



universität
wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Sank you wery much for the tirdy pounts!“

„The pronunciation of specific English consonant sounds by native Finns and Austrians“

verfasst von / submitted by

Mirja Leena Kemppe, BA MSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Master of Arts (MA)

Wien, 2016 / Vienna, 2016

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on
the student record sheet:

A 066 812

Studienrichtung lt. Studienblatt /
degree programme as it appears on
the student record sheet:

Masterstudium English Language and Linguistics
UG2002

Betreut von / Supervisor:

Univ.-Prof. Mag. Dr. Nikolaus Ritt

Table of Contents

List of figures.....	iv
List of tables.....	vi
1 Introduction.....	1
2 Theory.....	3
2.1 Age and the acquisition of phonetic structures.....	3
2.2 Researched sounds.....	7
2.2.1 Voiced labial-velar approximant /w/.....	7
2.2.2 Voiceless and voiced (inter-)dental fricatives /θ/ and /ð/.....	7
2.2.3 Voiced plosives /b/, /d/, and /g/ in word-final position.....	8
2.2.4 Voiceless alveolar plosive /t/ in word-medial prevocalic position.....	9
2.3 The first languages and the stereotypical pronunciation of English.....	11
2.3.1 Finnish.....	11
2.3.2 Austrian German.....	13
3 Methods.....	15
3.1 Informants.....	15
3.2 Pronunciation tasks.....	18
3.3 Recording procedure.....	20
3.4 Visualisations of audio data.....	21
4 Results.....	22
4.1 Voiced labial-velar approximant /w/.....	22
4.2 Voiceless and voiced (inter-)dental fricatives /θ/ and /ð/.....	36
4.3 Voiced plosives /b/, /d/, and /g/ in word-final position.....	55
4.4 Voiceless alveolar plosive /t/ in word-medial prevocalic position.....	74
5 Discussion.....	78
6 Conclusions.....	83
7 References.....	86
8 Appendices.....	90
8.1 Appendix A: The word list.....	90
8.2 Appendix B: The text passage.....	91
8.3 Appendix C: The analysed strings from the text passage.....	92
8.4 Appendix D: Questionnaire.....	93

8.5 Appendix E: The Finnish informants.....	95
8.6 Appendix F: The Austrian informants.....	97
8.7 Appendix G: Statistics of individual informant performance.....	99
8.8 Appendix H: Abstract.....	101
8.9 Appendix I: Zusammenfassung.....	102

List of figures

Figure 1.....	10
Figure 2.....	13
Figure 3.....	15
Figure 4.....	23
Figure 5.....	24
Figure 6.....	25
Figure 7.....	26
Figure 8.....	27
Figure 9.....	28
Figure 10.....	29
Figure 11.....	29
Figure 12.....	30
Figure 13.....	31
Figure 14.....	31
Figure 15.....	32
Figure 16.....	32
Figure 17.....	33
Figure 18.....	33
Figure 19.....	34
Figure 20.....	34
Figure 21.....	35
Figure 22.....	35
Figure 23.....	37
Figure 24.....	38
Figure 25.....	40
Figure 26.....	41
Figure 27.....	43
Figure 28.....	43
Figure 29.....	44
Figure 30.....	45
Figure 31.....	45
Figure 32.....	46
Figure 33.....	47
Figure 34.....	48
Figure 35.....	48

Figure 36.....	49
Figure 37.....	49
Figure 38.....	50
Figure 39.....	50
Figure 40.....	51
Figure 41.....	52
Figure 42.....	53
Figure 43.....	53
Figure 44.....	54
Figure 45.....	54
Figure 46.....	55
Figure 47.....	56
Figure 48.....	57
Figure 49.....	59
Figure 50.....	61
Figure 51.....	62
Figure 52.....	63
Figure 53.....	64
Figure 54.....	65
Figure 55.....	66
Figure 56.....	67
Figure 57.....	68
Figure 58.....	69
Figure 59.....	70
Figure 60.....	71
Figure 61.....	72
Figure 62.....	73
Figure 63.....	74
Figure 64.....	75
Figure 65.....	76
Figure 66.....	77
Figure 67.....	82
Figure 68.....	83

Index of Tables

Table 1: p-values from the χ^2 -test for /w/.....	24
Table 2: p-values from the χ^2 -test for ‘th’, /θ/ and /ð/.....	42
Table 3: p-values from the χ^2 -test for the word-final voiced plosives /b/, /d/ and /g/.....	58
Table 4: p-values from the χ^2 -test for the word-medial prevocalic /t/.....	77
Table 5: p-values from the χ^2 -test for the occurrences of all researched sounds in both pronunciation tasks.....	85

1 Introduction

Stampe's theory about natural phonology states that phonological acquisition is based on suppressing and applying natural articulatory processes, such as fortition, lenition and aspiration, to produce language specific sounds (1979). In learning a new language (L2), the processes of the first language (L1) are often applied to the phonology of the L2. A complete failure to learn to suppress and to apply the processes of L2 results in a stereotypical foreign accent. Foreign accents become problematic when the accented pronunciation leads to misunderstandings and/or distracts the listener. The probability of misunderstandings increases when the words pronounced in non-standard-like manner resemble other words from the same word class e.g. *think* and *sink*, or *back* and *bag*. This study on the pronunciation of English aims to answer the following research questions by examining the informants' pronunciation of specific English sounds:

1. How prevalent are the stereotypical Finnish- and German-style pronunciations among young adults of both nationalities?
2. Are there any differences between the two informant groups?

According to Scovel (1988), there is a critical period of language learning which applies only to pronunciation acquisition, i.e. the lack of exposure to the phonetic structures of a language during childhood and early adolescence often results in a foreign accent. Several studies, such as by Flege et al. (1995), Flege et al. (1999) and Piske et al. (2001), imply that early exposure to foreign languages, in both educational and naturalistic settings, is beneficial especially to the acquisition of the phonetic structures. In her approach based on mental categories, Bialystok suggests that differences between children and adults in acquiring a new language and its phonetic structures result from differing preferences for learning methods rather than from any maturational changes in the brain such as the lateralization (1997).

The main hypothesis of this study is two-fold. Firstly, mainly due to increased exposure to English through different media and higher quality English tuition, the pronunciation samples from young adults are expected to contain less instances of the stereotypical non-native-like pronunciation regardless of the informant's nationality. Secondly, resulting from the low percentage of synchronised films and television

programmes in Finland, the pronunciation of the Finnish informants is presumed to be more native-like and to have more instances of typically American pronunciation.

The purpose of this study is to analyse how specific English sounds are pronounced by native Austrians and Finns in order to assess the prevalence of stereotypical features and to evaluate the possible effects of exposure to English through different media in non-institutional settings. The influence of non-institutional exposure to English through non-synchronised films and television programmes on learners' pronunciation has received little attention in previous studies. The expected correlation between a more native-like accent and the exposure to spoken English through TV would highlight the benefits of non-synchronised programmes to the acquisition of new phonetic structures.

The pronunciation analysis is based on samples from two groups of informants, each group consisting of more than 20 informants. All informants belong to a specific age group (born between 1979 and 1987) and have a similar educational background (at least a bachelor's degree from a regular university or from a university of applied sciences). The only differences between the two informant groups are the informant's nationality (L1) and the different amounts of English input through non-institutional settings. Both of the informants' first languages were taken into account when choosing the sounds investigated in this study. To avoid discrimination between the two informant groups, only sounds that do not naturally appear in either of the L1s (standard Finnish and standard Austrian German) were chosen. These sounds include the voiced labiovelar approximant /w/, both the voiceless and the voiced (inter-)dental fricatives /θ/ and /ð/, and the voiced plosives /b/, /d/, and /g/ in word-final position. The influence of exposure to American English is examined by analysing the pronunciation of the voiceless alveolar plosive /t/ in word-medial prevocalic positions.

The pronunciation samples were collected by recording the informants' performance in two pronunciation tasks: a word list and a short text passage. In most of the previous studies the accents of the informants have been evaluated by a panel of native speakers. In this study, the evaluation of the samples is done primarily in a more objective manner. First, the collected samples are transformed into waveforms and spectrograms using Praat (Boersma & Weenik 2015). The structures of the investigated sounds in these visualisations of audio data are then compared with the structures of both standard varieties (Received Pronunciation (RP) and General American (GA)) in order to identify possible differences and similarities in the pronunciation.

Following this introduction, section 2 presents the theoretical background for this study including a short review of the theories related to the relationship between age and language acquisition (2.1) and a detailed description of the standard-like pronunciation of the researched sounds (2.2). The stereotypical pronunciation of the investigated phonemes by both informant groups and the influence of the first languages are described in subsection 2.3. In section 3, the research methods applied in this study are presented. The two informant groups are introduced in subsection 3.1. This is followed by subsections about the pronunciation tasks (3.2) and the recording process (3.3). The final subsection in section 3 is an introductory paragraph about the visualisations of audio data and the software used to transform the recorded samples into images (3.4). Section 4 presents the results of the analyses of the pronunciation samples. It is divided into four subsections, one for each researched sound: /w/ in 4.1, (inter-)dental fricatives in 4.2, word-final voiced plosives in 4.3, and word-medial prevocalic /t/ in 4.4. Finally, the results are discussed in section 5 and the main conclusions are presented in section 6.

2 Theory

2.1 Age and the acquisition of phonetic structures

The left hemisphere of the human brain is specialised in both production and comprehension of languages. Since using a language is not an innate ability – for human infants are not able to communicate via language from the moment on they are born – it is assumed that this specialisation of the left hemisphere, i.e. the lateralization of the brain, does not begin until early childhood. Even though there are some definite suggestions about the duration of this lateralization process, it is generally assumed to continue until puberty. During this time, also referred to as the critical period of language acquisition, the human brain is said to be optimally prepared for language input and learning a new language (Yule 2010:165). It is presumed that if children are not exposed to languages during this period, the probability that they will later acquire a language is very low (Lightbown & Spada 2006:17). The critical period for language acquisition has also been shown to apply to foreign language learning. For example, a study on the language skill of immigrants to the United States revealed a strong correlation between the age of acquisition and a high level of language proficiency, thus supporting the critical period hypothesis (CPH) in foreign language acquisition (Patkowski 1980).

Some theories concerning the critical period hypothesis suggest that especially the acquisition of the phonetic structures of a foreign language is closely linked with the learners' age. Scovel claims that the CPH only applies to acquiring phonology and that those who start learning a foreign language after this critical period will never attain a native-like accent (1988:185). Even though Scovel does not exclude the small possibility of highly talented or "superexceptional" adult learners, he clearly asserts that the probability of acquiring a pronunciation resembling that of a native speaker after the critical period is very low (1988:181).

Several empirical studies seem to support Scovel's theory. Flege, Munro & McKay (1995) investigated the possible connection between perceived foreign accent and the age of learning. In their experiment, a judge panel of native English speakers assessed the accents of 240 native speakers of Italian on a continuous scale. Even though foreign accents could be perceived among subjects who had begun learning English before puberty, the results of this study strongly support the CPH. According to their results, "virtually all of the NI [native Italian] subjects who began learning English after the age of 15 yr were so [foreign accented] classified" (Flege et al. 1995:3132).

In a later study on the perceived foreign accent of non-native English speakers, Flege, Yeni-Komshian & Liu (1999) researched the pronunciation of native Koreans who had immigrated to the United States of America with the age of arrival in the U.S. ranging between 1 and 23 years. The subjects' pronunciation samples of short English sentences containing a wide range of English consonant and vowel sounds were assessed by a panel of ten English native speakers using a 9-point rating scale ranging from 'very strong accent' to 'no accent'. The results of this study revealed a strong positive correlation between the age of arrival and the proficiency in English. This correlation was more prominent in the pronunciation of the subjects than in the domain of morphosyntax. The results showed, that the earlier the informants had moved to the United States and thus the more they had been exposed to English, the lower the probability of their having a perceivable foreign accent. Flege, Yeni-Komshian & Liu suggest that the different stages in the development of the native language phonetic system affect the acquisition of the phonetic structures of the foreign language by influencing the ways these two phonological systems interact (1999:99).

Another study supporting the CPH in pronunciation acquisition observed the pronunciation of 90 subjects, of which 72 were non-native speakers of English and 18 native English control subjects (Piske et al. 2001). Similar to the study by Flege, Munro &

McKay (1995), the non-native informants were native speakers of Italian residing in Canada. In this study, informants were divided into groups based on their age of learning, i.e. the age of arrival in Canada, gender, as well as the self-reported use of Italian. In addition, factors such as the length of residence in a foreign country and subjects' self-estimated Italian skills. According to the results, gender had little influence on the differences in perceived pronunciation, whereas both age of learning and the use of Italian strongly correlated with the informants' accents. Of these two factors, the effects of the age of learning proved more significant to the subject's accent. Subjects' length of residence and Italian skills were found not to affect the accent independently from the age of learning.

The results of the three reviewed studies provide corroborative evidence for the CPH in pronunciation acquisition. Instead of simply assessing foreign accents, some researchers have devoted their time to finding evidence against the hypothesis. They have conducted experiments to test whether adult, non-native learners, either trained or self-taught, can pass as native speakers. Ioup et al. (1994) tested the language skills of an Englishwoman who had learned to communicate in spoken Egyptian Arabic without instruction in a naturalistic environment. The study aimed to assess not only the speech production of the subject, but also her skill in recognising accents and her knowledge of specific syntactic rules. The subject's spoken language skills were evaluated by 13 teachers of Arabic as a foreign language. In terms of speech production, including pronunciation, intonation, vocabulary, as well as discourse markers, 62 per cent of the judges regarded the subject as native-like which would suggest that it is possible to reach native-like proficiency as an adult learner (1994:80). Nevertheless, it was the instances of non-native pronunciation, such as incorrect pronunciation of certain consonants and words, that caused the remaining 38 per cent of the judges to rate her as a non-native speaker. This, in turn, supports the critical period hypothesis in acquiring foreign language phonology. It could be hypothesised that since the judges had to concentrate on all features of the subject's speech production at once, some non-native features in pronunciation, additional to the ones observed, could have been left unnoticed.

The CPH of pronunciation acquisition states that acquiring a native-like accent after the critical period, i.e. early puberty, is highly unlikely. The results of several studies, some of which were reviewed above, seem to support this theory. Early exposure to the phonology of a language does not exclude the possibility of retaining a foreign accent, though it is shown to increase the level of proficiency. Even though some exceptions

regarding the acquisition of native-like pronunciation after the critical period have been observed, there is a general consensus about the importance of age in the process of acquiring second language phonology. In the light of their study, Bongaerts et al. consider the term critical period as too strict and claim that:

[...] it would seem better to replace the term *critical* period, which excludes the possibility that there are late learners who can learn to speak a second language without a foreign accent, with the term *sensitive period*, which does not does not exclude this possibility and, at the same time, does not deny that there may be biological advantages to an early start. (1995:45)

In her study, Bialystok investigates the differences between adults and children in acquiring a second language. She criticises CPH by stating that “there is insufficient evidence to accept the claim that mastery of a second language is determined wholly, or even primarily, by maturational factors” (1997:116). Yet, based on both empirical and anecdotal evidence, children seem to be more successful in learning a second language than adults. According to Bialystok, this is not because of a biologically based sensitive period or maturational limits, but because the methods for learning change with age. Here she refers to the use of mental categories in learning new things. New features, such as the phonological structures of a second language, can either be included in the existing categories formed by the first language or the learner can create a new category for them. The difference between adults and children is that:

[...] adults tend to extend the existing categories (cf. assimilation) while children tend to create new categories (cf. accommodation). This may be because children are in the process of creating new categories all the time as they are learning new information and this option is natural, while adults are more used to consolidating knowledge and seeking overall similarity. The result, in any case, would be for adults to inappropriately extend first language rules (syntax, phonology, etc.) while children would not. (1997:132)

Regardless of it resulting from maturational factors, e.g. the lateralization, or from different, age-dependant learning methods, there is strong evidence that early exposure to foreign languages (including their phonological structures) can result in higher-level language proficiency. This suggests that exposure to foreign language phonology at an early age, both in instructional and naturalistic settings, benefits the learner in terms of increasing the chances of achieving a more native-like accent.

2.2 Researched sounds

2.2.1 *Voiced labial-velar approximant /w/*

The sound [w], usually spelled with the letter ‘w’ in English, is sometimes referred to as a semivowel (like in Ladefoged 1993:229) for it resembles the English high back vowel [u] phonetically but functions like a consonant. Ladefoged further describes it as a nonsyllabic version of [u] (ibid.). The ‘w’-sound is an approximant which means that the vocal tract is constricted without fully obstructing the air stream. [w] is a sonorant which means that this constriction of the vocal tract is not as narrow as to cause turbulence. Most sounds typically have a single place of articulation i.e. the vocal tract is narrowed only in one place. The sound [w] is a doubly articulated consonant i.e. it is formed in two simultaneous primary places of articulation: the lips are rounded and protruded (labio) and the back of the tongue is raised towards the soft palate or velum (velar) (Ashby & Maidment 2005:40). The vowel-like characteristics of [w] are clearly visible in the spectrograms: similar to vowels, its formant structures are readily distinguishable (Johnson, 2003:138). In the spectrographic analysis, the sound [w] can be identified best by the rising second formant (Ladefoged & Disner 2012:53). Johnson further describes the [w] as “a glide leading into the vowel” (2003:139) and argues that because of this, the duration of the [w] sound is difficult to measure. One cannot accurately define where the [w] stops and the following vowel begins. He describes two options for separating [w] from the vowel: either the rising second formant is seen as a part of the vowel, or the vowel begins only when the formants are more or less in a steady state. In the spectrograms of this study the latter options is used to mark the phoneme [w]. The sound [w] only appears in word-initial and -medial positions preceding a vowel sound, even though the corresponding letter ‘w’ also exists in word-final positions like in the words *law* ([lɔː]) and *row* ([ɹəʊ] or [ɹaʊ]).

2.2.2 *Voiceless and voiced (inter-)dental fricatives /θ/ and /ð/*

Unlike approximants, fricatives are produced by obstructing the air flow with the articulators so that the air is forced through a narrow passage. As a result the air flow becomes turbulent and audible friction can be heard (Roach 2009:39). In the English language, there are two fricative sounds that are both spelled with the letter combination ‘th’. These sounds, [θ] and [ð], are non-sibilant fricatives that are produced by placing the tongue (active articulator) between (inter-) or more commonly behind the upper front teeth (passive articulator). According to Ladefoged and Maddieson, the interdental

pronunciation is more common in American English than in the English and Scottish accents (1996:143). They investigated the position of the tongue when pronouncing the (inter-)dental fricatives: 90% of the Californian college students produced the sounds interdental, whereas the majority of the English and Scottish informants (90%) pronounced the sounds by placing their tongues behind the upper incisors. [θ] is produced voicelessly i.e. the vocal cords are not vibrating when it is pronounced. During the pronunciation of [ð], the voiced counterpart, the vocal cords are simultaneously vibrating. The [ð] may also be pronounced with less voicing or voicelessly, especially in word-initial and -final positions (Roach 2009:40). According to Roach, the English voiced (inter-)dental fricative has often so little friction noise that it resembles “a weak dental plosive” (2005:45). The voiced sounds are often referred to as lenis i.e. weak (meaning short, unaspirated and sometimes voiced) for they are not necessarily always clearly voiced (Ashby & Maidment 2005:95). In comparison, the word fortis (strong, loud, aspirated and always voiceless) is used when referring to voiceless consonants. The vowels preceding [ð] are often noticeably longer than the ones followed by the fortis fricative [θ] (Roach 2009:40). This vowel-shortening effect of voiceless consonants is relatively common in English and it is called pre-fortis clipping (Ashby & Maidment 2005:97).

Both the voiceless and the voiced (inter-)dental fricatives are weak sounds with relatively low amplitude. The energy is spread over a wide range of frequencies, although according to Ladefoged and Disner, this energy is centred in a higher frequency range than the energy of the equally weak labiodental fricatives [f] and [v] (2012:56). Additionally, the formants can sometimes be used to distinguish these sounds from others. The fourth formant of the adjacent vowel of [θ] and [ð] is often found above 4000Hz and the second formant is “fairly level at around 1,250 Hz” (Ladefoged & Disner 2012:57). The voiced (inter-)dental fricative can be differentiated from its voiceless counterpart by the vertical striations in the spectrograms and the periodic waveform. Ladefoged and Disner also mention that the differences between (inter-)dental and labiodental fricatives are often very small, which is why “it is not very surprising that English is one of the few languages that uses both of these sounds” (ibid.). Both (inter-)dental fricative sounds can appear in all three positions: word-initially, -medially and -finally.

2.2.3 Voiced plosives /b/, /d/, and /g/ in word-final position

Plosives are also known as stops because at one point of the pronunciation the air flow through the vocal tract is completely stopped. In producing a [b], a bilabial plosive, both

lips are pressed together to block the air. In the case of a [d], the blockage is created by touching the alveolar ridge with the tip of the tongue (alveolar plosive). In pronouncing a [g], the back of the tongue makes contact with the soft palate (velar plosive) thus stopping the air flow. A stop sound consists of at least three phases: the closing phase (closure), the compression phase and the release phase. English voiceless stops may include a post-release phase in which the release burst is followed by an extra release of air (aspiration) (Plag et al. 2007:46). All of the three sounds described in this subsection are voiced, which means that the vocal cords vibrate during the release burst and often during the compression phase. This may be visible in the spectrograms as a low-frequency voicing bar and as periodicity in the waveforms. According to Roach, if the word-final voiced plosives are voiced, the voicing is often only present at the beginning of the compression phase (2009:28). The burst that immediately follows the release can be very weak and sometimes the word-final stops are unreleased when said in a normal, unemphatic way (Ladefoged & Disner 2012:52). Like with the (inter-)dental fricatives described above, these lenis plosives may also be pronounced with weak voicing or even voicelessly. In these cases, the duration of the preceding vowel is the only way to distinguish between voiced (lenis) and voiceless (fortis) plosives. The vowels preceding /p/, /t/ and /k/ are “much shorter” than vowels that are followed by lenis stops (Roach 2009:28). Figure 1 presents the differences between voiced and voiceless pronunciation of the alveolar plosives. In addition to the longer duration of the preceding vowel, a voicing bar is also visible in the spectrogram of the voiced sound. The English voiced plosives are never aspirated, so in addition to the vowel duration, visible aspiration in the spectrograms is a sign of a voiceless plosive.

2.2.4 Voiceless alveolar plosive /t/ in word-medial prevocalic position

In RP, these post-stress prevocalic voiceless plosives are slightly aspirated and pronounced as a [t^h]. Similar to the voiced [d] above, [t^h] is produced by first blocking the air flow with the tongue touching the alveolar ridge, then releasing it with a burst of additional air. Commonly in GA, and also in some other varieties, this type of a /t/ is pronounced differently. Roach calls this American style of pronouncing a /t/ in word-medial prevocalic positions the “flapped r” (2009:164), Ladefoged and Maddieson simply refer to it as a flap (1996:231), although earlier Ladefoged refers to it as a tap (1993:168) and Ashby and Maidment describe it as a voiced alveolar tap (2005:59). In this paper, the word ‘flap’ ([ɾ]) is used to refer to this allophone of the word-medial prevocalic /t/. Although the names

vary, the manner of pronouncing is generally described as follows: the tip of the tongue makes a very brief contact with the alveolar ridge. According to Ashby and Maidment, the manner of pronouncing a flap is identical to that of a voiced alveolar plosive [d] except that it is noticeably faster (2005:59). Unlike during the compression phase of plosive sounds, there is not enough time for the air pressure to increase during the closure of a flap. As a result, the release burst after the short closure is very weak and often not visible.

