquiry*, **34**(1):91–125.

Wiik, K. (1965). *Finnish and English vowels: A Comparison with Special Reference to the Learning Problems Met by Native Speakers of Finnish Learning English*. University of Turku, Turku.

Wiik, K. and Lehiste, I. (1968). Vowel quantity in Finnish disyllabic words. In *Congressus Secondus Internationalis Fenno-Ugristarum*, pages 569–574. Societas Fenno-Ugrica, Helsinki.

Ylinen, S., Shestakova, A., Alku, P., and 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.

Ylitalo, R. (2004). Toisen tavun vokaalin puolipidennyksestä Oulunseutulaisten puheessa. *Virittäjä*, **108**(3):414–422.

Ylitalo, R. (2009). *The realisation of prominence in three varieties of standard spoken Finnish*. PhD thesis, University of Oulu.