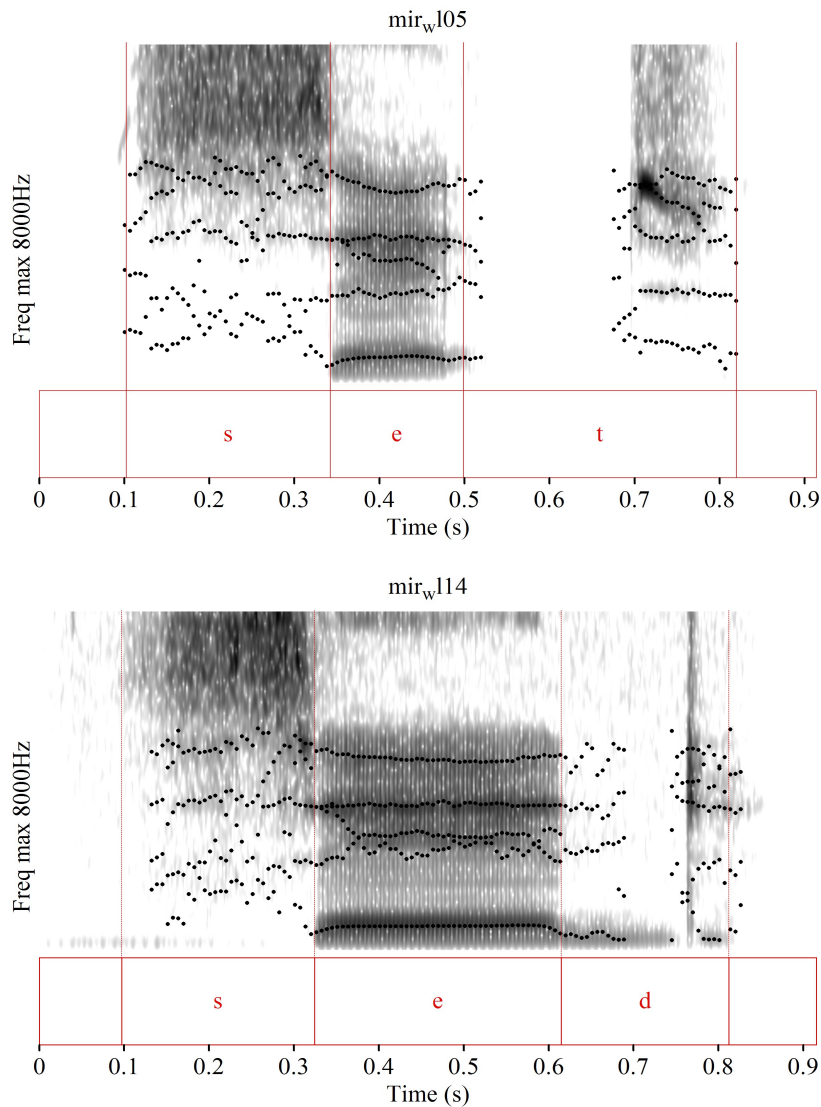


Figure 1: Spectrograms of the word list words 5 and 14 as pronounced by the author. The words *set* (above) and *said* (below) are transcribed in red. Formants for vowels and sonorants are marked with black dots.

2.3 The first languages and the stereotypical pronunciation of English

Several theories, including the theory of natural phonology by Stampe (1979) and that of Bialystok described in section 2.1, suggest that the structures of the first language (L1; mother tongue) and of other previously learned languages affect the acquisition of new foreign languages. The influences of L1 on phonetic structures and phonology is assumed to be greater, for

[...] the acquisition of a new sound structure includes too the learning of new patterns of articulation and perception. Such patterns involve physiological aspects of language behaviour, which, it is assumed, are more resistant to change and adjustment than the more ‘cognitive’ aspects of language behaviour associated with knowledge of higher levels of linguistic structure, such as syntax. (James, 1988:30)

It can thus be assumed that the phonological structures i.e. the suppressed and applied processes of the L1s, Finnish and Austrian German, have influenced the informants’ acquisition of English phonology. In order to be able to distinguish typically Finnish or German features in the samples, the phonological structures of the two L1s and the stereotypical accents related to the investigated sounds are introduced and discussed in the following two subsections.

2.3.1 Finnish

Both the labiovelar approximant and the (inter-)dental fricatives are not part of the Finnish sound system, although it is mentioned that the sound [w] may occur following diphthongs ending with a [u] as in *vauva* ([vauwa]) ‘baby’ (Suomi, Toivanen & Ylitalo 2008:31). Also the voiced plosives /b/ and /g/ are considered ‘foreign’ for they only appear in loanwords such as *bussi* ‘bus’ and *graniitti* ‘granite’. In the stereotypical Finnish accent, the English [w] is often replaced with the sound corresponding to the Finnish letter ‘v’ which is pronounced as a labiodental approximant [v]. This Finnish sound is sometimes also used in pronouncing English words containing the letter ‘v’ (Tyndall 2013:6). Actually, the name of the letter ‘w’ in Finnish, *kaksois-v* ([kaksois-ve:], ‘double v’), can be seen to indicate the Finnish-style pronunciation of the phoneme /w/. A visualisation of the pronunciation of the Finnish letter ‘v’ is presented in Figure 2.

Finnish has a phonemic orthography, which means that the spelling or the written symbols (graphemes) correspond to the phonemes of the language. In other words, Finnish is spoken as it is written. This may be the reason why the (inter-)dental fricatives are stereotypically pronounced as a Finnish-style voiceless unaspirated dentalalveolar plosive [t̪] (Figure 2) which corresponds to the Finnish letter ‘t’. In trying to include the letter ‘h’

in ‘th’ into the pronunciation, /θ/ and /ð/ may both be pronounced with a voiceless aspirated alveolar plosive [t^h]. In addition, Finnish speakers of English sometimes pronounce /ð/ as [d] or even as a flap (Tyndall 2013:6). Additionally, in less ‘extreme’ Finnish accents of English, [ð] is often devoiced and replaced with [θ] (Korpela 2015: chapter 40/4).

As mentioned above, /b/ and /g/ do not originally occur in the Finnish language but have entered it through loanwords (Suomi, Toivanen & Ylitalo 2008:35). These sounds are pronounced as voiced counterparts of the bilabial and the velar plosives. Some speakers of Finnish may pronounce these sounds voicelessly, e.g. words like *baletti* ‘ballet’ and *grilli* ‘grill’ as [palet̪i] and [kril̪i]. Suomi, Toivanen & Ylitalo describe the Finnish [d] as a voiced apical alveolar. According to them, its manner of pronunciation is something between a plosive and a flap. Unlike [b] and [g], [d] does appear in native Finnish words, though only word-internally. In this study, only the pronunciation of word-final voiced plosives is investigated. The only plosive that can occur word-finally in fully native Finnish words is the voiceless dentalveolar [t̪]. This might suggest that there is a tendency of devoicing the word-final plosives. In describing the common ‘mistakes’, Korpela generalises by stating that “[the] [c]ontrast between voiced stops [...] and unvoiced stops [...] tends to be insufficient: the voiced stops are incompletely voiced or even unvoiced [...]” (2015: chapter 40/4). The pronunciation of both the voiceless non-aspirated dentalveolar plosive [t̪] and the voiced alveolar plosive [d] is illustrated in Figure 2. The spectrogram of a [d] resembles the flap in Figure 64, although the release burst is more visible.

The fourth sound researched in this paper is the word-medial prevocalic voiceless alveolar plosive /t/. In Finnish, the letter ‘t’ commonly occurs in word-medial prevocalic positions, e.g. in words like *lato* ‘barn’, *äiti* ‘mother’ and *lauta* ‘plank’. As described above, the sound corresponding to the letter ‘t’ in Finnish is a voiceless dentalveolar plosive [t̪]. This sound is generally not aspirated. The only observations of aspiration in Finnish plosives have been reported by Suomi: some of his informants pronounced the [t̪] in the sequence /t̪ n/ with aspiration (2009:406). Based on this information, a person with a strong Finnish accent would pronounce the /t/ in the English words such as *city* and *thirty* with a [t̪].

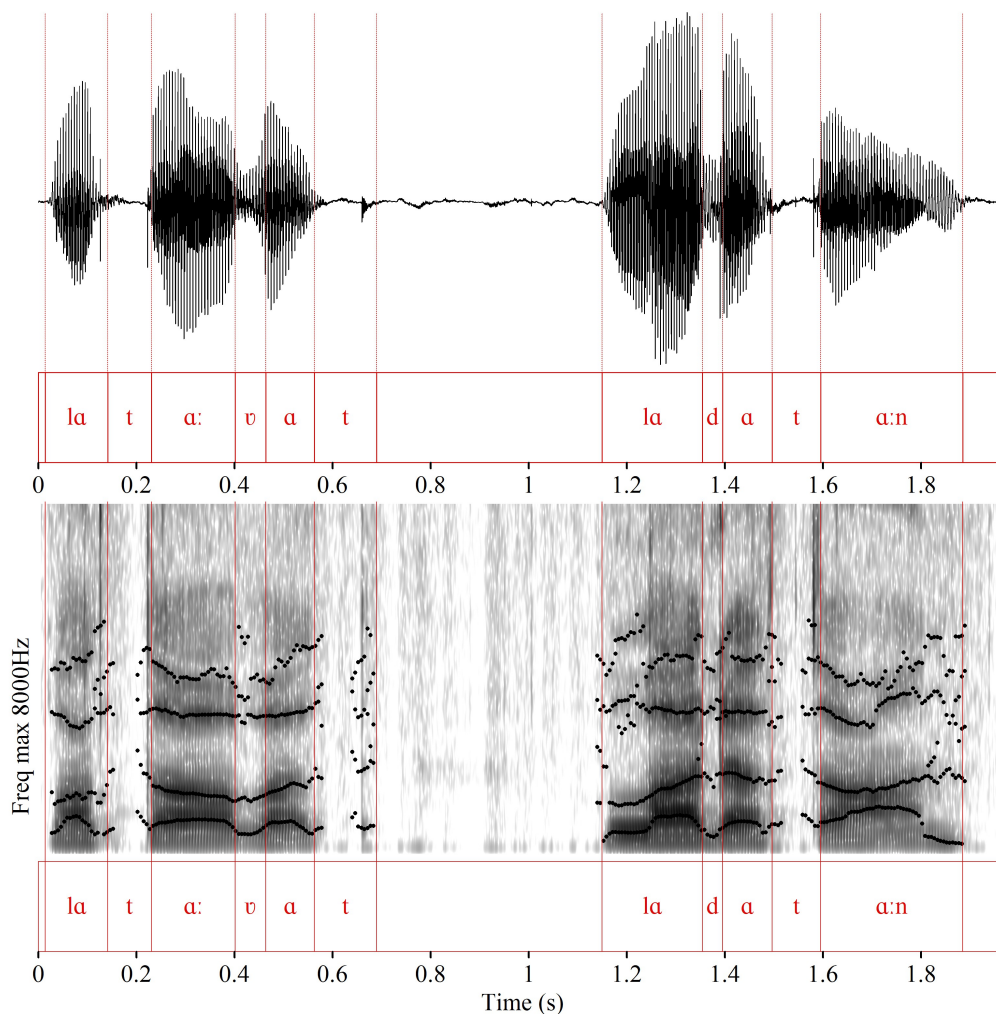


Figure 2: The waveform (above) and the spectrogram (below) of the Finnish words *lataavat* ‘(they) load’ (3rd person plural present tense) and *ladataan* ‘is/are loaded’ (present passive) as pronounced by the author. The words are marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

2.3.2 Austrian German

Originally and in its ‘purest’ form, the consonant inventory of Finnish consisted only of 11 consonant sounds. German has more consonant sounds and many of them resemble the English phonemes. Similar to Finnish, the labiovelar approximant [w] does not occur in standard German or standard Austrian German. The letter ‘w’ is widely used, though it corresponds with the phoneme /v/. German words spelled with a ‘w’ are thus pronounced with a voiced labiodental fricative [v]. In the typically German accent of English, the /w/ is pronounced as a [v] (Kresta 2015:127).

Like [w], the (inter-)dental fricatives are also not part of the standard German (or Austrian German) consonant sounds, even though the letter combination ‘th’ occurs at

least in place names (e.g. *Thüringen*, *Thalgau*) as well as in loanwords (e.g. *Sympathie* ‘sympathy’, *Theorie* ‘theory’). These words are all pronounced with an aspirated voiceless alveolar plosive [t^h]. Whereas the Finns, presumably influenced by the orthography, stereotypically change the manner of articulation and produce a [t̪] when trying to pronounce the ‘th’-sounds, a German-speaker stereotypically alters the place of articulation by shifting the tongue backwards thus pronouncing /θ/ and /ð/ as alveolar fricatives [s] and [z] (Kresta 2015:127).

According to Biersack, the voiced plosives of English and German can be described as identical (2002:57). Like in English, the German voiced plosives can be pronounced voicelessly. There is one significant difference: when unvoiced, the English voiced plosives preserve their lenis character i.e. they remain weak and non-aspirated and the duration of the preceding vowel continues to be longer than when followed by a fortis plosive. In German, the process of devoicing the word-final voiced obstruents fully neutralises the contrast between voiceless (fortis) and voiced (lenis) sounds (Plag et al. 2007:41). Unlike in English, the voiced sounds lose their lenis character and become fortis. The German voiced plosives /b/, /d/ and /g/ are commonly pronounced as [p^h], [t^h] and [k^h] in word-final position (Biersack 2002:57). It can thus be assumed that a person with a strong German accent applies this rule when pronouncing English words such as *cab* like [k^hæp^h], *said* like [set^h] and *blog* like [blɒk^h]. In comparison to standard German, the process of word-final devoicing is not that strong in standard Austrian German. According to Muhr (2000:47), the plosives /b/, /d/ and /g/ are only slightly strengthened (“leichte Fortisierung”) and either weakly or not at all aspirated. The results of Ehrlich’s empirical study on standard Austrian German pronunciation support Muhr’s definition (Ehrlich 2009:96). At least 75% of all word-final voiced plosives were pronounced with aspiration, and around 80% of the samples showed signs of weaker word-final devoicing. In Figure 3 the visualisations of the Austrian pronunciation two German words, *Rad* ‘wheel’ and *Rat* ‘advice, council’, are presented. There is only a small difference in the length of the preceding vowels (0,03 seconds) and both word-final plosives were pronounced with aspiration.

The fourth and the last of the investigated sounds is the word-medial prevocalic /t/. Biersack states that the voiceless plosives in German generally occur aspirated in all positions and claims that aspiration is the main distinction between German voiced and voiceless plosives (2002:57). According to Ammon et al., these sounds are also aspirated in standard Austrian German (2004:LVII). All words in Ehrlich’s study with a word-

medial prevocalic /t/ were pronounced with weak aspiration by the Austrian informants (2009:95). It can thus be presumed that Austrian German speakers, who fully rely on the pronunciation rules of their first language, would pronounce the English words *city* and *matter* with aspirated voiceless alveolar plosives.

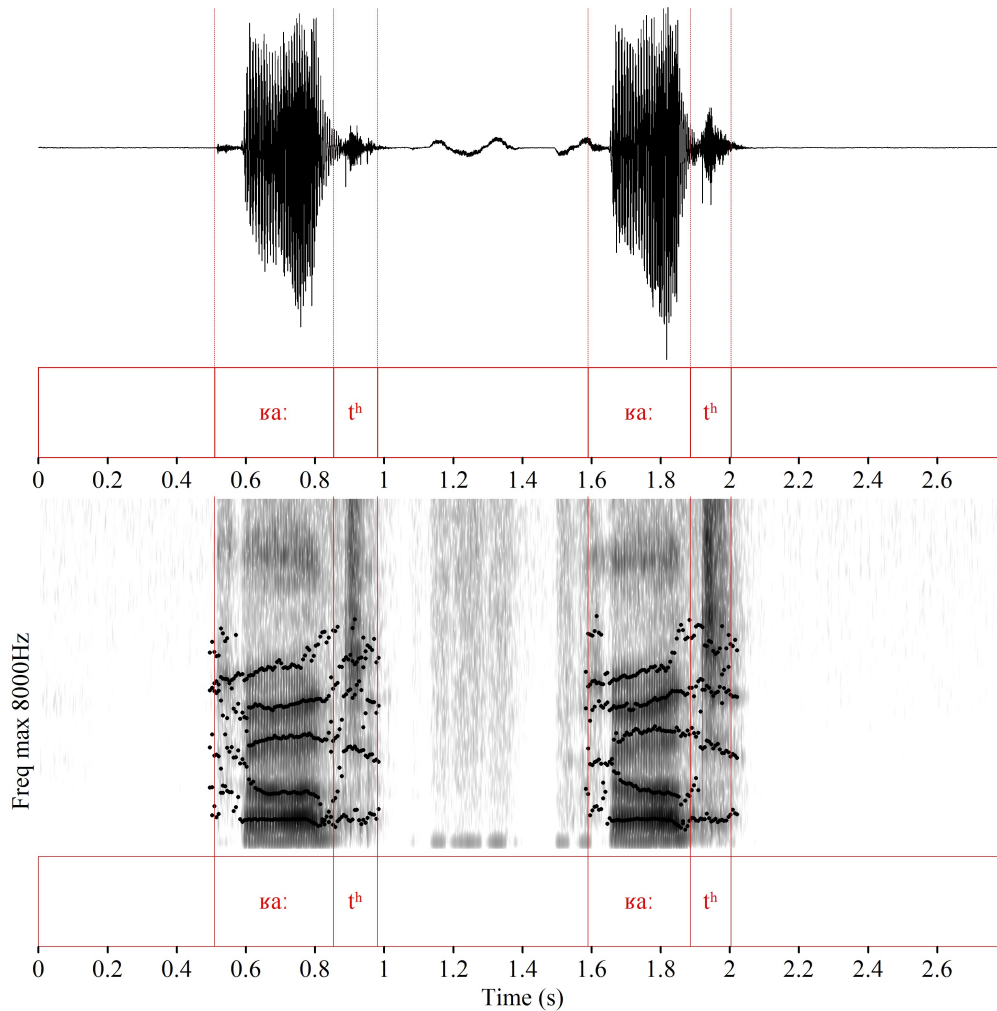


Figure 3: The waveforms (above) and the spectrograms (below) of the German words *Rad* ‘wheel’ (left) and *Rat* ‘advice, council’ (right) as pronounced by informant A22. The words are marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

3 Methods

3.1 Informants

The informants for this study were selected on the basis of three main criteria: first language, age and level of education. In this study the English pronunciation of native Finnish-speaking Finns and Austrians is investigated, so the first language of the

informants was either Finnish or (Austrian) German. The age range of the informants was strongly influenced by the practical issues related to data collection. Since acquiring pronunciation samples from total strangers proved almost impossible, the majority of the informants are either friends or acquaintances and belong to the same age group as the author. This age group, birth year 1983 ± 4 years, was thus chosen as the age range for this study and consists of informants born between 1979 and 1987. The age difference between the youngest and the oldest informant is thus 8 years. In addition to the two criteria described, all suitable informants were required to have at least a bachelor's degree, either from a regular university or from a university of applied sciences (hence UAS; 'Fachhochschule' in Austria and 'ammattikorkeakoulu' in Finland).

All collected samples were divided into two groups based on the L1 of the informant. All informants were asked to fill in a short questionnaire about their personal details and English use (Appendix D). The first group, referred to as Group 1 or with the words 'Finns' and 'Finnish', consists of 25 Finnish informants. Originally, recordings from 26 informants were received, but the quality of sample 24 was so low that it had to be excluded from the analysis. A typical Group 1 informant was born in 1982 and has a master's degree from a regular university. Four out of 25 have finished their doctoral degrees and seven of the Finnish informants have a degree from an UAS. Due to the relatively uniform school system in Finland, most Group 1 informants started learning English at the age of 9 (average value 9,3 years) and had a total of 10 years (average value 9,8 years) of English tuition at school (primary and secondary i.e. Finnish comprehensive school and upper secondary school, excluding institutions of higher education). Additionally, two thirds of Group 1 informants had at least one English course during their studies and eight of them has spent some time in an English-speaking country. According to their answers, Group 1 informants use their English language skills related to pronunciation, i.e. speaking and listening, at least on a weekly basis. Most informants actively listen to English every day and speak English at least once a week. Based on the estimations of Group 1 informants, most of them communicate in English either with non-native speakers of English (52%) or equally with both non-native and native speakers (44%). More details on Group 1 informants can be found in Appendix E.

The second group (hence Group 2 or 'Austrians' and 'Austrian') consists of 21 Austrian informants. Similar to Group 1, one sample (number 6) had to be excluded from the analysis due to poor quality. Group 2 is slightly younger than Group 1. A typical Austrian informant was born in 1984. Similar to Group 1, most Austrian informants have a

master's degree from a regular university. In contrast to Group 1, Group 2 is more homogeneous in terms of education. Less Group 2 informants have degrees from a UAS and only one has a doctorate degree. Typically, a Group 2 informant started learning English at the age of nine (average 8,9 years), although there is more variety between individual Group 2 informants. The average amount of English tuition at school is 9,1 years, which is almost a year less than that of Group 1. Similar to Group 1, the Austrians actively listen to English every day and speak in English on a weekly basis. Based on their own estimations, almost two thirds (62%) of Group 2 informants communicate in English mostly with non-native speakers of English. 33% claim to communicate in English equally with both non-native and native speakers. Details on individual Group 2 informants are presented in Appendix F.

The main difference between the two groups is the informants' L1 (or nationality). The groups also differ in terms of the total amount of English tuition at school as well as the exposure to English in informal situations. In Finland since 1994, average learners learn English as their first foreign language (hence L2; known as A1-language in Finland) a total amount of 24 weekly lessons per year (Ranta 2004; Könölä et al. 2001:37). One weekly lesson per year equals to 38 regular lessons, so the average Finnish learners of English have a total of 912 English lessons during their school careers (comprehensive school grades 3-9 and 3 years of upper secondary school i.e. 'lukio'). Between 1985 and 1994, the amount of weekly English lessons per year was 23. The total number of English lessons for average Austrian learners, who learned English as their L2 in the 1980's and 90's and went to a grammar school ('allgemeinbildende höhere Schule' or 'AHS') after four years of primary school ('Volksschule'), is 870 regular lessons or 29 weekly hours ('Wochenstunde'; equals to ca. 30 lessons) (de Cillia and Krumm 2010:155-156). When the length of a regular lesson (45min in Finland and 50min in Austria) is taken into account, the Finnish informants learned English at school on average a total of 684 hours and the Austrians 725 hours.

Both in Finland and in Austria, English is taught as a foreign language and the model language for teaching is that of a native speaker (Ranta 2004:35; Hebenstreit 1997:37). In Finland, British English was mentioned as the official target variety to be taught in Finnish schools in the curricula between 1960's and 80's, though since 1994, a specific target variety has not been named in the general curricula for the comprehensive and the upper secondary schools (Ranta 2004:36). Similar official target variety specifications could not be found for Austria, but according to Hebenstreit's study most

Austrian teachers (at the time the study was conducted) speak or try to speak with an RP accent (1997:61). Additionally, a relatively recent study revealed that the majority of Finns consider British English as the most pleasant of all English varieties (Leppänen et al. 2009:56). According to a study by Dalton-Puffer, Kaltenböck and Smit, RP remains the preferred English accent in Austria (1997:126).

Another difference between the Finnish and the Austrian informants is the exposure to spoken English in informal settings i.e. outside the classroom. Since the late 1980's, when the Finland's first commercial TV channel was founded (Ruhanen 2009), and throughout the informants' critical/sensitive period for language learning generally in the 1990's, the Finnish informants have increasingly been exposed to English through TV-programmes and films. Although some children's programmes and films are dubbed, most foreign films and TV programmes are generally subtitled (Safar et al. 2011:6-8). As a result, a large proportion of the broadcast material is in a foreign language and it can be assumed that not only the viewers watching, but also the people within the range of the sound, are exposed to the language, either deliberately or subconsciously. Since the majority of the foreign programmes and films originate in North-America, the variety to which the viewers are exposed is GA or any of the other American Englishes (Tilastokeskus 2005). In Austria, films and TV programmes are predominantly dubbed i.e. in German (Safar et al. 2011:6-8), so it can be assumed that during childhood and adolescence, the Austrian informants of this study were mostly exposed to English in institutional settings only.

3.2 Pronunciation tasks

Pronunciation tasks used in studies on pronunciation are often classified in terms of formality. According to Labov's principle of attention, the more attention the speakers pay to their speech, the more "frozen" or "ritualistic" the language (1972:112). Reading out a list consisting of pairs of words that differ phonetically only by a single phoneme is considered to be the most formal and thus most artificial type of testing pronunciation. These so called minimal pairs force the informant to pay special attention to the pronunciation of each phoneme in order to distinguish the words as separate lexical items. The second most formal way to examine pronunciation is the word list task. In comparison to minimal pairs, reading out a word list of seemingly random words requires less attention since the phonetic differences between the words are greater. Both of these tasks, minimal pairs and word lists, consist of single words pronounced separately without any or only

with little interference from other words. This often results in clear articulation of each individual phoneme. At the other end of the formality scale of pronunciation tasks is the free conversation. In an ideal situation, free conversation equals the informant's natural way of speaking with all the additional noises and other 'imperfections'. More importantly, the attention to the correct pronunciation of separate phonemes decreases as a result from an increase in speed. In addition, the informants' attention is increasingly focused on the content of the conversation instead of correct pronunciation. Formality-wise, reading a text passage is situated in between the word list task and free conversation. It includes prosody and interaction between words that a word list lacks. Similar to free conversation, the speed is greater than when reading aloud a list of words. When reading a text passage, the amount of attention the reader pays to the smaller units of language decreases as the reader focuses on larger entities such as sentences and rhythm instead of individual words and their pronunciation.

Of the tasks of different formality levels described above, the word list and the text passage tasks were used in collecting the data for this study. Although free conversation would have provided the most realistic idea of the informants' pronunciation, the decision of excluding it from the tasks was based on following reasons. The primary reason for excluding free conversation from the test methods was the high amount of the researched sounds. It would have been difficult to obtain enough occurrences of all sounds from free conversation or even from an interview. Even when asking specific questions to guide the informants into using certain words containing the wanted phonemes, the uncertainty of collecting enough comparable data would have been relatively high. In the case of a word list and a text passage, the amount of occurrences could easily be controlled. Another reason for choosing these two pronunciation tasks is the Observer's Paradox (Labov 1972:113). Regardless of the task type, the situation of collecting data by using a recorder invokes a level of formality thus affecting the language used by the informants. To attain the most realistic data, the samples should be collected without the informants' being aware of it. In addition to the obvious ethical issues arising from recording people's speech without their permission, obtaining such data for this study would have been very difficult. Finding suitable situations to record enough Finns and Austrians of very specific backgrounds and age groups speaking English, a foreign language, would have practically been impossible within such a limited time frame.

Both pronunciation tasks, the word list and the text passage were created specifically for the purposes of this study. The word list consists of 39 simple, generally

well-known words of which all but three are either mono- or bisyllabic (Appendix A). The list contains at least two instances of each researched phoneme in all possible positions. In addition to words containing one or more of the sounds investigated in this paper, some minimal pairs, like *wag* and *whack*, were included. The minimal pairs are not introduced to the informants as such, i.e. the pairs are separated and simply listed along with the other words. By not highlighting the minimal difference between these words, the formality of the task is assumed to decrease to the same level as that of a word list. In some words, the minimal pair -like structures are less obvious. These words were included in the list for having phonetic structures (syllables) that differ from the syllables containing one of the researched sounds only by one phoneme, e.g. [weɪ] in *waiter* and [veɪ] in *vein*, or [θɪn] in *thin* vs. [fɪn] in *Finland*. The text passage is a short email or a letter consisting of 195 words (Appendix B). The choice of an everyday-topic and relatively casual style aim at making the informants feel more at ease thus decreasing the situation's level of formality. Similar to the word list, the text passage contains as many instances of the investigated phonetic structures as possible.

3.3 Recording procedure

Collecting a large enough an amount of data for this study would also have been difficult and very time-consuming, had all the samples been collected in face-to-face situations. Thanks to modern technology, most of the informants recorded their samples independently using an online recorder (123apps LCC 2015). Some of the samples were collected personally. To minimize the influence of the Observer's Paradox, these informants were asked to record their samples in private i.e. in another room. From a negative point of view, this uncontrolled recording procedure gave the informants a chance to familiarise themselves with and to practice the pronunciation tasks before recording their samples. This may have influenced the data making the pronunciation of some informants more artificial. Additionally, resulting from a variety of microphones used in recording the samples, the collected audio data varies considerably in quality. Of a total of 48 samples, two had to be discarded due to quality issues. Regardless of these negative aspects, this type of independent recording was considered the most practical way to gather the required amount of audio samples for this particular paper. Recording the samples using online software and collecting them via email was practical and time-saving for both parties. Instead of trying to set appointments for a meeting, the informants could choose the most suitable time and location themselves. Additionally, thanks to this arrangement,

any extra travelling costs could be avoided. In addition to these more practical aspects, allowing the informants to record their pronunciation samples unobserved can be seen as beneficial to the quality of the collected data. According to Labov “[a]ny systematic observation of a speaker defines a formal context where more than the minimum attention is paid to speech” (1984:29). Based on this methodological axiom, it can be assumed that the absence of an interviewer/observer decreased the formality of the situation and resulted in more realistic pronunciation.

3.4 Visualisations of audio data

The pronunciation samples collected from the informants were processed and analysed with version 5.4.22 of Praat, a programme specially developed for linguistic analysis of audio data. The software package of Praat was created by Paul Boersma and David Weenik of the University of Amsterdam (2015). Praat is also maintained by them, and new, improved versions are launched occasionally. Versions of the programme for all the main operating systems are available online at praat.org and can be downloaded free of charge.

The pronunciation analysis is mainly based on spectrograms (or sonograms) and on the waveforms of the recordings. In order to analyse the visual structure of the pronunciation samples, the sound files (.mp3 or .wav) were visualised as waveforms and transformed into spectrograms using Praat. Both of these are visualisations of acoustic signals. Spectrograms represent the amplitude of the sound waves, i.e. the loudness, in light-to-dark colour scale: the higher the amplitude, the darker the spectrogram. The frequency (or the pitch in Hz) of the recorded sample forms the vertical axis of the graph and time in seconds (s) is represented on the horizontal axis. Spectrograms of certain sounds (vowels and other sonorants) contain horizontal dark bands. These bands are called formants and they are “concentration[s] of acoustic energy around a particular frequency in the speech wave” (Wood 2005). Formants can be used to distinguish between different sounds which is why Praat has an algorithm that has been programmed to detect and mark the maxima in the sound spectrum i.e. the formants. Sometimes this algorithm falsely detects formants also in obstruent sounds. For example, Praat detected and marked formants also for the aspirated voiceless plosive in the word *wet* in Figure 4. The other visualisation of audio data used in this study, the waveform, presents the amplitude of the sound file as a function of time. Waveforms are especially useful for assessing the voicing of the sounds: the waveform of voiced sound is periodic, whereas the voiceless sounds appear as aperiodic i.e. irregular, zigzag-like waves.

4 Results

In the following four subsections, the main results of the analysis are presented in detail. The possible reasons that influenced the pronunciation and lead to these results are mostly discussed in section 5.

4.1 Voiced labial-velar approximant /w/

The analysis of the pronunciation of /w/ was purely based on the characteristic structure of /w/ in spectrograms. According to Ladefoged and Disner, the “most conspicuous aspect” of the /w/ is the rising second formant (2012:53). The steepness of this rise depends on the height of the second formant of the following vowel. These differences in steepness are presented in Figure 4, which shows the example spectrograms of /w/ preceding five different vowels in words *we*, *wet*, *war*, *woo* and *were*.

The analysis of the pronunciation samples revealed that the Finnish informants pronounced /w/ in a standard-like manner more often than the Austrians in both pronunciation tasks (Figure 5). The Finns produced a [w] in 83% of all the instances in both the word list and the text passage samples. In the Austrian word list samples, /w/ was less frequently pronounced as a [w]: 70% of the occurrences of /w/ were pronounced as a [w] and in the remaining 30% of the cases /w/ was pronounced as a [v]. In the Austrian text passage samples i.e. in the less formal task, the percentage of producing a [w] increased slightly by 7%. This difference between the two informant groups is significant at the 5% significance level ($\alpha=0,05$) for the p-values of the χ^2 -test for both the word list and the text passage data are smaller than the significance level α (Table 1). In other words, there is a significant correlation between the informant group and the pronunciation of /w/.

Figure 6 presents the percentages of pronouncing /w/ as [w] when preceding both rounded and non-rounded vowels in word-initial and -medial positions. In the word list samples, the Finnish informants produced a [w] more often when the /w/ preceded a non-rounded vowel. The word-initial /w/ was pronounced in a standard-like manner slightly more frequently than the /w/ in word-medial positions. Although the percentages of standard-like pronunciation suggest that the Finnish informants’ pronunciation of /w/ would depend on its position and the roundedness of the following vowel, this correlation is not statistically significant at the 5% significance level (Table 1). Thus, based on the Group 1 word list samples, the null hypothesis, i.e. producing a [w] is independent of both

the position of /w/ and the roundedness of the following vowel, cannot be discarded. The word list samples of the Austrian informants differed from the Finnish samples in terms of producing a [w] in different positions as well as when followed by a rounded or a non-rounded vowel. Words with a word-initial /w/ were more commonly pronounced with a [w] when the following vowel was non-rounded, whereas the word-medial /w/ was pronounced as a [w] in 81% of all occurrences in which the /w/ preceded a rounded vowel. Unlike the p-value of Group 1, the p-value of Group 2 samples is smaller than the significance level of $\alpha=0,05$ which means that Group 2 pronunciation of /w/ in the word list task correlates with the position of /w/ and the roundedness following vowel (Table 1).

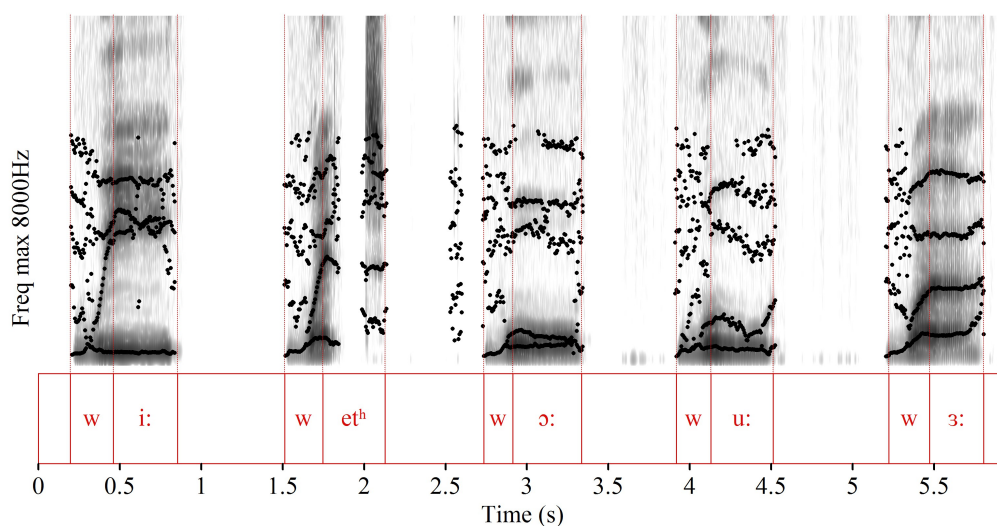


Figure 4: The spectrograms of the words *we*, *wet*, *war*, *woo* and *were* as pronounced by the author. These words are marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

In comparison to the word list task, in which the only similarity between the groups was the preference of producing a [w] in word-initial position when preceding a non-rounded vowel, the text passage samples resemble each other in terms of pronouncing /w/ as a [w] in different positions and with different roundedness of the following vowel (Figure 6). Both groups pronounced /w/ as a [w] more frequently in word-initial position: of all word-initial occurrences of /w/, 86,3% were pronounced as a [w] by the Finns and 80,0% by the Austrians. These percentages are significantly lower for the word-medial occurrences: Finns produced a [w] only in 72,0% and Austrians in 69,0% of the occurrences. Similar to the position of /w/, the roundedness of the following vowel seems to affect the pronunciation of /w/. In the text passage task, both groups produced a [w]

more frequently when /w/ was followed by a rounded vowel (Finns: 86,0%; Austrians: 84,5%) though the Austrians's preference of the rounded vowels was more noticeable. 73,8% of all instances of /w/ preceding a non-rounded vowel were pronounced as [w] by the Austrians whereas the same percentage for the Finns is 81,7%. This preference of word-initial /w/ followed by a rounded vowel in the text passage task is also statistically significant for both groups (Table 1).

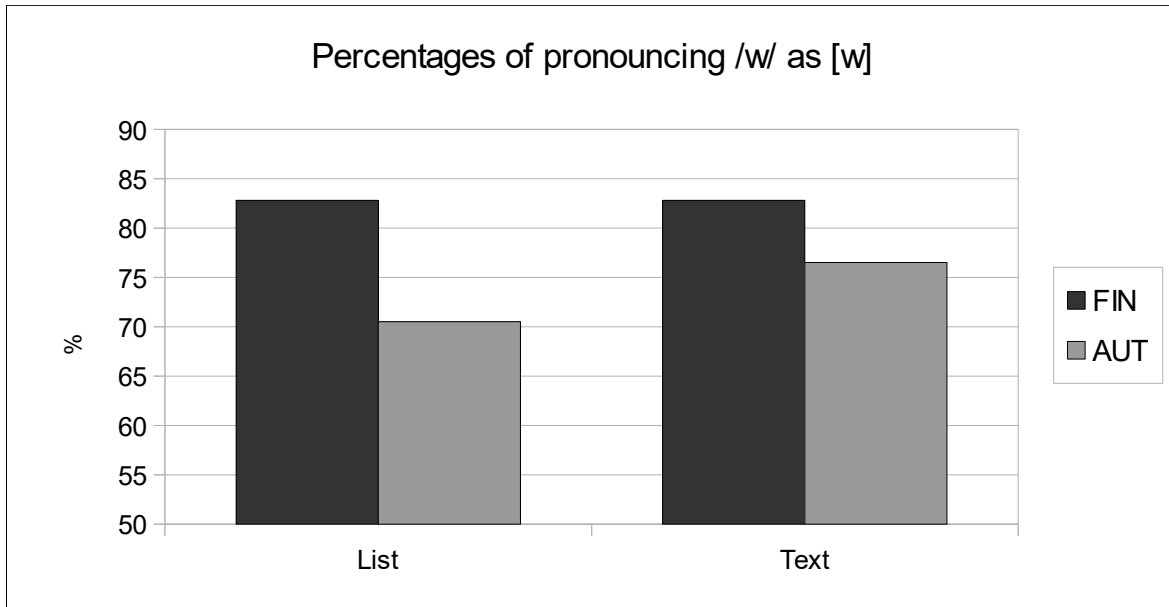


Figure 5: The percentages of pronouncing /w/ as [w] in both tasks: word list (left) and text passage (right).

Variables (Controlled x measured)	p-value from the χ^2 -test ($\alpha=5e-02$)
Group x Pronunciation (LIST)	2,50e-03
Group x Pronunciation (TEXT)	4,35e-02
Group x Pronunciation (TOTAL)	2,28e-10
FIN: Pronunciation x Position/Roundedness (LIST)	16,41e-02
FIN: Pronunciation x Position/Roundedness (TEXT)	7,02e-03
AUT: Pronunciation x Position/Roundedness (LIST)	1,38e-03
AUT: Pronunciation x Position/Roundedness (TEXT)	4,33e-02

Table 1: p-values from the χ^2 -test for /w/.

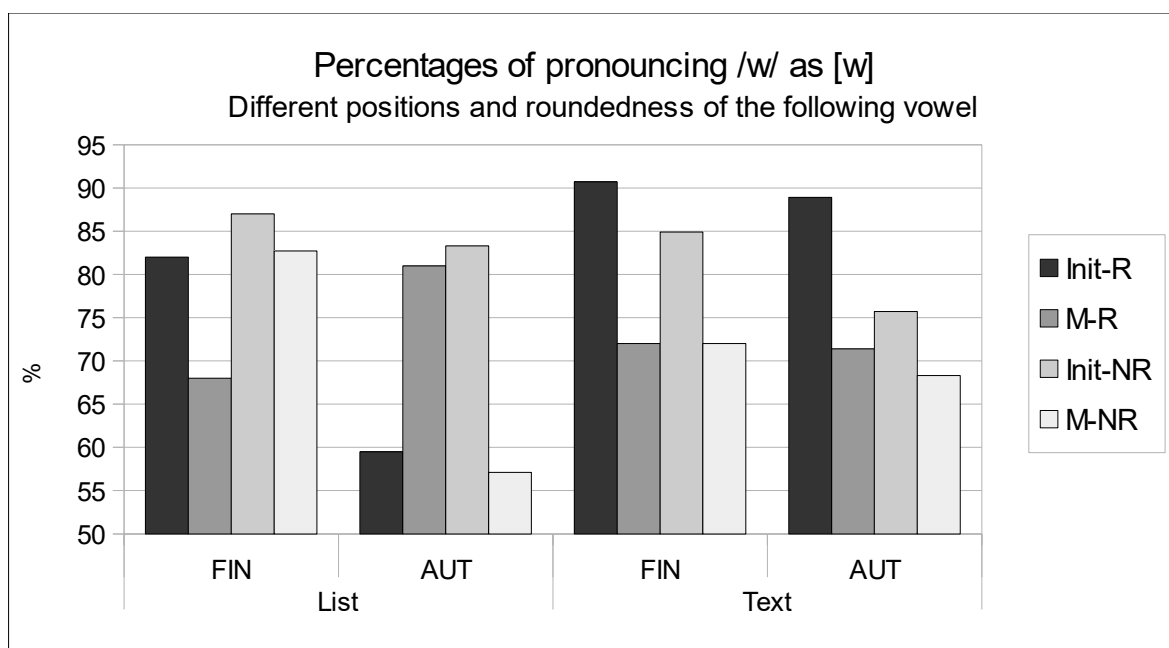


Figure 6: The percentages (y-axis) of pronouncing /w/ as [w] in different positions and varying roundedness of the following vowel in both pronunciation tasks (word list on the left and text passage on the right): word-initial position followed by a rounded vowel (Init-R), word-medial followed by a rounded vowel (M-R), word-initial followed by a non-rounded vowel (Init-NR), and word-medial followed by a non-rounded vowel (M-NR).

According to the analysis of the word list samples, the very first word, *worth*, was the “most difficult” one i.e. the lowest number of informants pronounced the word in a standard-like manner. 68,0% of the Finnish and only 38,1% of the Austrian informants pronounced *worth* with a [w]. Figure 7 illustrates the non-standard-like pronunciation of *worth* by an Austrian informant. Other “difficult” words were *wag* and *award* (Figure 8) for the Finnish informants and *forward* and *otherwise* for the Austrians. The words most frequently pronounced with a [w] by the Finns were *with* and *water*: in both instances [w] was produced by 24 out of 25 Finnish informants (96%). The spectrogram of *water* as pronounced by a Finnish informant is presented in Figure 9. The majority of the Austrians (18 of 21 informants i.e. 85,7%) pronounced the /w/ in *with* (spectrogram in Figure 10) and *whack* as a [w]. Only the word *waiter* was pronounced more often in a standard-like manner by Group A (90,5%). The text passage words most frequently pronounced with a [w] by Group 1 include *with* (string 12, Appendix C), *we’ll* (string 14), *-winning* (string 3) and *would* (string 8). The /w/ in all of these four words was pronounced as a [w] by 24 out of 25 Groups 1 informants. An example of the standard-like pronunciation of a /w/ in *with* by a Finnish informant is presented in Figure 11. All of the 21 Austrian informants produced a [w] in pronouncing *waters* (string 11; spectrogram in Figure 12), and a [w] could be detected in 20 out of 21 samples of *wake-boarding* (string 12). *Forward* (string

25), *will* (string 21) and *Taiwan* (string 4) were the three words from the text passage with the lowest percentages of standard-like pronunciation by the Finnish informants. At least one third of the Finnish informants pronounced these words with a [v]. Figure 13 illustrates this non-standard-like pronunciation of *Taiwan*. Producing a [w] in the beginning of *will* (both strings 14 and 21; Appendix C) and in the middle of *Taiwan* was also difficult for the Austrian informants. 12 or less of the total of 21 Group 2 informants pronounced these words in a standard-like manner, i.e. with a [w]. Figure 14 presents an example of the word *will* pronounced in non-standard-like manner by an Austrian informant.

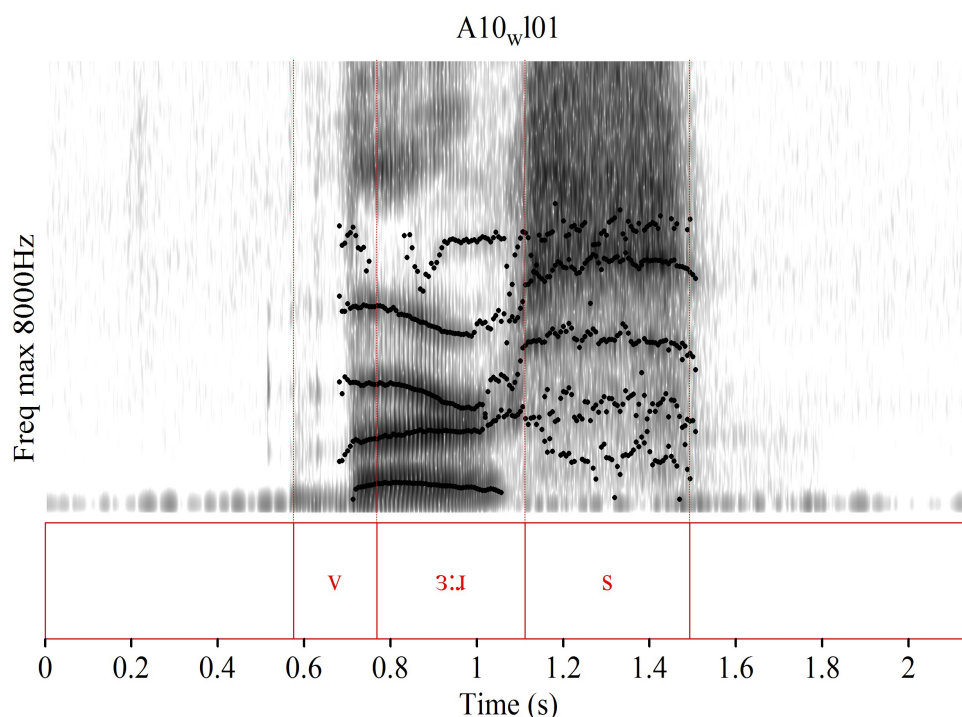


Figure 7: Spectrogram of word 1 of the word list as pronounced by informant A10. The word *worth* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

The analyses of the data from both pronunciation tasks revealed an interesting phenomenon related to the pronunciation of word-initial /v/: it was often pronounced as a [w] by informants from both groups (Figure 15). This hypercorrection was more common among the Finnish informants and occurred more frequently in the text passage samples as the graph in Figure 15 illustrates. A [w] was produced especially when the word beginning with a /v/ was preceded by a word with a word-initial /w/. Two such strings were intentionally included in the text passage: *would very* (string 16; Appendix C) and *We'll visit* (string 19). 40% of all Finnish informants produced a [w] when pronouncing the word *very* and almost every second Austrian informant (47,6%) pronounced the word *visit* with a

[w] (Figure 16). The spectrograms of these two strings are presented in Figures 17 and 18. Also the two words in the word list as well as the word *van* in string 28 (Appendix C) were pronounced with a [w] by some of the informants. Seven Finnish informants (28%) produced a [w] in the beginning of *van*. Both the standard-like and the [w]-pronunciation of the word *van* by two Finnish informants are illustrated in Figures 19 and 20. The standard-like pronunciation of the voiced labiodental fricative in *vein* by an Austrian informant is presented in Figure 21. The spectrogram of the same word in Figure 22 illustrates the non-standard-like pronunciation by another Austrian informant.

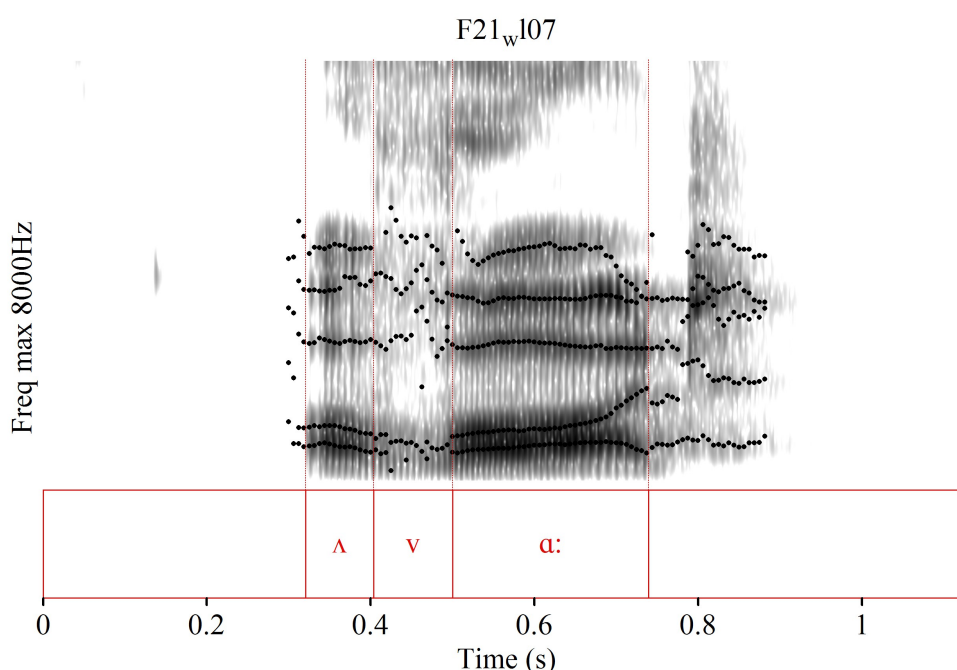


Figure 8: Spectrogram of word 7 of the word list as pronounced by informant F21. The first three phones of the word *award* are marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

On average, a typical Finnish informant produced the voiced labiovelar approximant in 8,3 out of ten word list occurrences. This arithmetic mean is lower than both the median and the mode i.e. the distribution is negatively skewed. According to both the median and the mode, i.e. the most frequently occurring value in the Group 1 word list data for /w/, the majority of Group 1 informants (13 Finns) pronounced all ten words in a standard-like manner (Appendix G). The average amount of produced standard-like sounds per Group 2 informant was slightly lower than the Finnish mean: seven out of ten ‘w’-words were pronounced with a [w]. Also the Austrian distribution is slightly skewed to the right, i.e. more than half of the Austrian informants produced a [w] more often than the mean value suggests. 38% of Group 2 informants pronounced at least 9 out of ten w-

occurrences with a [w]. This difference between the two groups in individual informant's performance decreased in the text passage task: an average Finn produced a [w] in 13,3 and a typical Austrian in 12,2 out of 16 instances of /w/. This time the distribution of Group 1 performance is more markedly negatively skewed than that of Group 2 for the difference between mode and mean values is 2,7 (Appendix G). Almost 50% of Group 1 informants (12 out of 25) pronounced at least 15 of the 16 instances of /w/ in a standard-like manner. In comparison, five out of 21 (24%) Group 2 informants produced a [w] in at least 15 out of 16 occurrences of /w/.

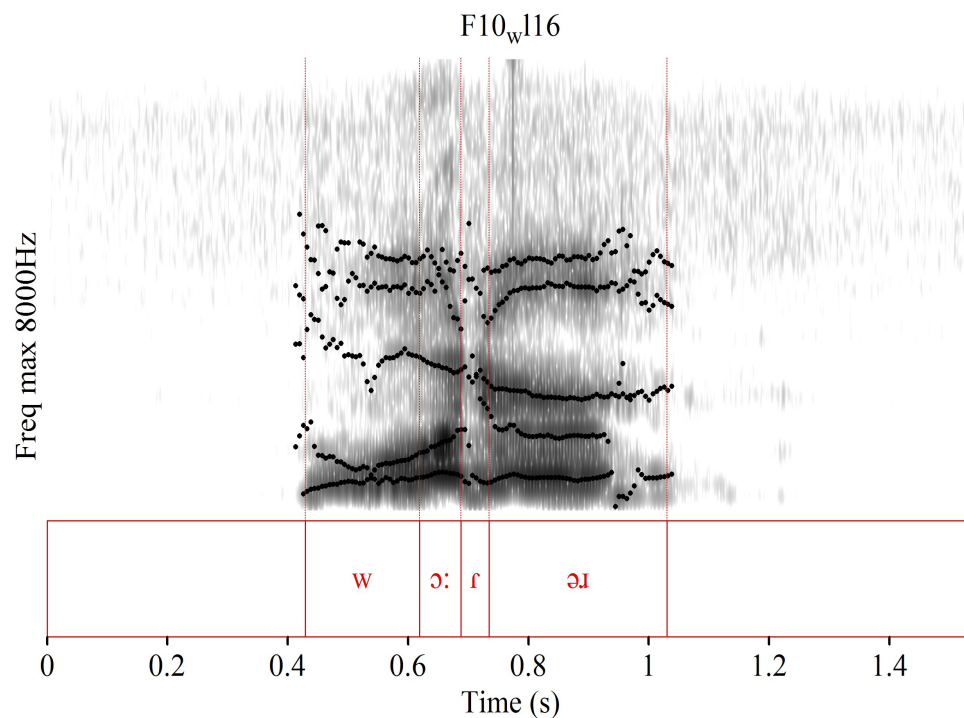


Figure 9: Spectrogram of word 16 of the word list as pronounced by informant F10. The word *water* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

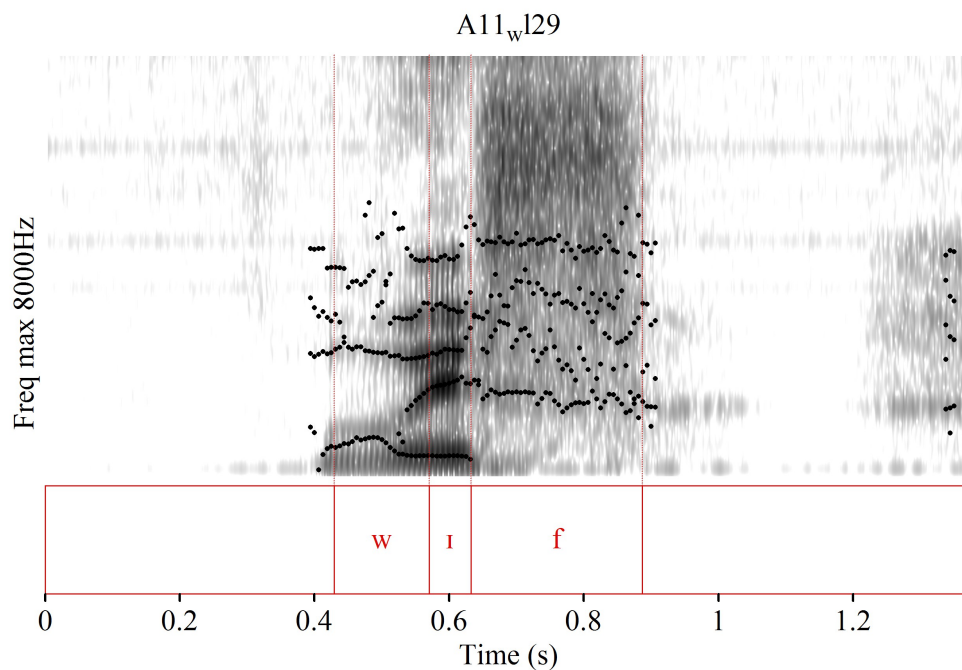


Figure 10: Spectrogram of word 29 of the word list as pronounced by informant A11. The word *with* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

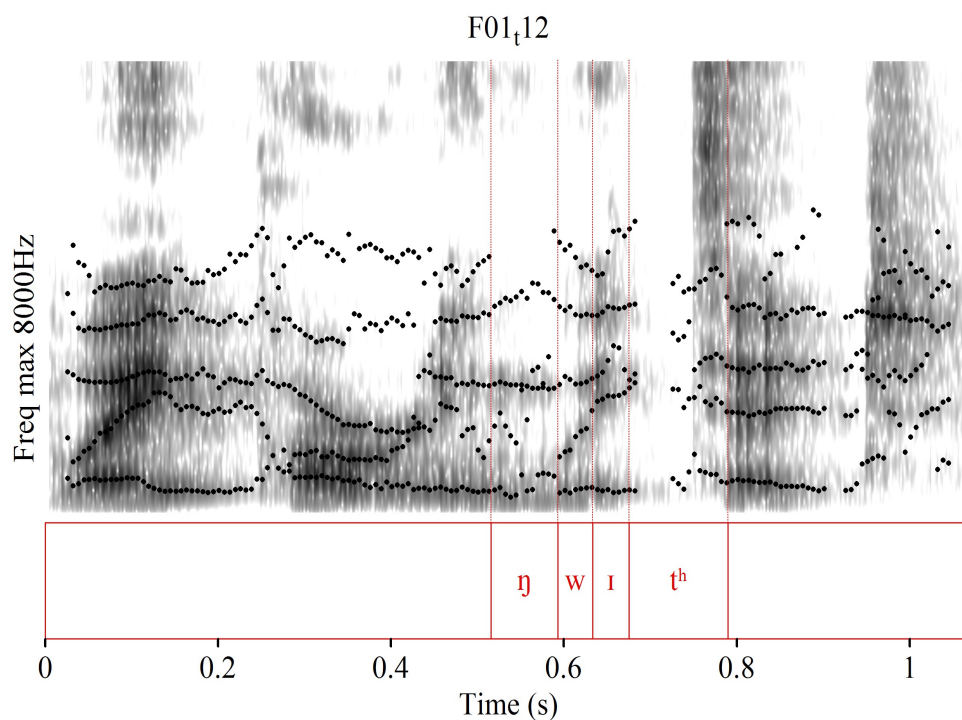


Figure 11: Spectrogram of string 12 of the text passage as pronounced by informant F01. The last sound of the word *wake-boarding* and the word *with* are marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

A20_t11

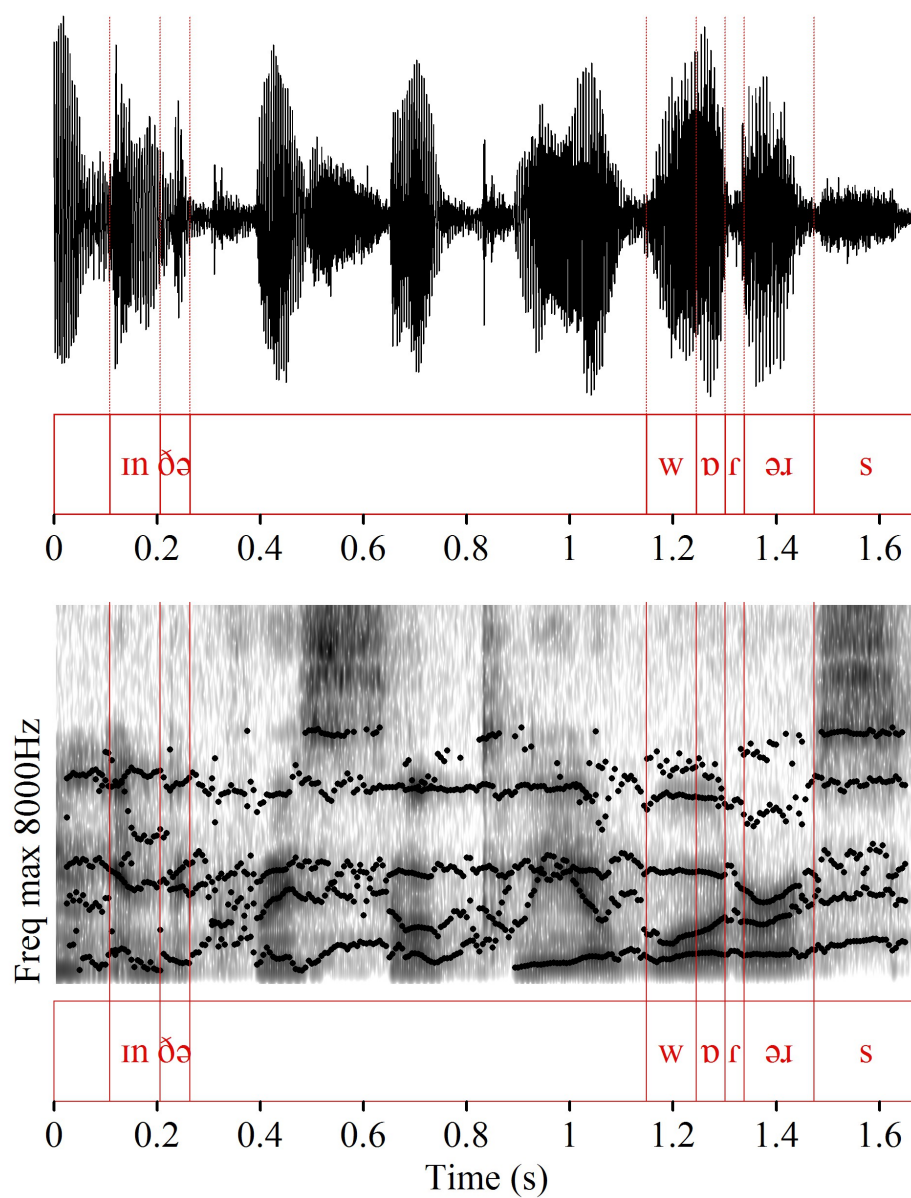


Figure 12: The waveform (above) and the spectrogram (below) of string 11 of the text passage as pronounced by informant A20. The words *in the* and *waters* are marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

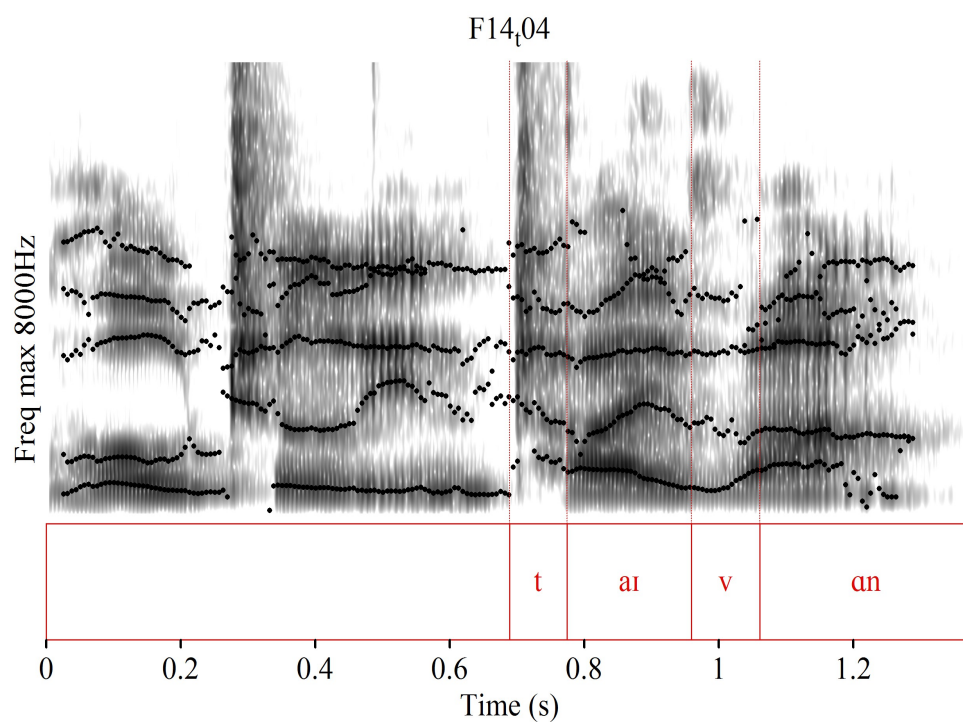


Figure 13: Spectrogram of string 4 of the text passage as pronounced by informant F14. The word *Taiwan* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

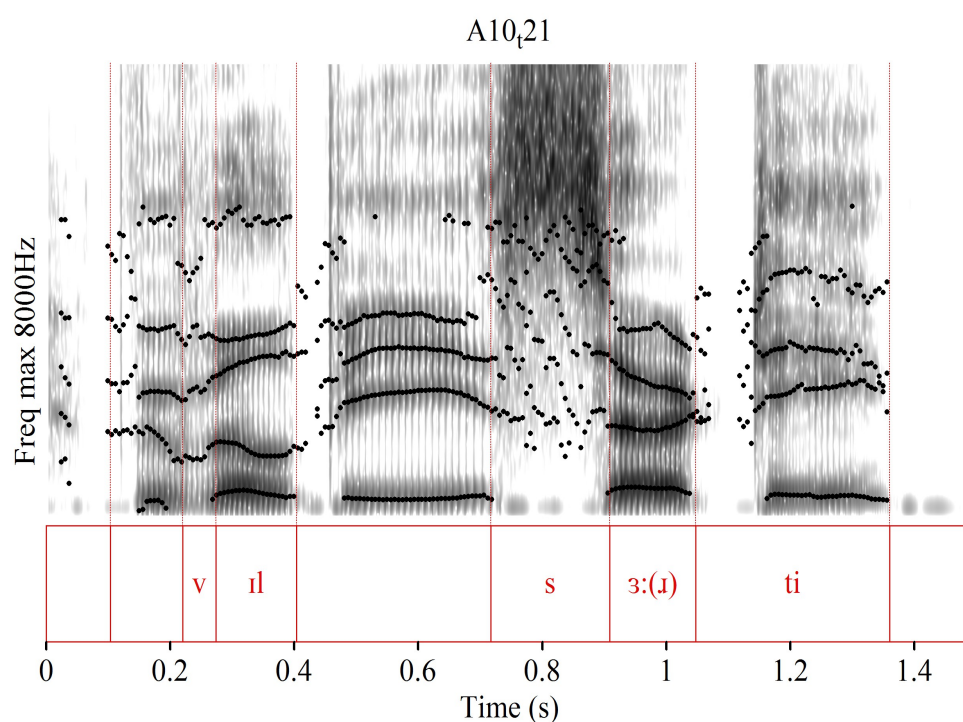


Figure 14: Spectrogram of string 21 of the text passage as pronounced by informant A10. The words *will* and *thirty* are marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

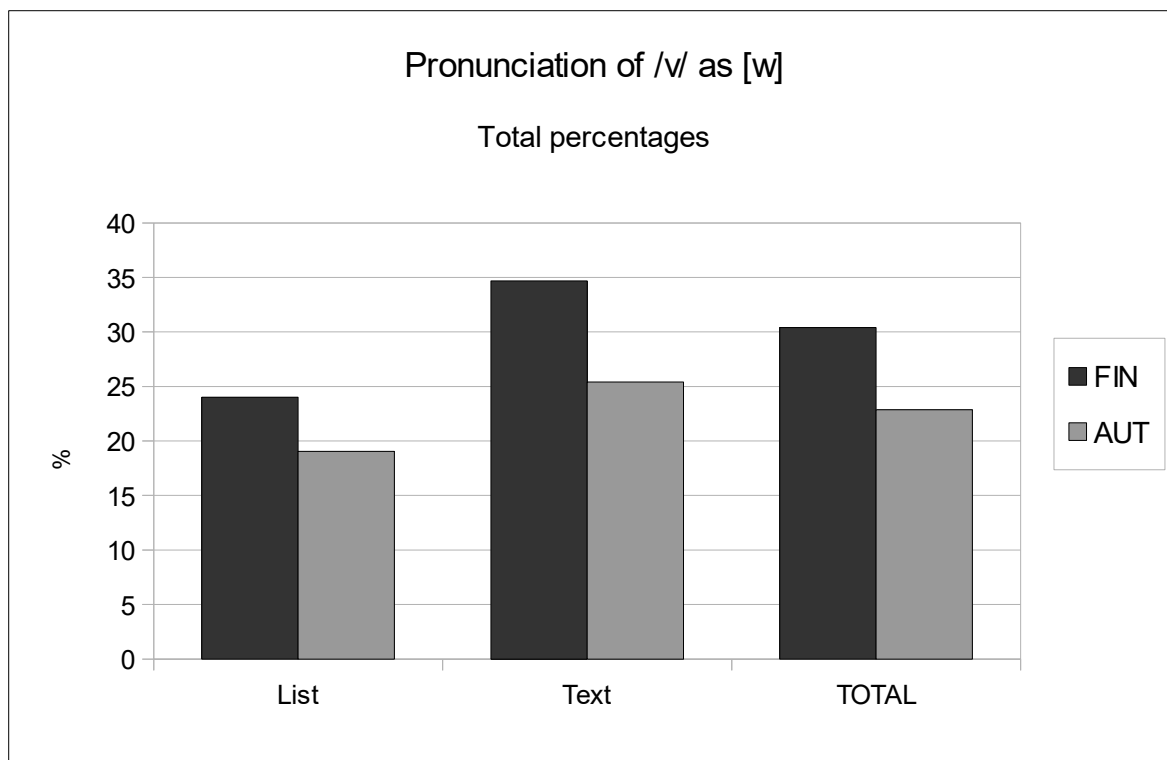


Figure 15: Pronunciation of /v/ as [w]: word list (left), text passage (middle) and total percentage (right).

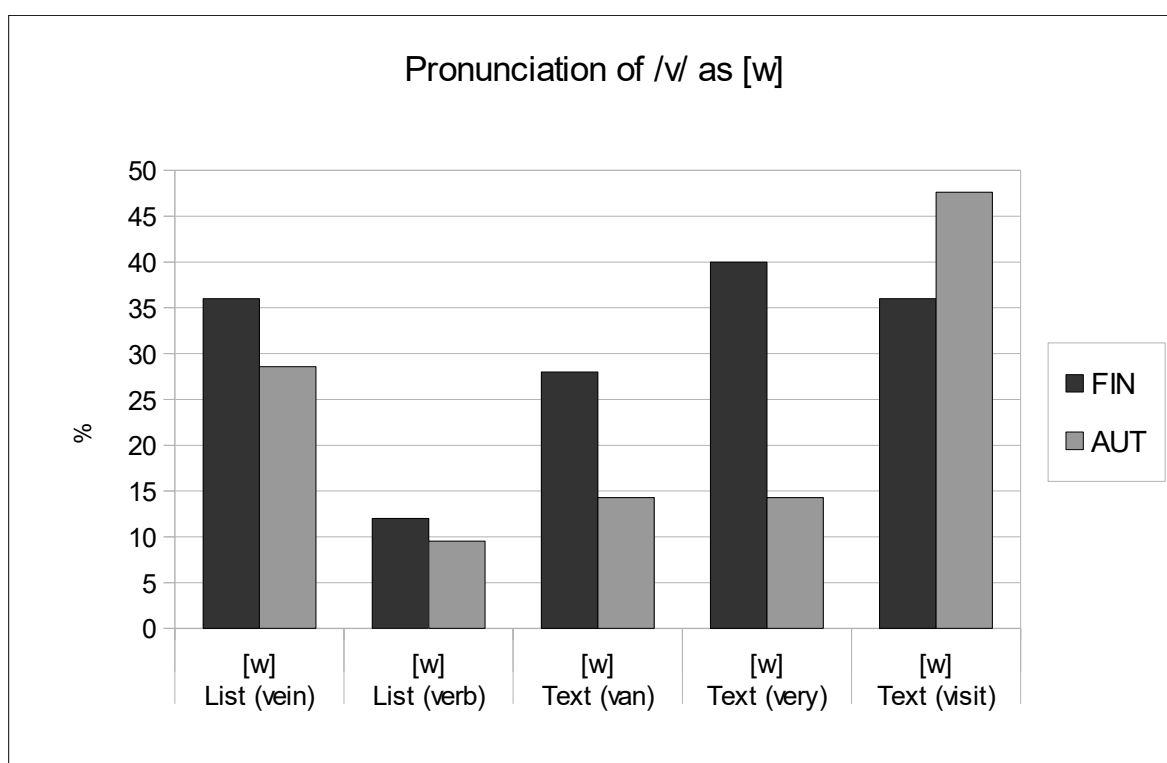


Figure 16: /v/ in word-initial position: the percentages of producing a [w] in both tasks.

F07_t16

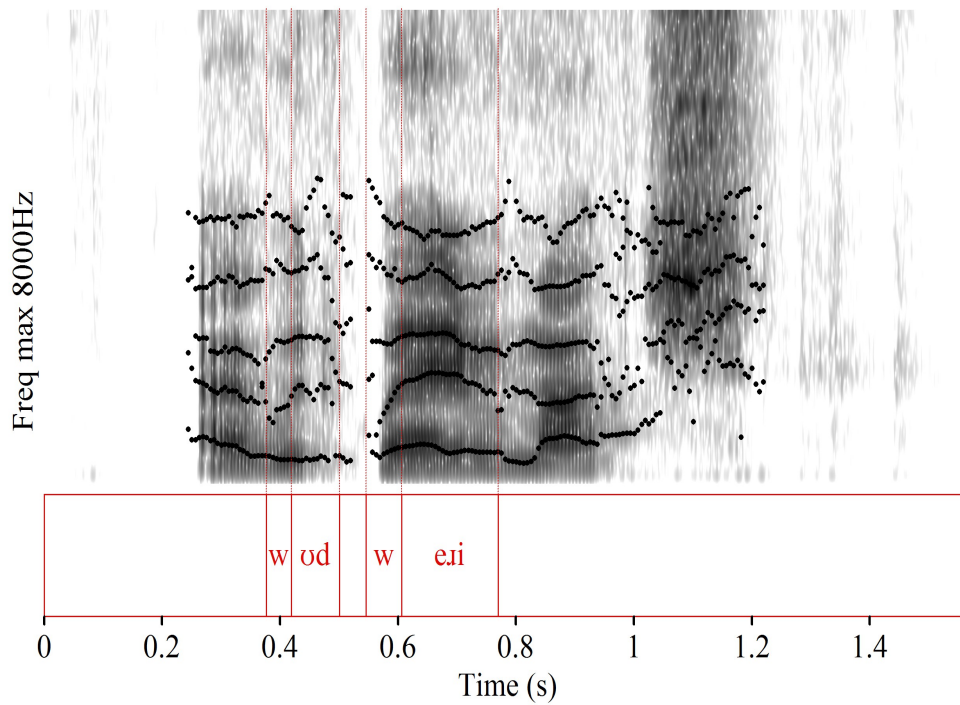


Figure 17: Spectrogram of string 16 of the text passage as pronounced by informant F10. The words *would* and *very* are marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

A01_t19

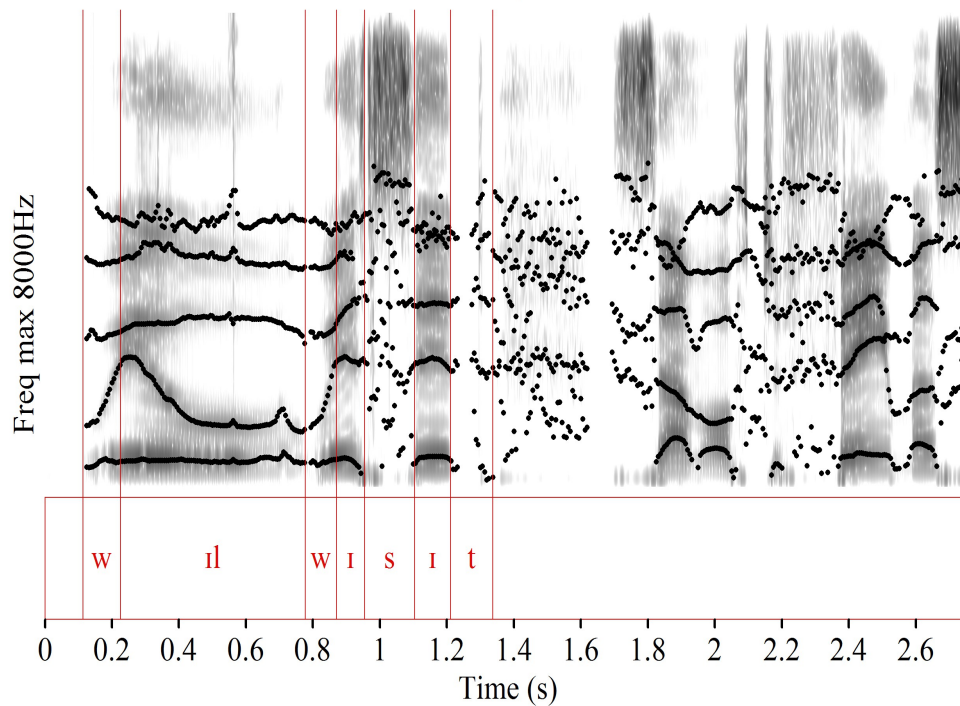


Figure 18: Spectrogram of string 19 of the text passage as pronounced by informant A01. The words *we'll* and *visit* are marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

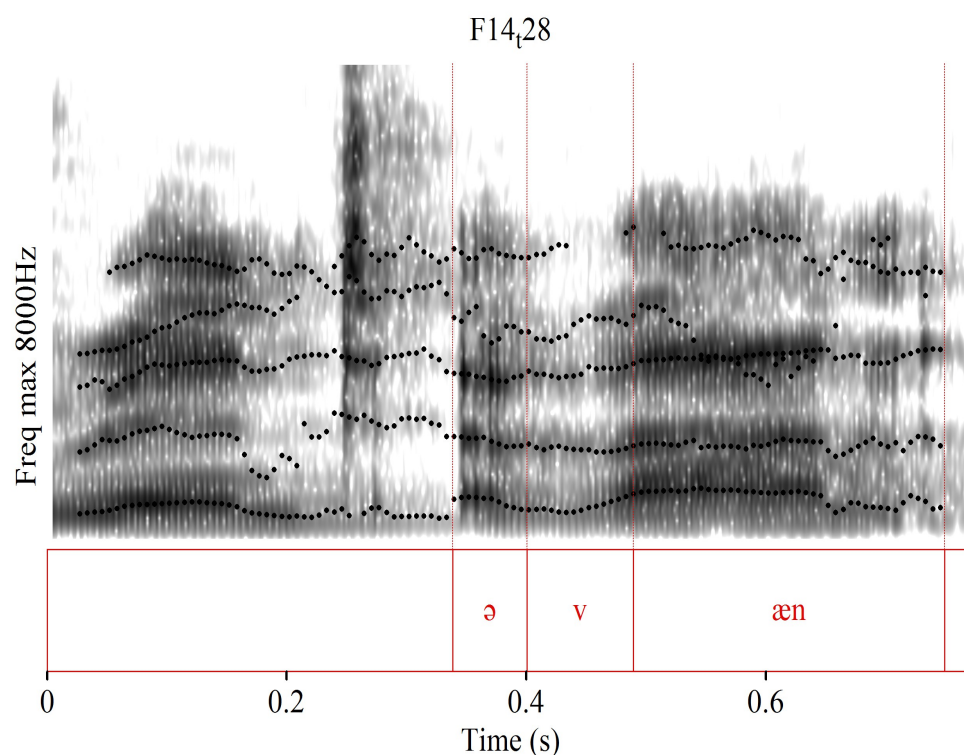


Figure 19: Spectrogram of string 28 of the text passage as pronounced by informant F14. The noun phrase *a van* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

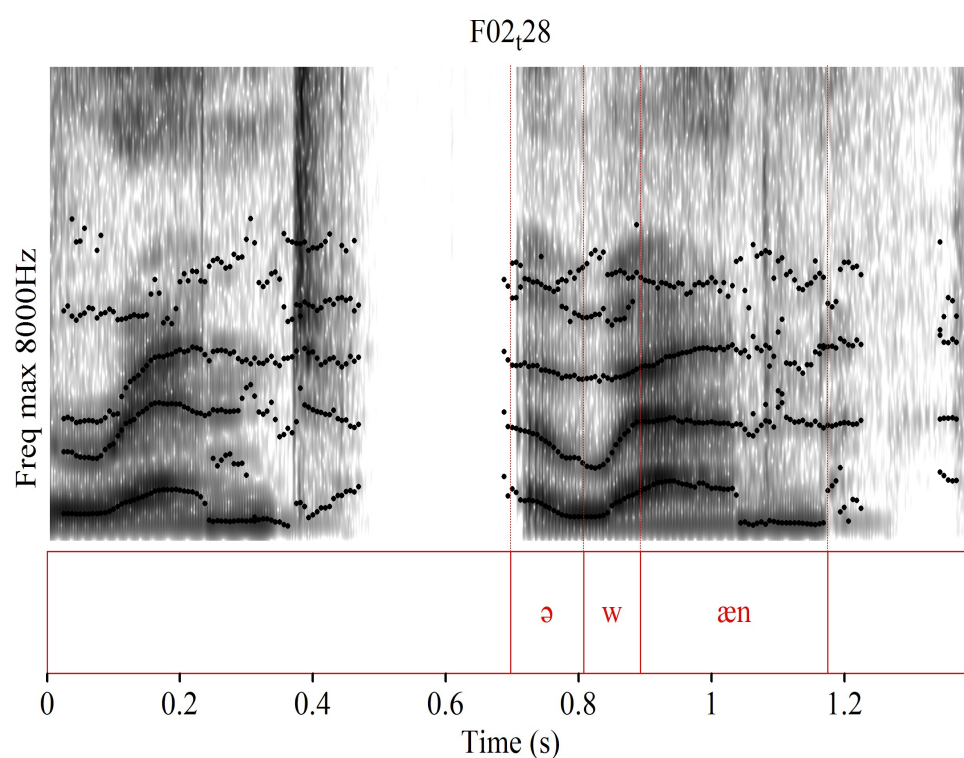


Figure 20: Spectrogram of string 28 of the text passage as pronounced by informant F02. The noun phrase *a van* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

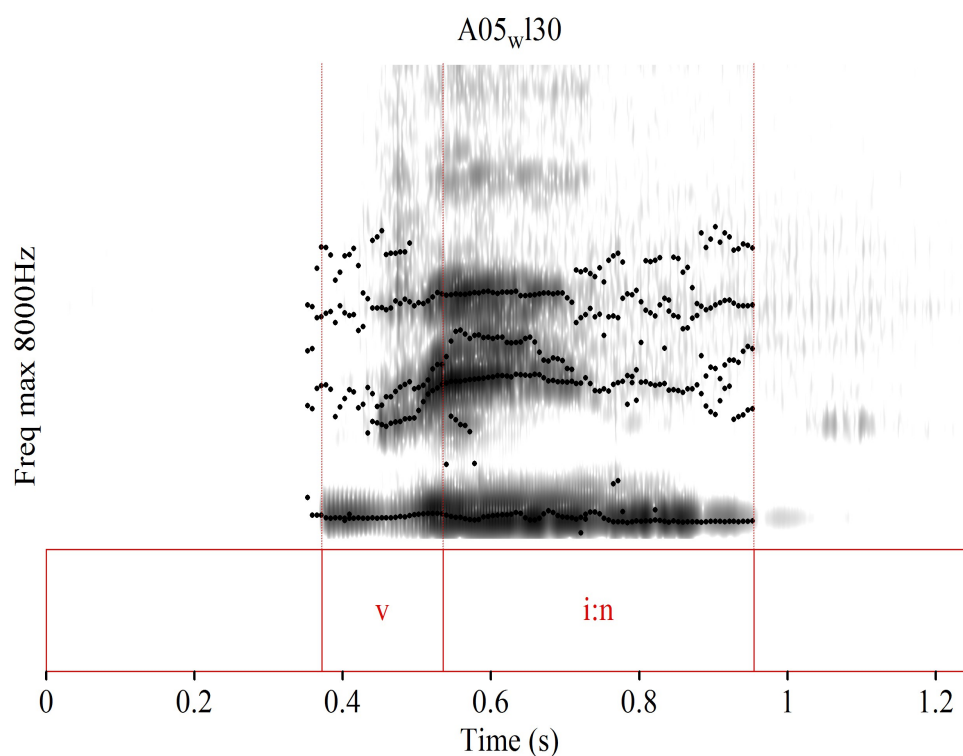


Figure 21: Spectrogram of word 30 of the word list as pronounced by informant A05. The word *vein* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

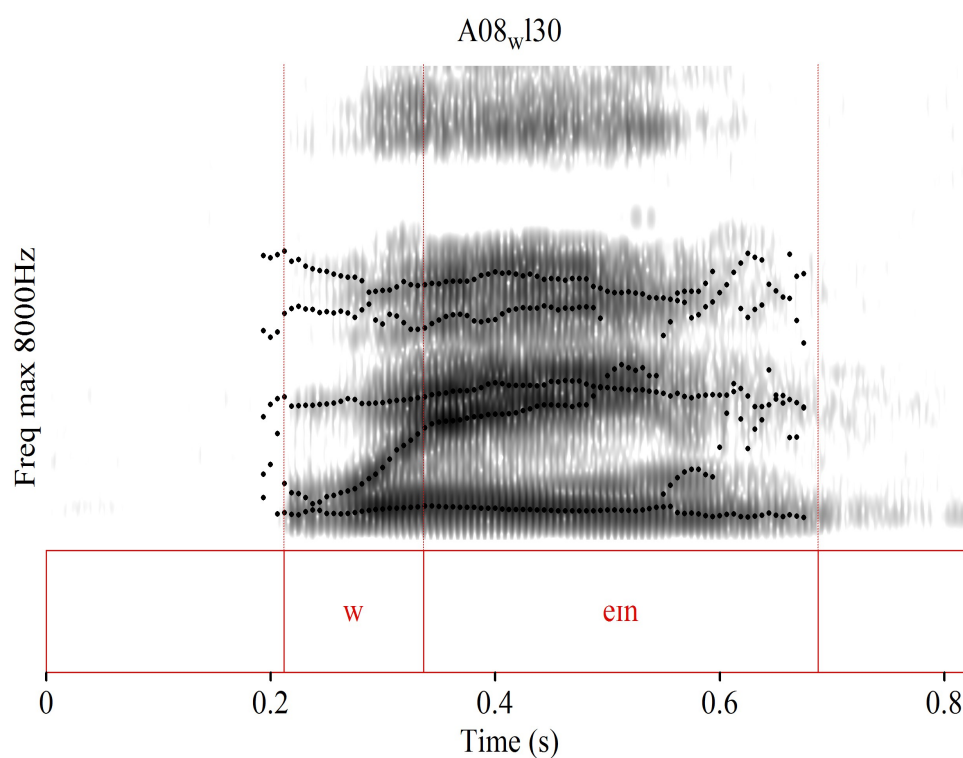


Figure 22: Spectrogram of word 30 of the word list as pronounced by informant A08. The word *vein* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

4.2 Voiceless and voiced (inter-)dental fricatives /θ/ and /ð/

The identification of the (inter-)dental fricatives from the pronunciation samples was not as straightforward as identifying the /w/ with its distinct formant structure. As mentioned earlier, the main difference between /θ/ and /ð/ is the voicing and the lenis character of the latter. The /ð/ sound is generally shorter and weaker than /θ/. On the spectrogram of a strongly voiced /ð/ sound, a voicing bar is visible and the waveform is periodic. The alternative forms of pronunciation were distinguished from the standard-like pronunciation by applying the following criteria. The manner of pronunciation of both the stereotypical Finnish dentalveolar [t̪] and the aspirated [tʰ] differs from that of the (inter-)dental fricatives. Both of these sounds are plosives and can be recognised from the visualisations by their typical structure, which includes closing, compression and release phases (in the case of a [tʰ] also post-release phase is included). The most prominent features of these voiceless sounds are the total stricture, i.e. silence, during the compression phase and the following release burst. These typical features are visible in Figures 23 (second and third from the right) and 24 (second from the right).

Differentiating between the other alternative pronunciations and the [θ] and the [ð] proved to be more difficult. All of the other common alternatives were fricatives and differed from the standard-like sounds only through place of pronunciation and loudness. The stereotypical Austrian replacement for the (inter-)dental fricatives, [s], is an alveolar fricative and being a sibilant sound, it is typically louder than the researched sounds. This loudness is visible both in the spectrograms and in the waveforms: the spectrogram of [s] is darker and the amplitude of the aperiodic waves is higher (the right-most visualisations in Figures 23 and 24). The (inter-)dental fricatives are relatively weak sounds and very similar to the labiodental fricatives /f/ and /v/. As mentioned in section 2.2.2, the main criteria for differentiating the (inter-)dental fricatives from the labiodental ones are the fricative noise and related formant transitions: in Figure 23, for example, both the second and the third formant rise rapidly after [f], but after [θ] only the second formant rises. A similar rise is visible in Figure 24 after [f] but not after [θ]. The differences in both fricative noise and formant transitions are often very small. In search for more reliable and objective ways of differentiating between different fricatives, Jongman, Wayland & Wong compared various acoustic cues and tested how accurately they predicted the place of articulation in English fricatives (2000). Even though the results of their testing suggest that a combination of several different variables can reliably be used in differentiating all

four places of articulation, the correct group membership (i.e. place of articulation) for both (inter-)dental and labiodental fricatives could be predicted only in two thirds of all instances in their study (Jongman, Wayland and Wong 2000:1261). The percentages of correct prediction of non-sibilant fricatives lowered noticeably when only three of the most effective acoustic properties were used in distinguishing between the four places of articulation: on average the correct place could be predicted only in 50% of the occurrences (ibid.). Since the main focus of this study is to analyse the pronunciation of four different groups of sounds and not only that of the (inter-)dental fricatives, these complicated methods of differentiating between (inter-)dental and labiodental fricatives were not applied. In cases, where no distinctions could be made on the basis of the visual realisations of the fricative sounds, the analysis was done subjectively relying on the hearing of the author.

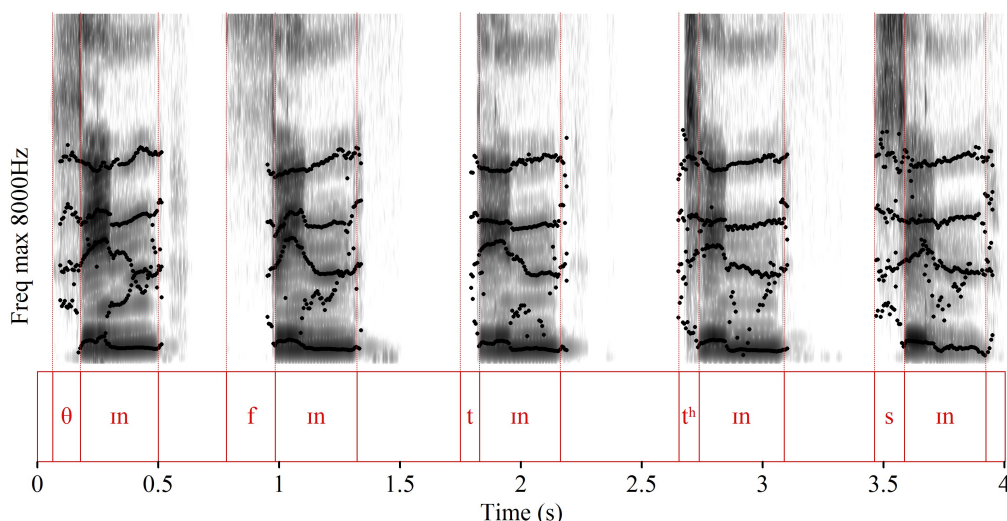


Figure 23: Spectrograms of the words *thin* (left-most), *fin*, *tin* (with a dentalveolar plosive [t]), *tin* (with aspiration) and *sin* as pronounced by the author. The words are marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

The Finnish informants pronounced the letter combination ‘th’ in a standard-like manner more often than the Austrians in both pronunciation tasks, although the difference is relatively small in the word list task (Figure 25). The Finns produced the standard-like sound in 62,7% of all word list instances and in 60,7% of all text passage occurrences. The same percentages for the Austrians are 60,3% and 45,1%. Due to the small difference, the correlation between group and the pronunciation of the th-sounds of the word list task is not statistically significant at the chosen 5% significance level (Table 2). In contrast to the word list task, the differences in pronunciation in the text passage samples between the two

informant groups are clearly significant, for the p-value of 8.55e-06 from the χ^2 -test is markedly smaller than the significance level 0,05 (Table 2).

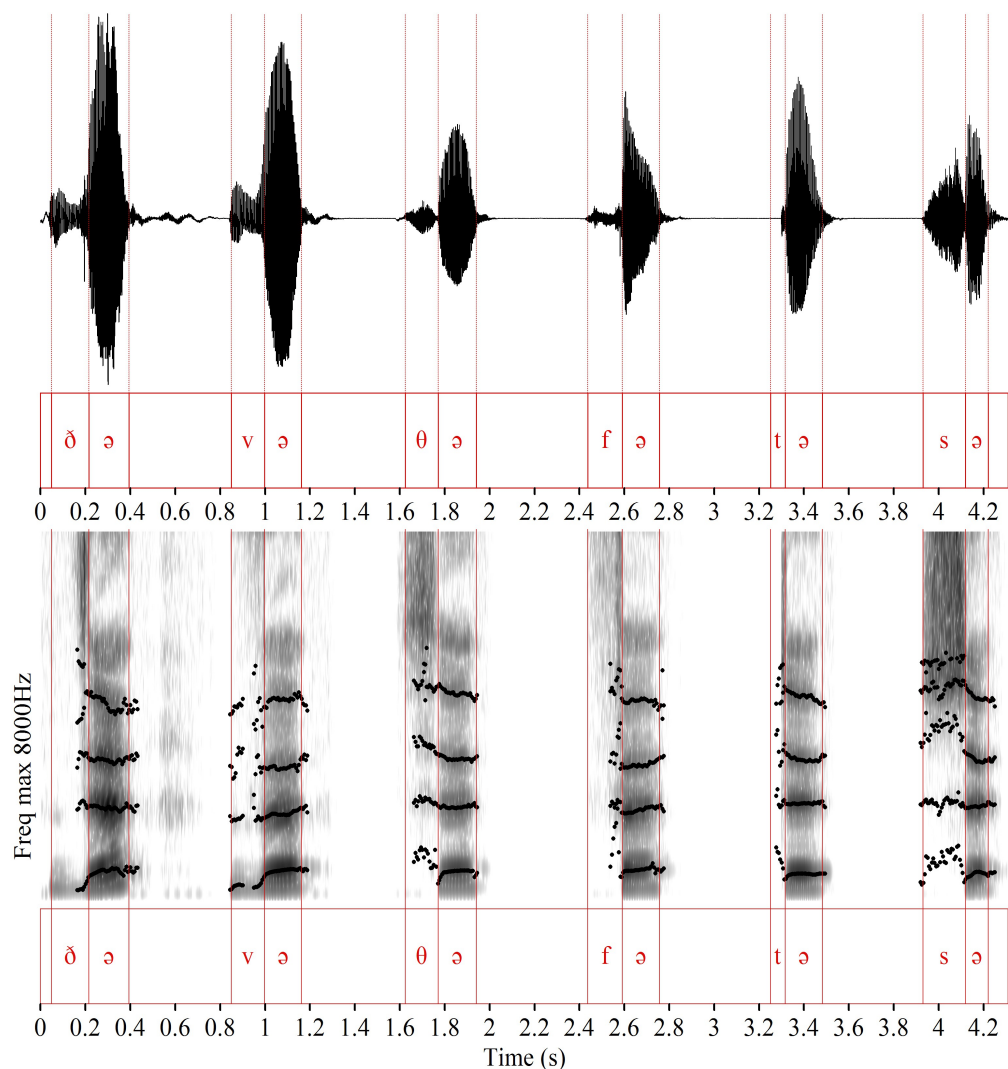


Figure 24: The waveforms (above) and the spectrograms of the word *the* as pronounced by the author in five alternative ways: [ðə] (left), [və], [θə], [tə], [fə] and [sə] (right). The sounds marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

The voiceless (inter-)dental fricative was produced more frequently in a standard-like manner than the /ð/: both groups produced a [θ] in around two thirds of all instances in the word list task (Group 1: 66%; Group 2: 64%). The difference between the two informant groups is greater in the text passage task: less instances of [θ] could be found in the Austrian text passage samples (53%), whereas the Finns' percentage of standard-like pronunciation increased ever so slightly as the task formality decreased. [θ] was produced by the Finnish informants in 66,5% of all instances in the text passage. The pie charts on

the left side in Figures 45 (Group 1) and 46 (Group 2) present the most common substitutions of /θ/ in both pronunciation tasks. These sounds include the dentalalveolar ‘t’ ([t̪]) and the aspirated voiceless alveolar plosive ([tʰ]) for Group 1 and both the voiceless labiodental and alveolar fricatives [f] and [s] for Group 2. Similar to the labiovelar approximant, the pronunciation of the (inter-)dental fricatives in different positions was analysed and the results can be seen in Figure 26. In the word list task both groups produced a [θ] most frequently in word-medial position (Group 1: 68%; Group 2: 73,8%). This was also the only position in which Group 2 samples included more occurrences of standard-like pronunciation than Group 1. The differences in the percentages between the groups were generally (with one exception) relatively small (2-13,7%). In the text samples, words with word-final /θ/ were most commonly pronounced with a [θ] by both groups: Finns produced a [θ] in almost 70% of the occurrences and 55,3% of the instances were pronounced as [θ] by the Austrians. Even though these percentages indicate a connection between standard-like pronunciation and the position of /θ/, they are not statistically significant at the chosen significance level i.e. there is not enough data to prove that these preferences of position could result from more than just random variation (Table 2).

The voiced (inter-)dental fricative was slightly less frequently pronounced in a standard-like manner than its voiceless counterpart: Group 1 informants produced a [ð] in 58% and Group 2 in 53% of all the instances in the word list data. Similar to /θ/, the Group 1 percentage increased slightly in the less formal task, whereas the Group 2 percentage decreased to 40%. The Finnish informants substituted the voiced (inter-)dental fricative in often with [θ], [t̪] and [tʰ] (pie charts on the right in Figure 45). In the Austrian samples, the most common alternatives for /ð/ were [θ], [t̪] and [r] or a flap-like [d] (right pie charts in Figure 46). Of all the positions in both tasks, standard-like pronunciation could most frequently be observed in the word list samples of the voiced (inter-)dental fricative in word-initial position. 82% of all Finnish and 76,2% of all Austrian occurrences were pronounced as a [ð]. In contrast, the word-final position of /ð/ proved to be the most difficult one for both groups. Since the word *with* can be pronounced with both the voiced and the voiceless (inter-)dental fricatives, the only word with /ð/ in this position is word 9 (*smooth*) of the word list. It was pronounced with a [ð] only by 3 Finnish and 2 Austrian informants. The majority of the informants (14/25 Finns and 7/21 Austrians) pronounced the word as [smu:θ]. In the text passage samples of Group 1, word-initial /ð/ was more frequently pronounced in a standard-like manner than word-medial /ð/. The opposite was true for Group 2: a [ð] was produced in 46% of the word-medial and in 38,1% of the word-

initial occurrences. According to the χ^2 -test, the word list samples of both groups indicate a significant correlation between the position and the pronunciation of /ð/ at the 5% significance level (Table 2). Similar correlation could not be detected in the text passage samples. More data would be required to ascertain whether the position affects the pronunciation of /ð/ also in less formal tasks.

The two most “difficult” word list words with a [θ] were *thirty* (word 8) for Group 1 and *worth* (word 1) for Group 2. 17 out of 25 Group 1 informants pronounced *thirty* in a standard-like manner (Figure 27). The initial sound was also pronounced as an aspirated voiceless alveolar stop (Figure 28) and as a Finnish-style voiceless dentalveolar stop (Figure 29). Nine Group 2 informants produced a [θ] when pronouncing the word *worth* (Figure 30). The alternative pronunciation of this word-final /θ/ included an [f] (Figure 31) and an [s] (Figure 7). As mentioned already above, the word *smooth* (word 9; Figure 32) was the word list word least frequently pronounced in a standard-like manner i.e. with a [ð]: the final /ð/ was pronounced as a [θ] by the majority of the informants as in Figure 33. Alternatively, some Finnish informants also produced a [t^h] (Figure 34) and a [t̪] (Figure 35) when pronouncing the word. The Austrian substitutions for this word-final sound included an [f] (Figure 36) as well as an [s] (Figure 37).

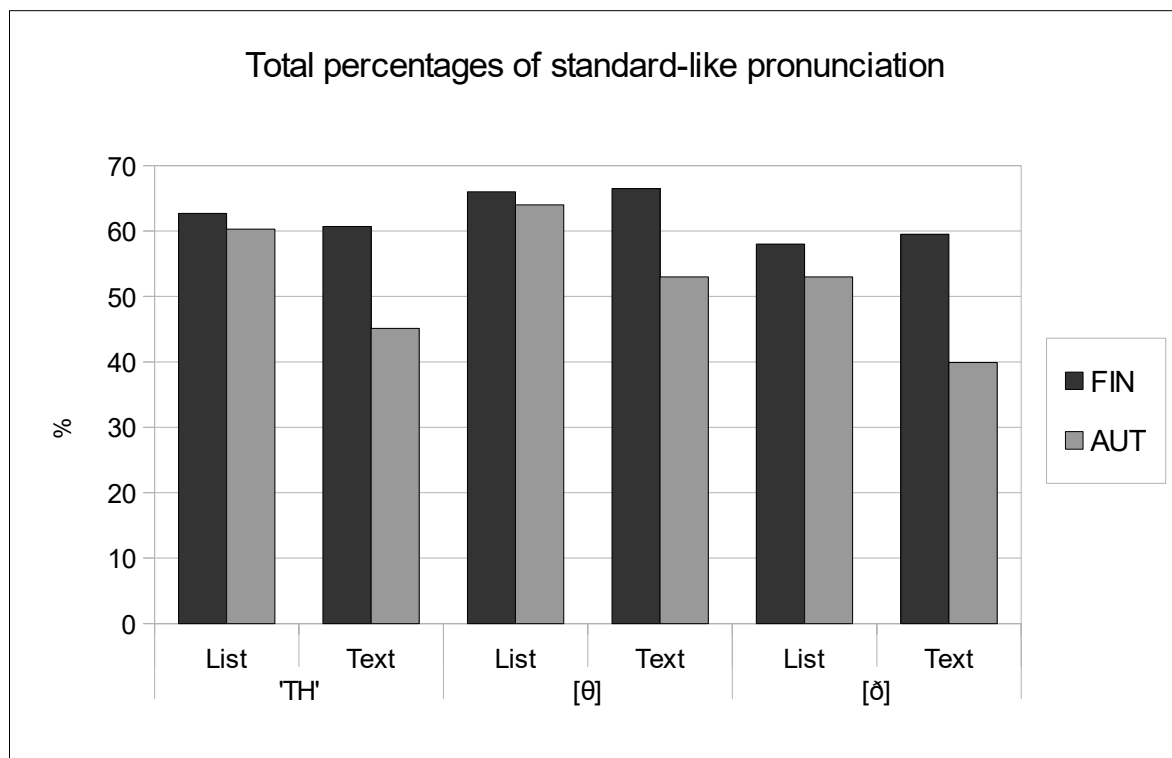


Figure 25: Total percentages of standard-like pronunciation of all ‘TH’-sounds (left), [θ] (middle) and [ð] (right).

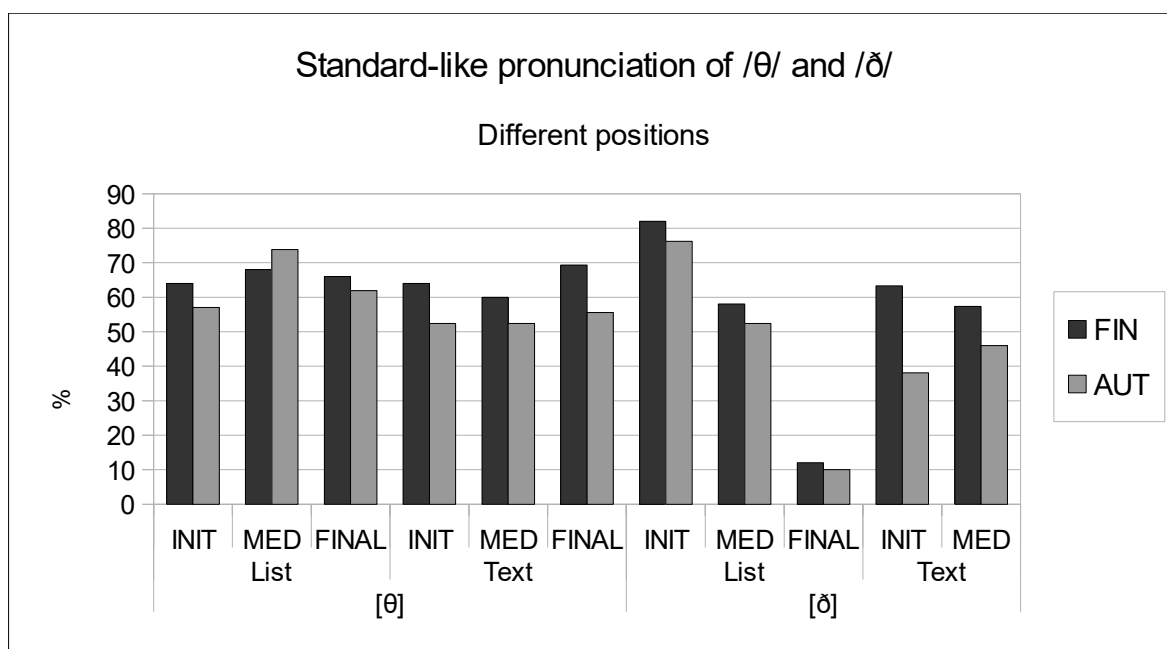


Figure 26: Percentages of standard-like pronunciation of [θ] (left) and [ð] (right) in word-initial (INIT), -medial (MED) and -final positions in both word list and text passage samples.

The text passage samples contained less instances of standard-like pronunciation than the more formal word list samples. The most “difficult” ‘th’-words for Group 1 were *birthday* and *soothing*: 13 out of 25 informants pronounced the first one with [θ] (Figure 38) and 9 produced a [ð] (Figure 40) when pronouncing the word *soothing*. Alternative pronunciations included elision of the fricative sound as in Figure 39 and the Finnish-style [t̪] (Figure 41). The text passage words *thirty* (string 21) and *the* (string 9) were pronounced least frequently in a standard-like manner by Group 2 informants. Figures 42 and 12 illustrate the pronunciation of these words with standard-like (inter-)dental fricatives. In addition to the standard-like pronunciation, Group 2 informants pronounced these word-initial sounds also with an [f] (Figure 43), an [s] (Figure 14) and with a [d] (Figure 44).

According to the Oxford Advanced Learner’s Dictionary (Turnbull 2011:1770), the word *with* (word 29 in the word list) can be pronounced with both [θ] and [ð]. This is why the analysis of this word was done separately i.e. the results were only included in the total percentage of ‘th’ pronunciation (Figure 25: left-most bars). 16 out of 25 (64%) Group 1 and 15 out of 21 (71%) Group 2 informants pronounced the word in the word list task as [wɪθ] or as [wɪð]. Most informants pronounced *with* with a voiceless (inter-)dental fricative, but also with an [f] like in Figure 10. The word *with* was also included in the text

Variables (Controlled x measured)	p-values from the χ^2-test ($\alpha=5e-02$)
Group x Pronunciation of TH (LIST)	63,31e-02
Group x Pronunciation of TH (TEXT)	8,55e-06
Group x Pronunciation of TH (TOTAL)	1,75e-04
Group x Pronunciation of /θ/ (LIST)	8,23e-02
Group x Pronunciation of /θ/ (TEXT)	4,93e-02
Group x Pronunciation of /θ/ (TOTAL)	8,89e-03
Group x Pronunciation of /ð/ (LIST)	52,36e-02
Group x Pronunciation of /ð/ (TEXT)	4,53e-05
Group x Pronunciation of /ð/ (TOTAL)	1,89e-04
FIN: Pronunciation x Position of /θ/ (LIST)	91,47e-02
FIN: Pronunciation x Position of /θ/ (TEXT)	55,12e-02
FIN: Pronunciation x Position of /ð/ (LIST)	5,00e-08
FIN: Pronunciation x Position of /ð/ (TEXT)	46,78e-02
AUT: Pronunciation x Position of /θ/ (LIST)	25,97e-02
AUT: Pronunciation x Position of /θ/ (TEXT)	92,96e-02
AUT: Pronunciation x Position of /ð/ (LIST)	3,68e-06
AUT: Pronunciation x Position of /ð/ (TEXT)	37,36e-02

Table 2: p-values from the χ^2 -test for 'th', /θ/ and /ð/.

passage (string 12) where it is followed by the word *Ted*. Most informants, both Finnish and Austrian, omitted the fricative sound completely like in Figure 11 where the transcribed [t^h] is the word-initial sound of the word *Ted*.

Between the two groups, there is very little difference in the performance of an individual informant in the word list task. On average, a Finnish informant pronounced the 'th' in a standard-like manner in 7,5 out of a total of 12 word list instances and the same mean value for a typical Austrian informant is 7,2. Similarly, the an average informant from both groups produced a [θ] in four out of six instances. The distributions of both the 'th' and /θ/ are slightly negatively skewed for both groups (Appendix G). Three out of five word list instances of voiced (inter-)dental fricative were pronounced with a [ð] by an average informant from both groups. This mean value describes the overall performance well, since the distributions of the groups are not skewed. As the analysis of the text passage samples already indicated, the difference in the performance of individual informants between Group 1 and Group 2 was greater in the less formal pronunciation task. Of the total of 17 'th'-words in the text, a Finnish informant pronounced on average

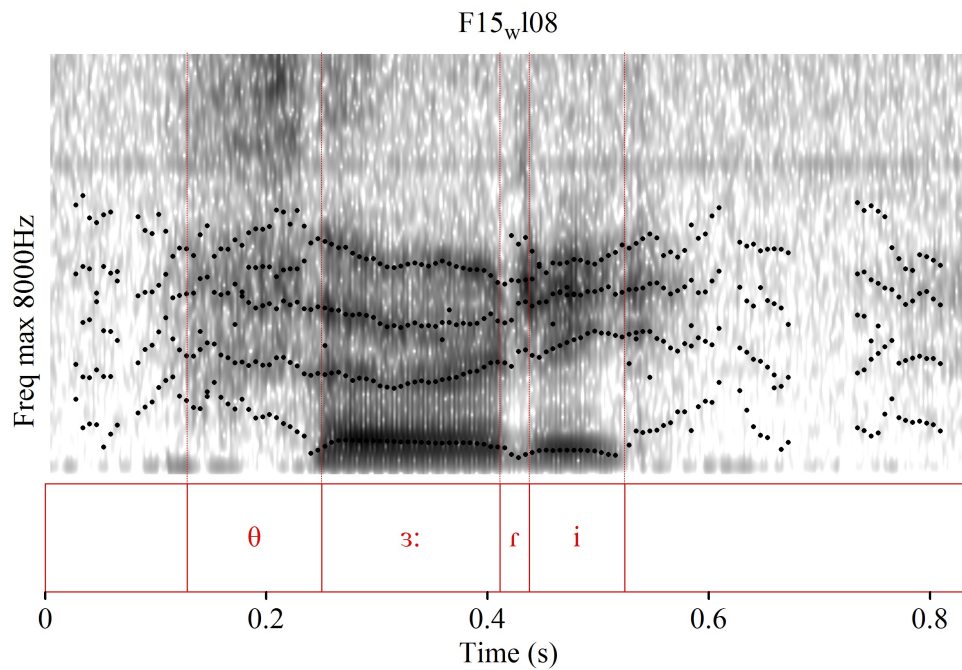


Figure 27: Spectrogram of word 8 of the word list as pronounced by informant F15. The word *thirty* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

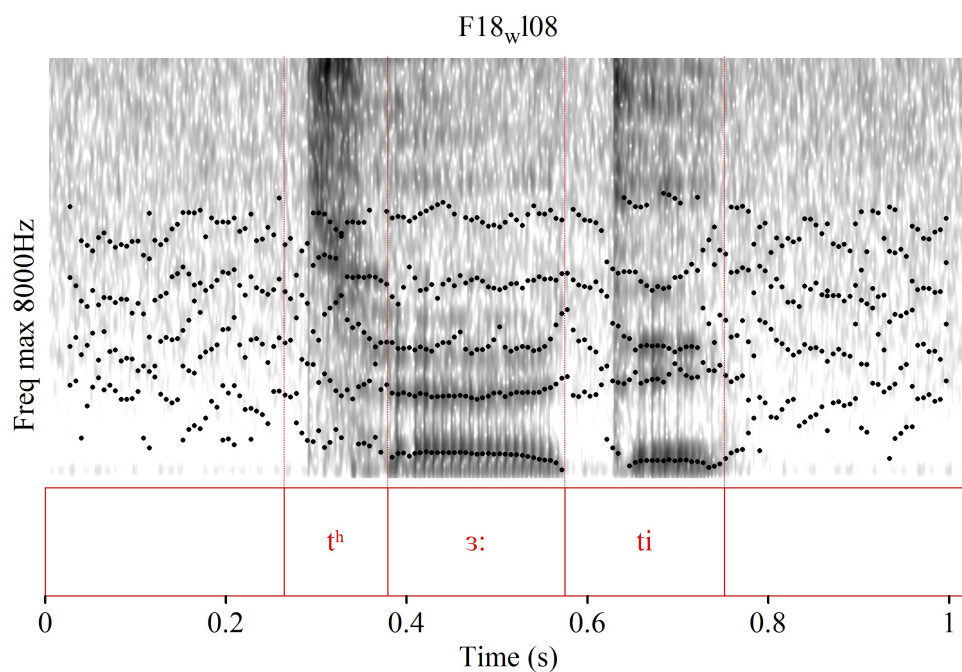


Figure 28: Spectrogram of word 8 of the word list as pronounced by informant F18. The word *thirty* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

10,4 instances with an (inter-)dental fricative. A typical Austrian informant pronounced a 7,3 out of 17 occurrences of ‘th’ in a standard-like manner. The differences within the two groups are also noteworthy: the range for Group 1, i.e. the difference between the highest

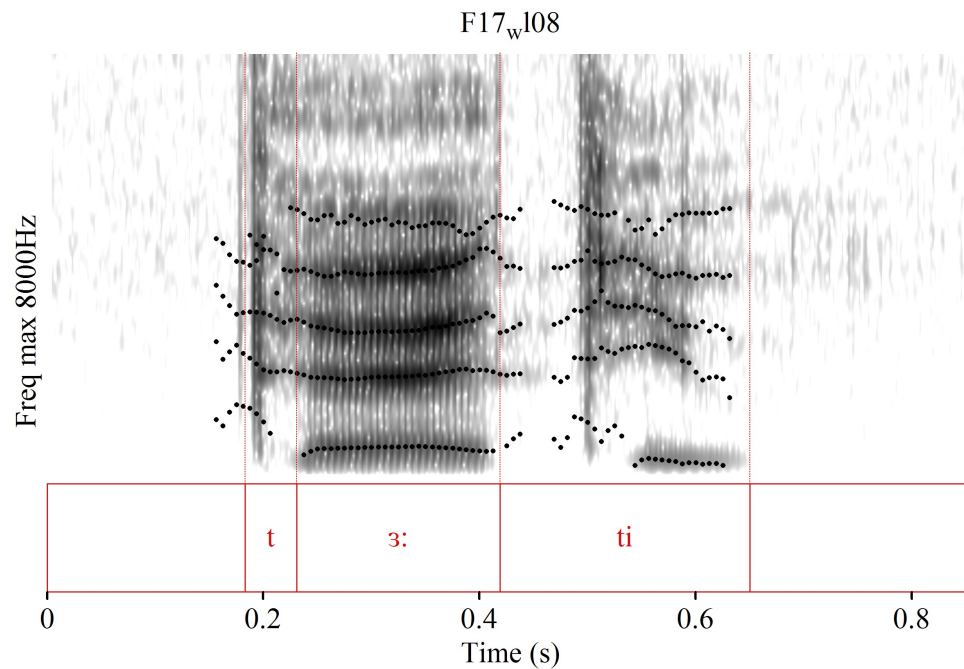


Figure 29: Spectrogram of word 8 of the word list as pronounced by informant F17. The word *thirty* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

and the lowest score of standard-like pronunciation within the group, is 1,45 times higher than the range of Group 2. This means that there was more variety among the Finnish informants.

The difference between Groups 1 and 2 was smaller in terms of the average performance of pronouncing the voiceless (inter-)dental fricative as a [θ]: a typical Austrian informant produced a [θ] in 3,5 and an average Finn in 4,6 out of seven instances. The performance distributions of the Finns are in both cases (total ‘th’ and /θ/) slightly more negatively skewed than the Austrian distributions (Appendix G). The majority of the Finns (13 out of 25 informants) pronounced six or all seven of the analysed occurrences of the voiceless (inter-)dental fricative with a [θ]. The percentage of Austrian informants with a performance score of at least six standard-like voiceless (inter)dental fricatives is relatively low: 14% or three out of 21 informants produced a [θ] in at least six out of seven instances in the text passage. In comparison to the scores of ‘th’ and /θ/, the average scores of producing a voiced (inter-)dental fricative are lower: Group 1 informants pronounced on average 5,5 and Group 2 informants 3,1 words out a total of 9 analysed occurrences of /ð/ in a standard-like manner. The performance distribution of Finnish informants is again skewed to the right, whereas the skewness of the Austrian performance distribution is

positive. Similar to the pronunciation of /θ/, a higher percentage of Group 1 informants produced a [ð] in at least eight out of nine instances: a third of all Finns and only one Austrian pronounced at least eight of the analysed words with a [ð].

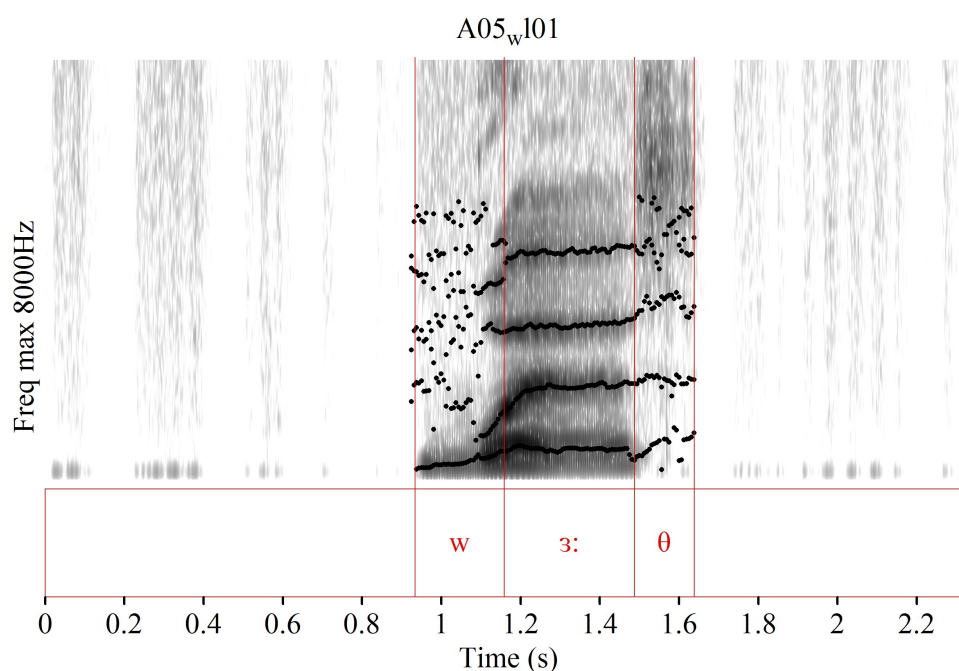


Figure 30: Spectrogram of word 1 of the word list as pronounced by informant A05. The word *worth* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

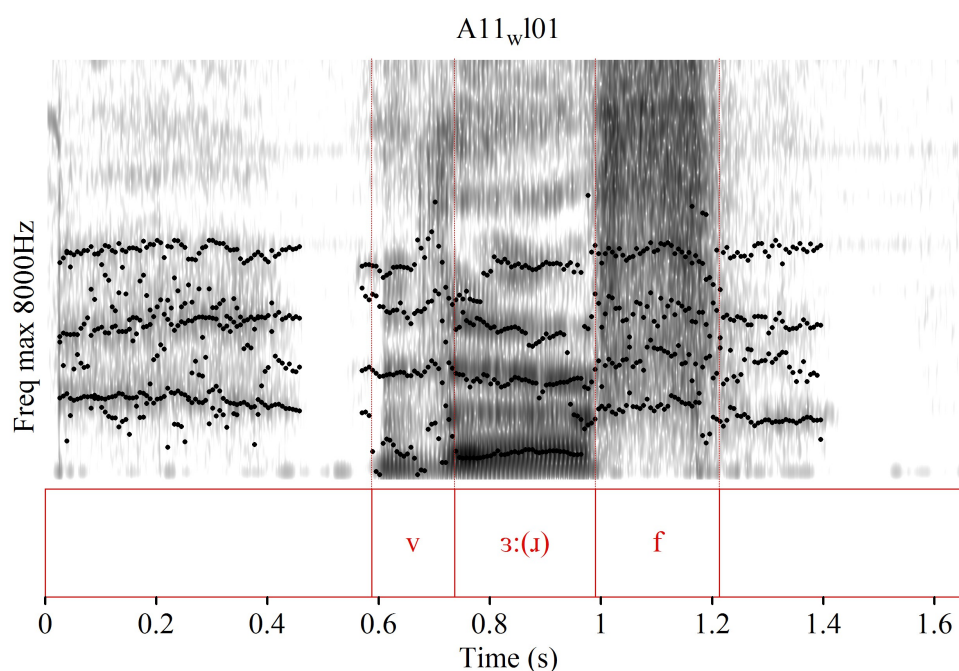


Figure 31: Spectrogram of word 1 of the word list as pronounced by informant A11. The word *worth* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

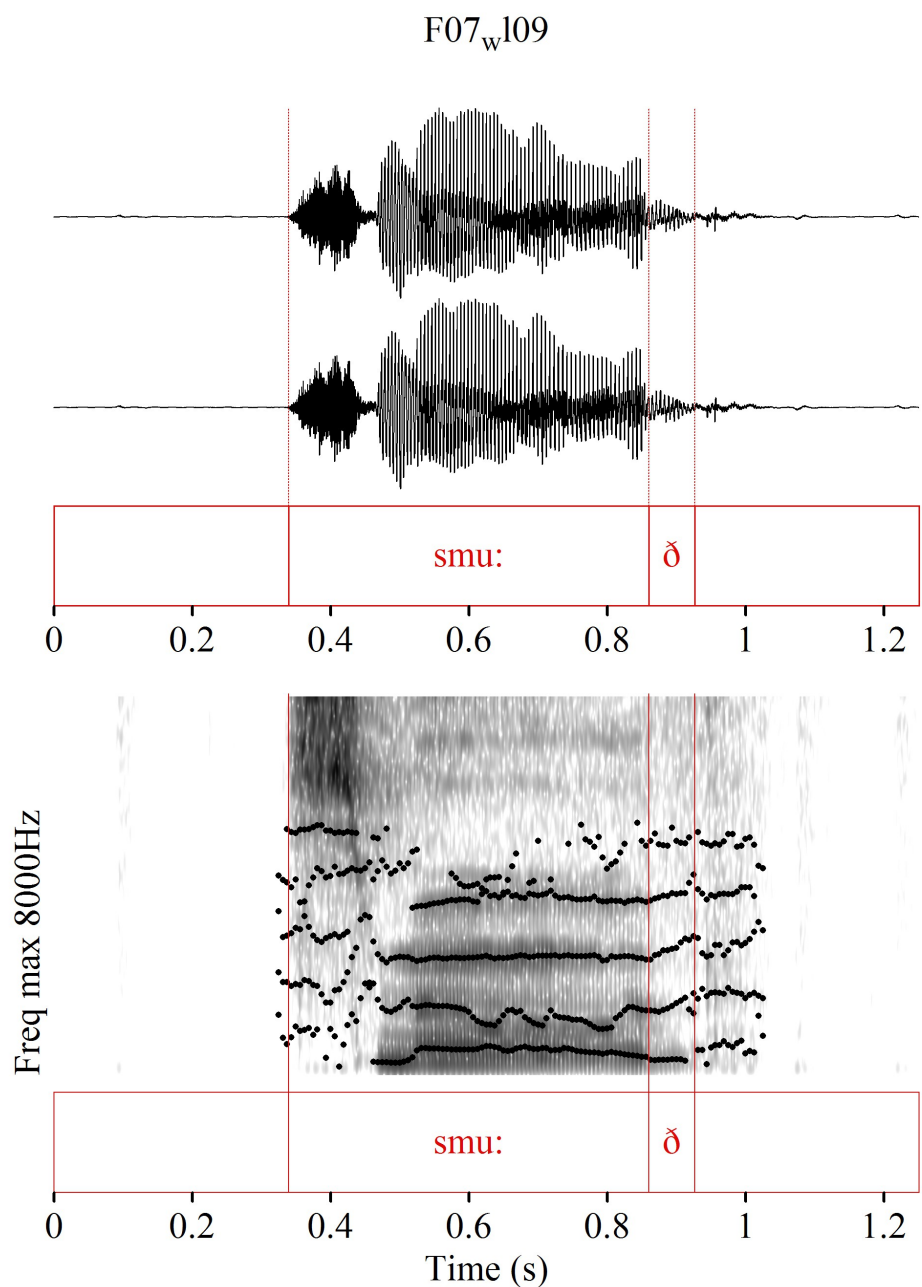


Figure 32: The waveform (above) and the spectrogram (below) of word 9 of the word list as pronounced by informant F07. The word *smooth* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

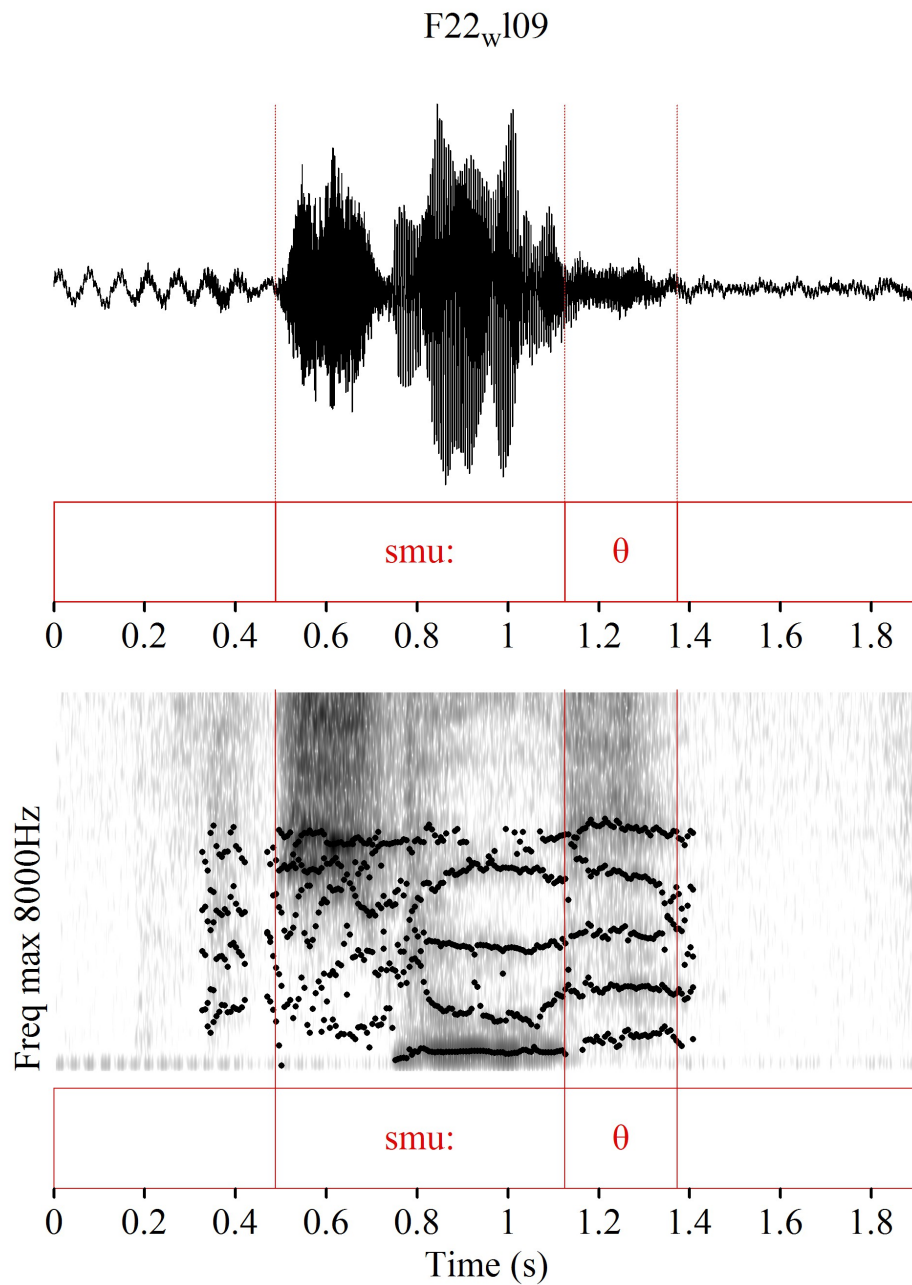


Figure 33: The waveform (above) and the spectrogram (below) of word 9 of the word list as pronounced by informant F22. The word *smooth* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

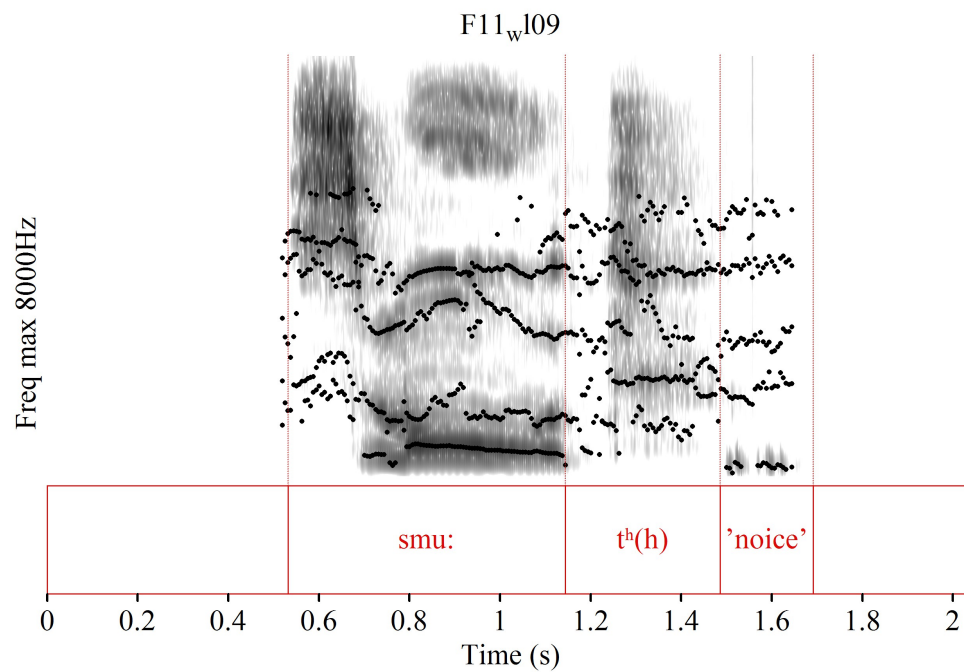


Figure 34: Spectrogram of word 9 of the word list as pronounced by informant F11. The word *smooth* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

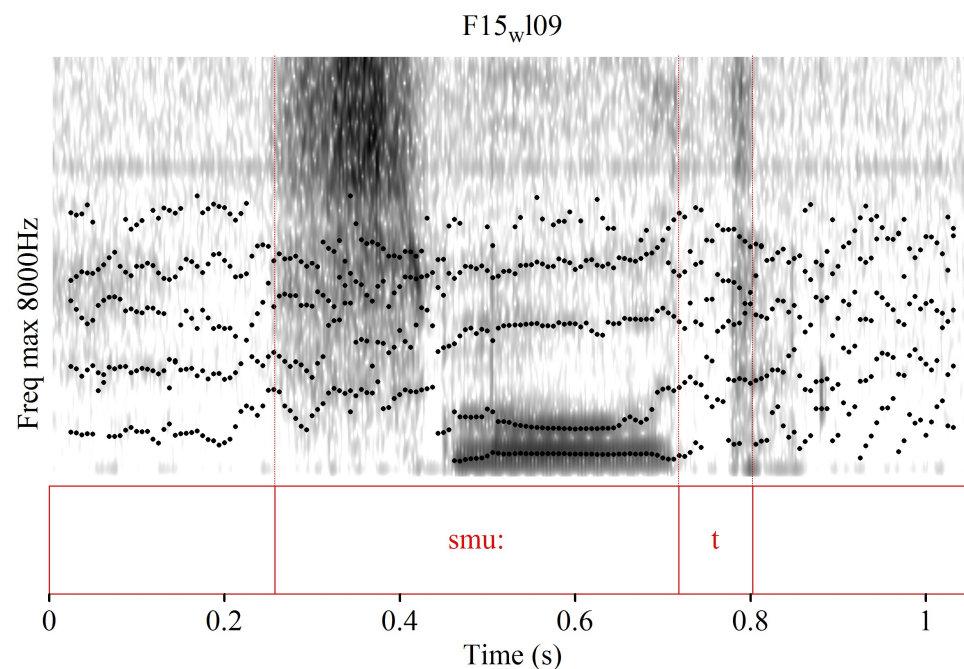


Figure 35: Spectrogram of word 9 of the word list as pronounced by informant F15. The word *smooth* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

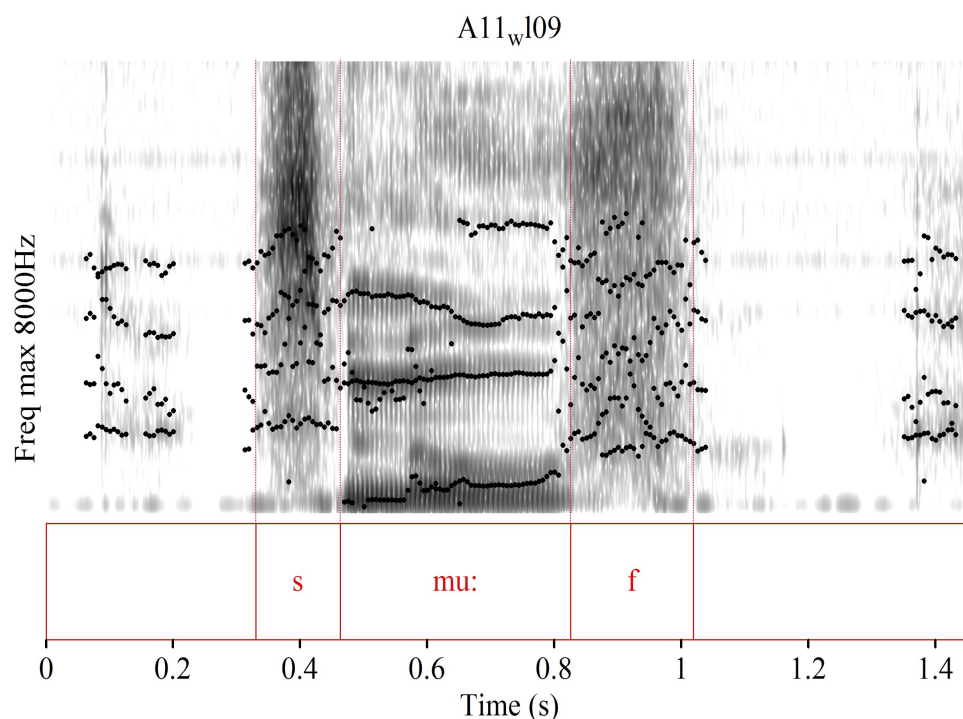


Figure 36: Spectrogram of word 9 of the word list as pronounced by informant A11. The word *smooth* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

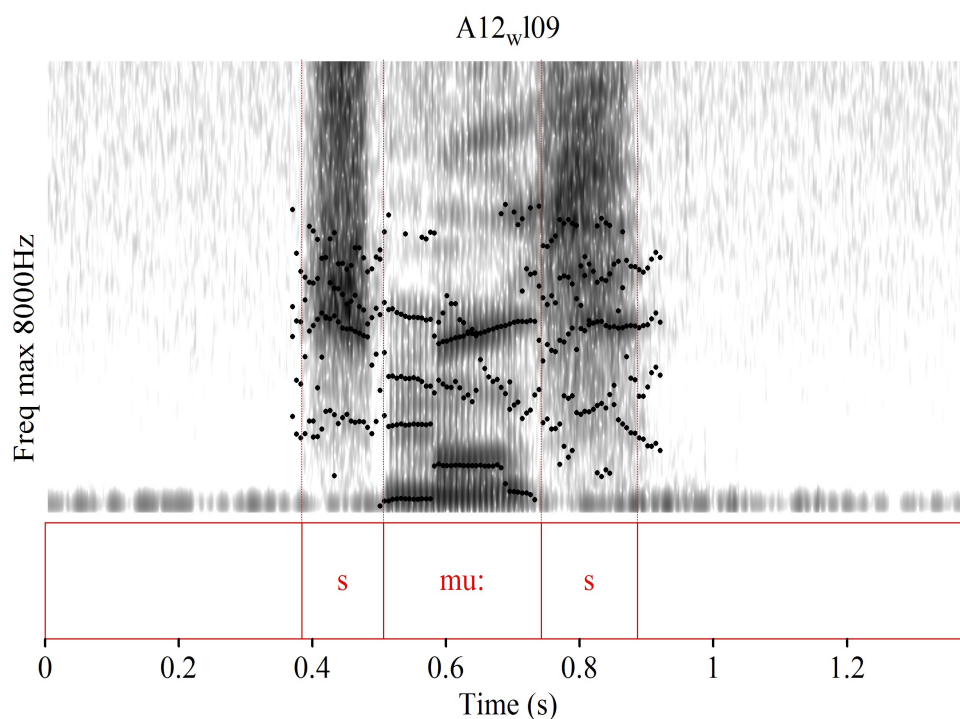


Figure 37: Spectrogram of word 9 of the word list as pronounced by informant A12. The word *smooth* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

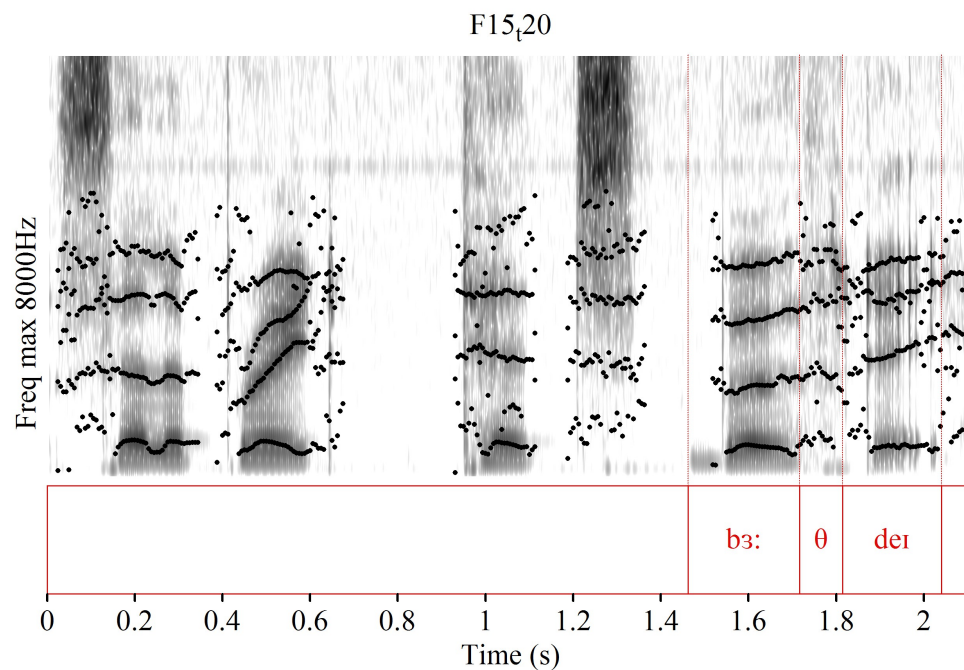


Figure 38: Spectrogram of string 20 of the text passage as pronounced by informant F15. The word *birthday* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

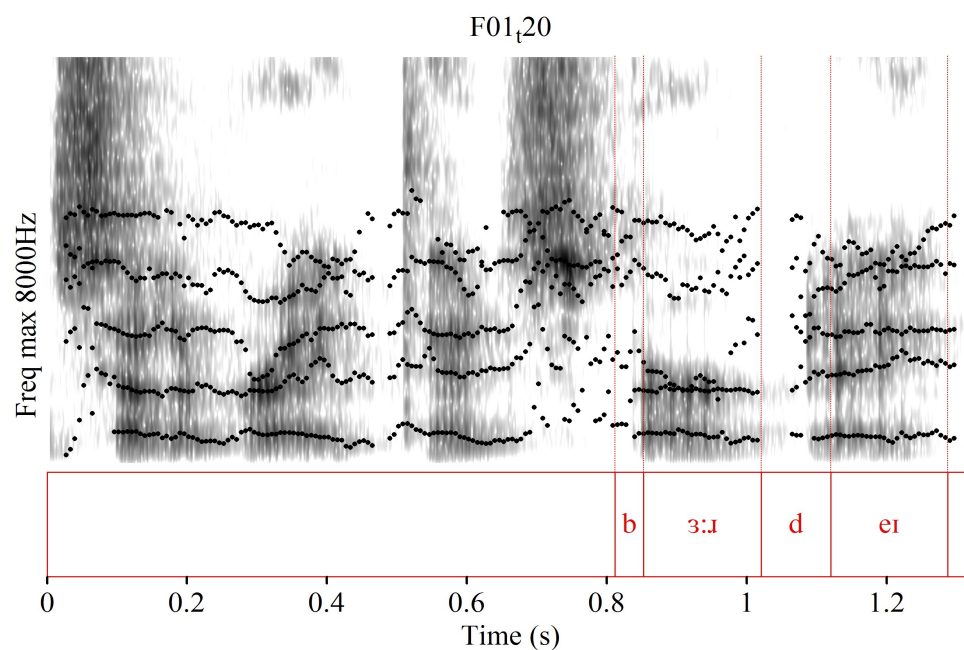


Figure 39: Spectrogram of string 20 of the text passage as pronounced by informant F01. The word *birthday* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

F14_t09

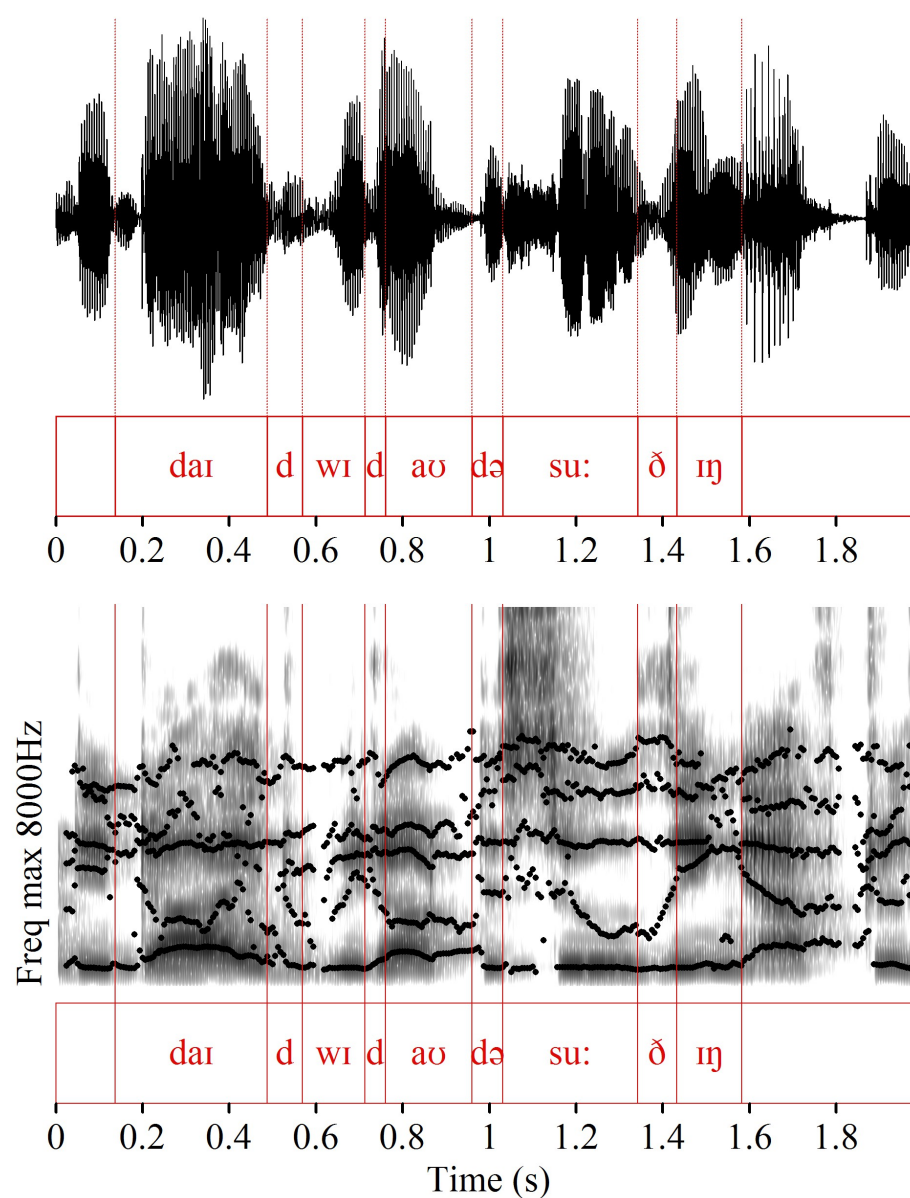


Figure 40: The waveform (above) and the spectrogram (below) of string 9 of the text passage as pronounced by informant F14. The phrase *died without the soothing* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

F17_t09

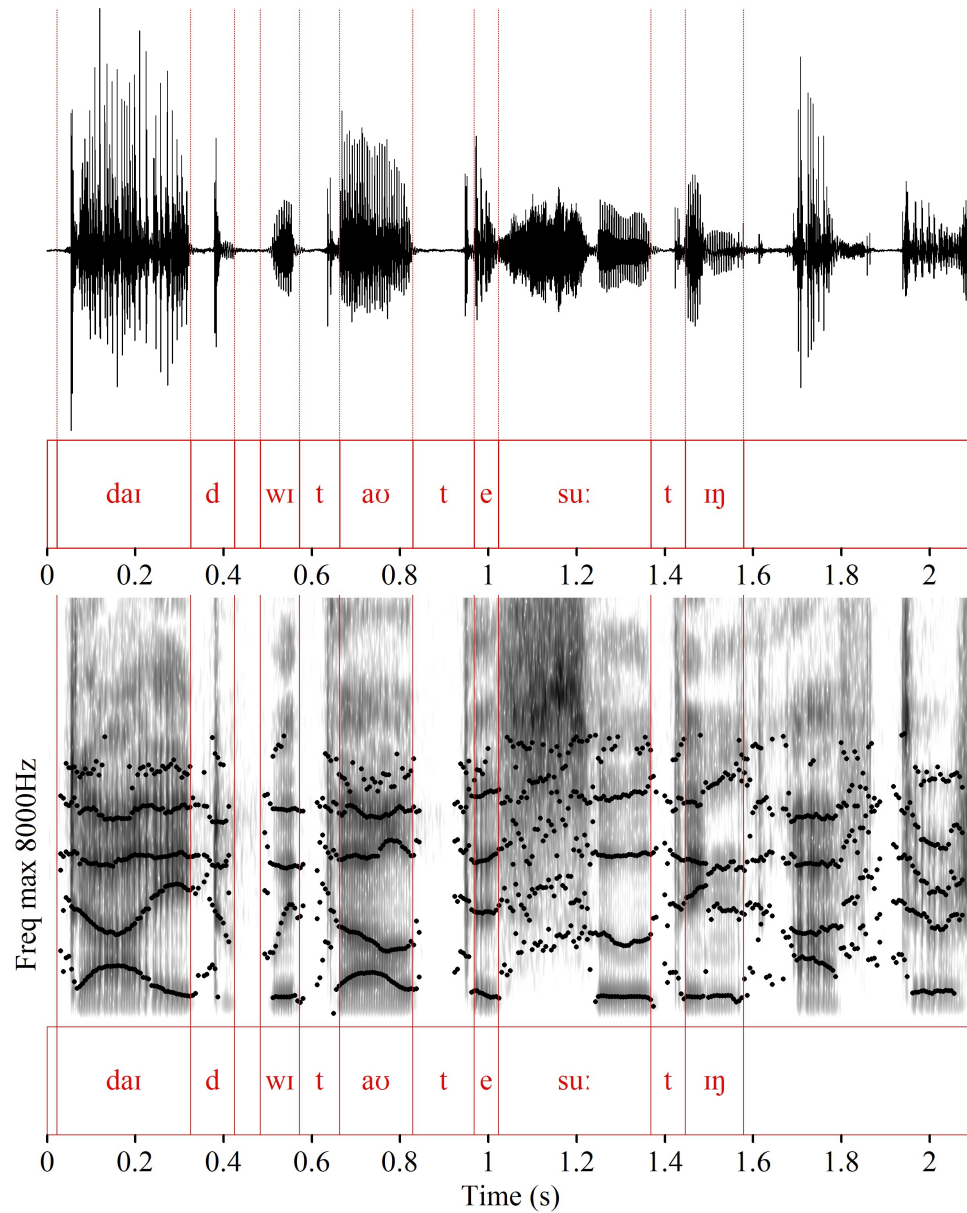


Figure 41: The waveform (above) and the spectrogram (below) of string 9 of the text passage as pronounced by informant F17. The phrase *died without the soothing* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

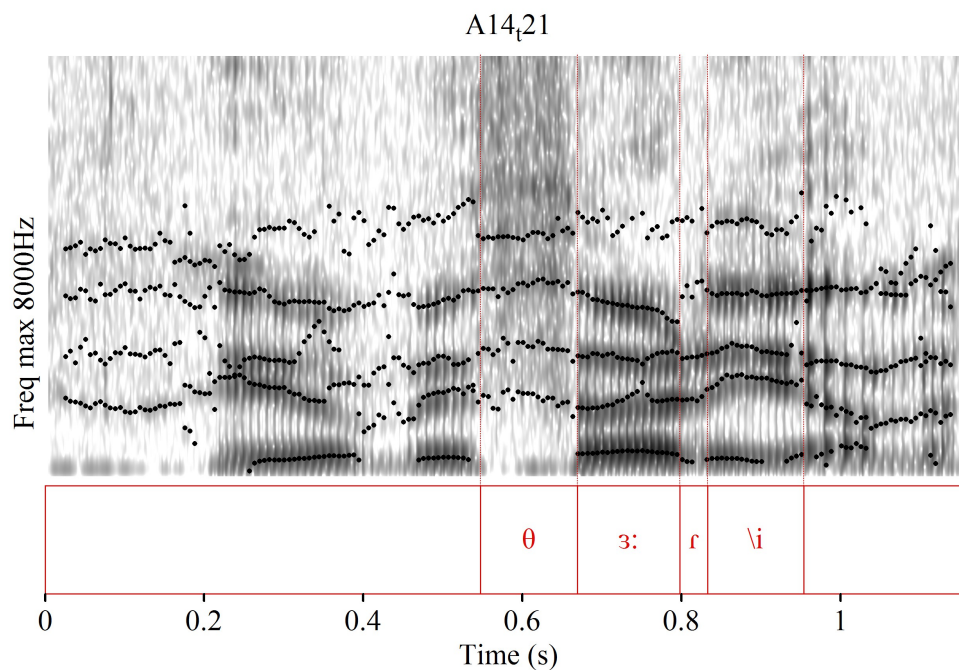


Figure 42: Spectrogram of string 21 of the text passage as pronounced by informant A14. The word *thirty* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

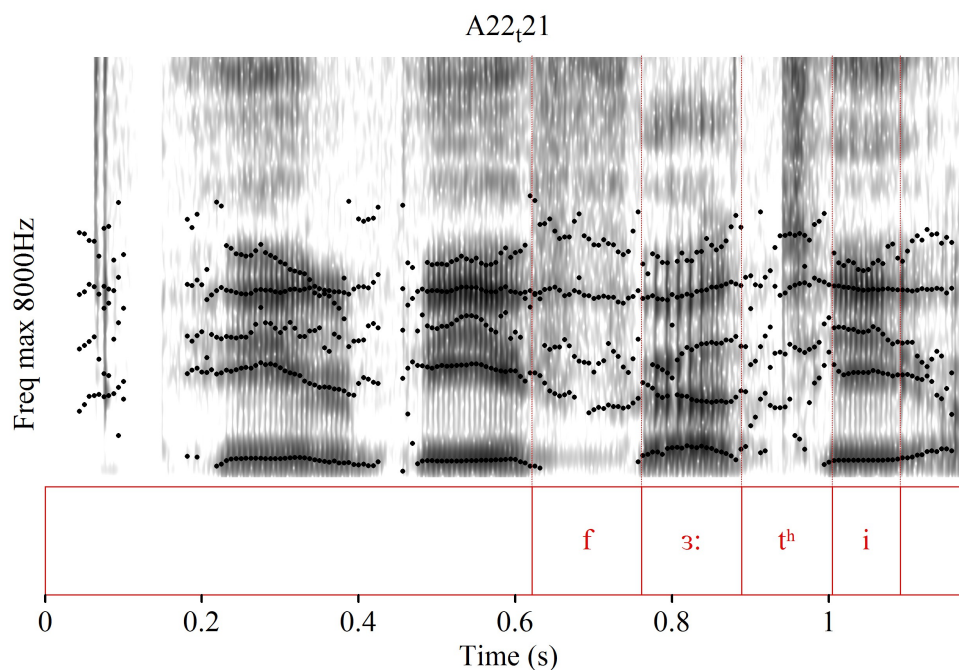


Figure 43: Spectrogram of string 21 of the text passage as pronounced by informant A22. The word *thirty* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

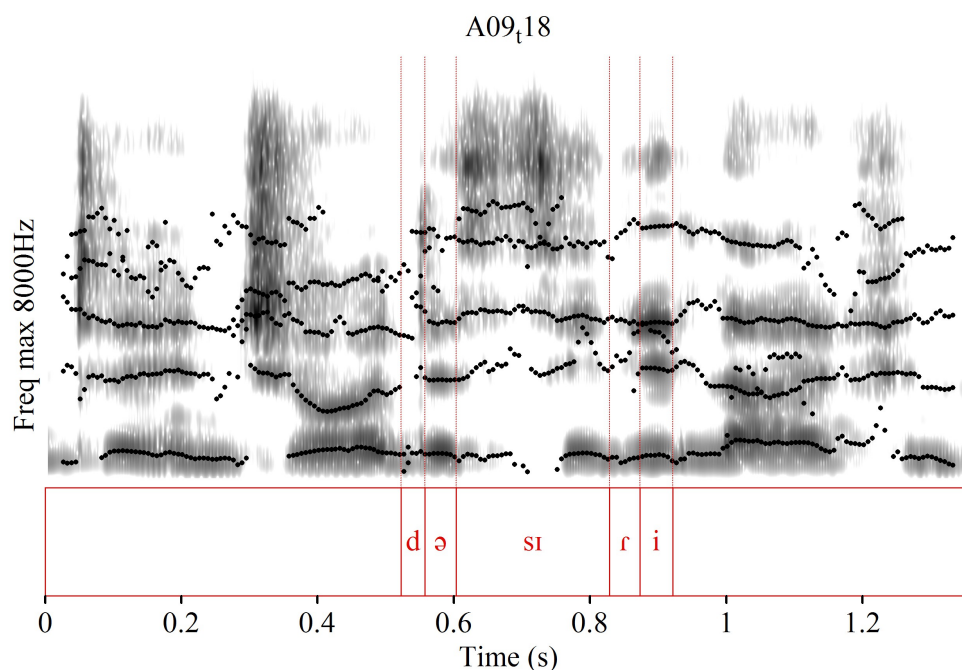


Figure 44: Spectrogram of string 18 of the text passage as pronounced by informant A09. The words *the* and *city* are marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

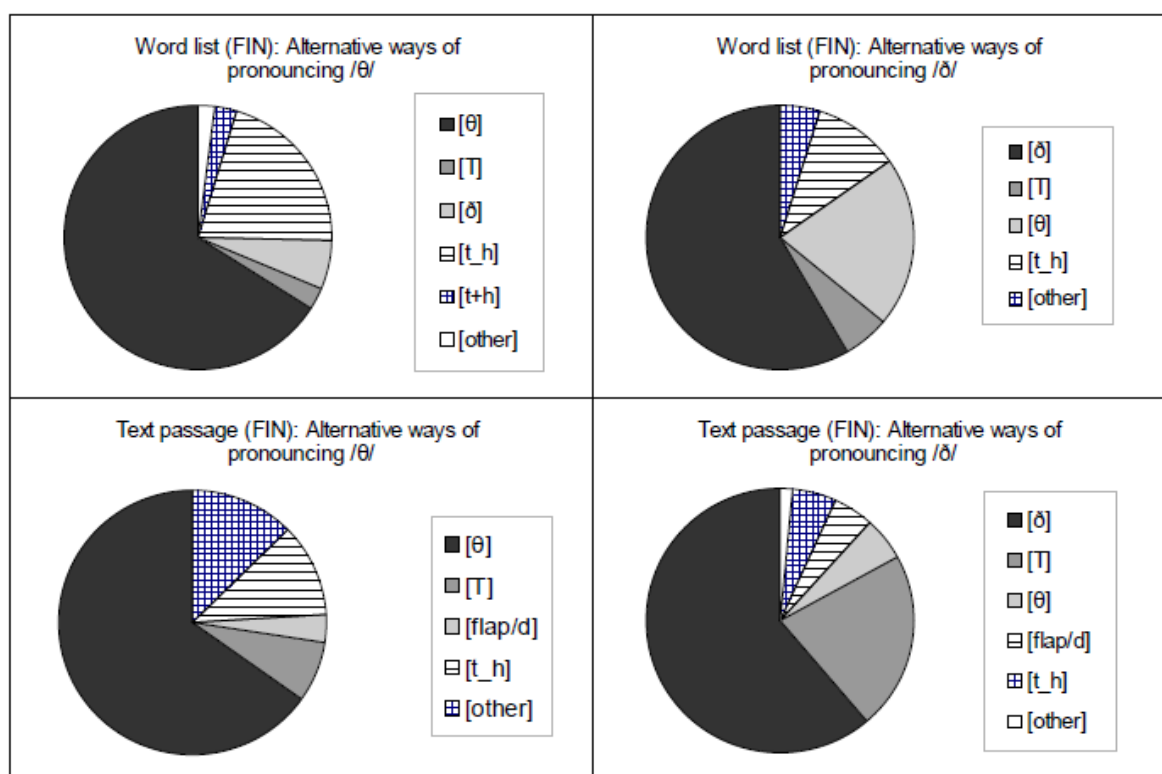


Figure 45: Group 1: Alternative ways of pronouncing /θ/ (left) and /ð/ (right). Results from both the word list (top) and the text passage task (bottom).

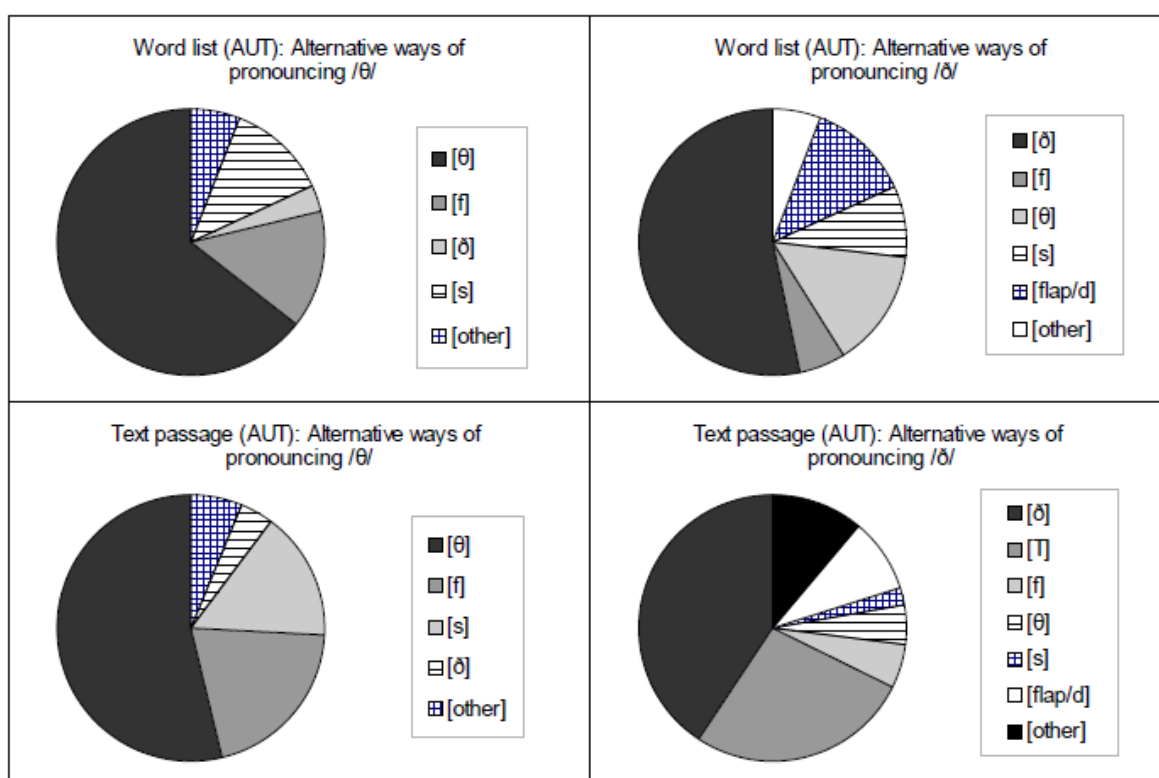


Figure 46: Group 2: Alternative ways of pronouncing /θ/ (left) and /ð/ (right). Results from both the word list (top) and the text passage task (bottom).

4.3 Voiced plosives /b/, /d/, and /g/ in word-final position

The analysis of the word-final voiced plosives was based on the following criteria: the absence of a post-release phase i.e. aspiration, and the visible voicing bar during closure in the spectrograms and periodic waves in the waveform. In this analysis, all word-final plosive sounds that were pronounced with aspiration were analysed as voiceless. The main criteria for differentiating between word-final non-aspirated voiceless and voiced plosives was the visible voicing (voicing bar and periodic waves) during the compression phase. Since all of the analysed word-final voiced plosives in the word list task had a voiceless counterpart (minimal pair), also the lenis character of the preceding vowel, that is the duration of the vowel sound, could additionally be assessed from the word list samples. Figure 47 illustrates the above listed criteria for differentiating between voiced, voiceless aspirated and voiceless non-aspirated sounds in pronouncing the word *said*. The spectrogram on the left represents the voiced pronunciation of the word-final /d/. The release burst of the [d] is short and non-aspirated. The voicing bar and the periodic waves are visible during the closure. Additionally, the duration of the vowel [e] preceding the [d] is twice as long as the durations of the vowels in the other two cases. The second

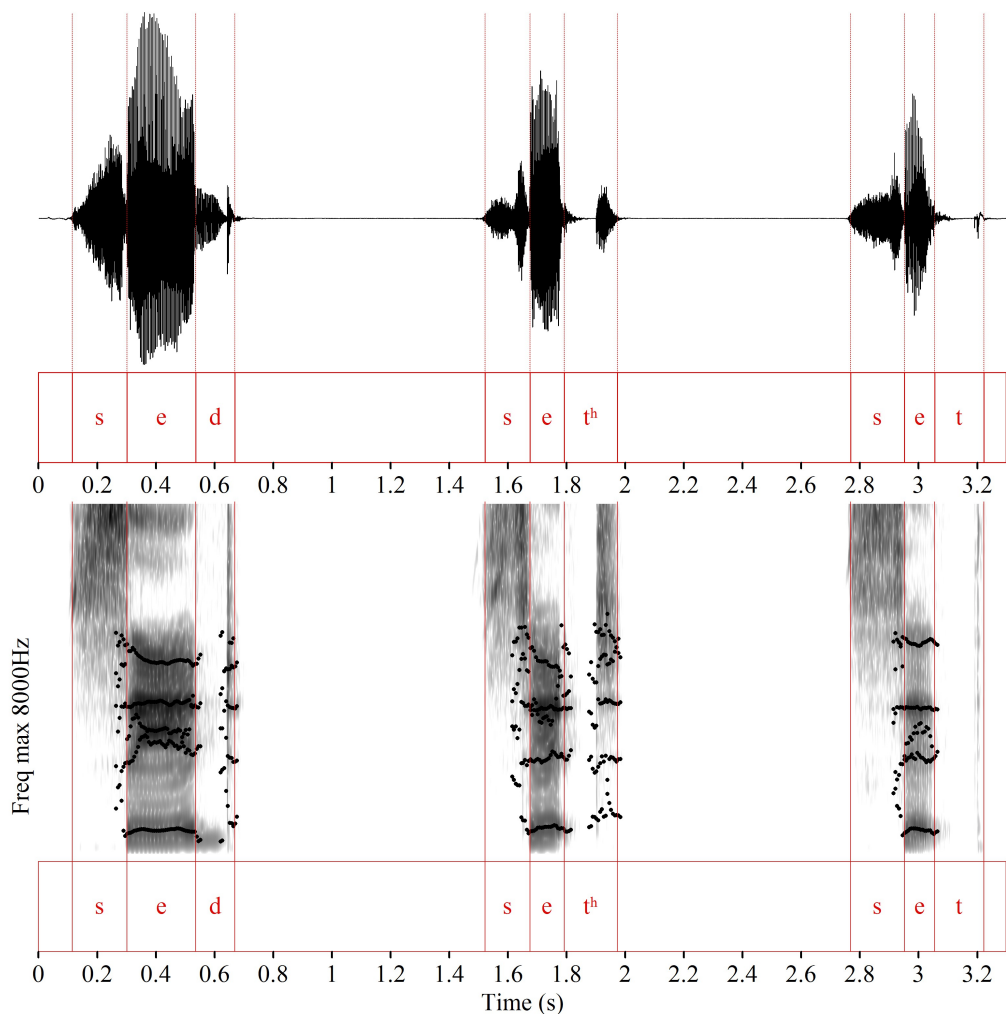


Figure 47: The waveforms (above) and the spectrograms of the word *said* as pronounced by the author in three alternative ways: [sed] (left), [seth] and [set] (right). The sounds are marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

allophone for /d/ in Figure 47 is the [t^h]. This time there is no sign of voicing during the closure and the release burst is followed by aspiration. The third allophone of /t/ is the voiceless non-aspirated [t]. The release burst resembles the first allophone, but this time there is no voicing during the compression phase and the preceding vowel is also shorter than the [e] before the voiced allophone.

Figure 48 presents the total percentage of voicing in both pronunciation tasks. The difference between the two groups was greater in the word list task. 80% of all word-final voiced plosives in the word list task were pronounced as voiced by Group 1 informants, whereas only third of all word list occurrences of word-final voiced plosives were produced with voicing by Group 2 informants. The above mentioned percentages include pronunciations in which voicing as periodic waves in the waveforms could be detected

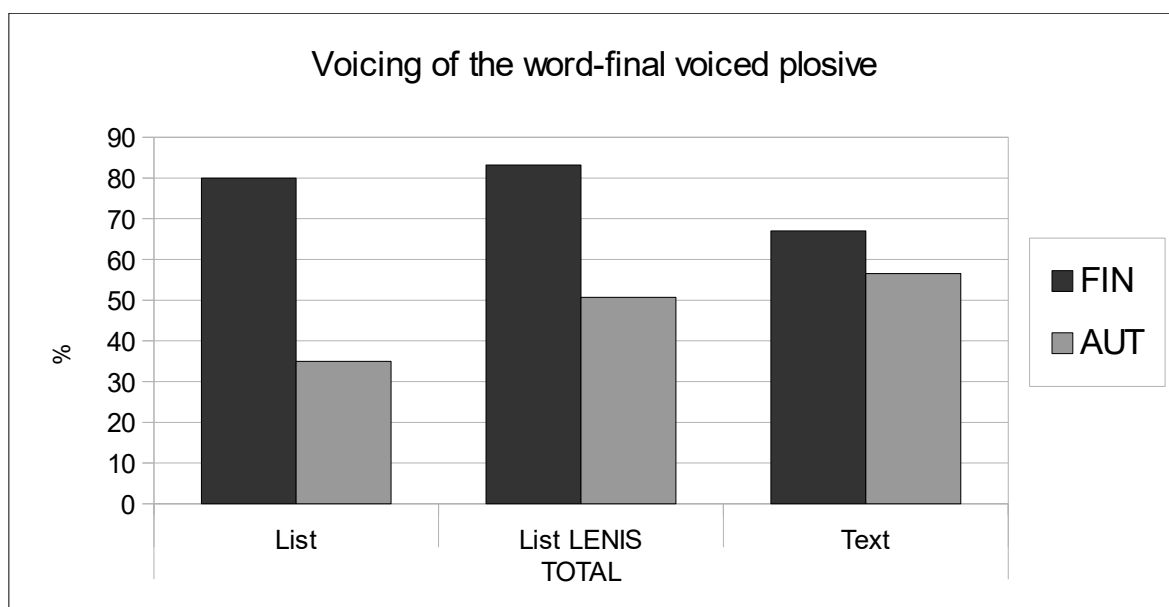


Figure 48: Total percentages of voicing the voiced word-final plosives in the word list (left) and in the text passage (right) tasks. The bars in the middle present the percentages of voicing and/or lenis character in the word list task.

from the visualisations. When the samples with no visual sign of voicing but with clearly (at least 1,5 times) longer preceding vowels are taken into account, the difference between the two informant groups decreases. 83,2% of all occurrences were pronounced with voicing and/or with a longer vowel by Group 1 and 50,7% by Group 2 informants (bars in the middle in Figure 48). In the text passage task, the difference between the groups decreased even further. The percentage of instances of voiced plosives in the Finnish samples decreased to 64,5%, while the Austrian percentage of voicing the word-final voiced plosives increased to 55,4%. All in all, the word-final voiced plosives were more commonly pronounced in a standard-like manner (voicing and/or longer vowel) by Group 1 informants. They produced a standard-like sounds in 72,3% of all instances (both tasks). The total percentage for Group 2 is 53,4%. According to the χ^2 -test, results from the analysis of the word list data are statistically significant at the 5% significance level. The p-value of all instances of word-final voiced plosives in both tasks also indicates a strong correlation between the informant group and the pronunciation. There is not enough data, though, to show that the pronunciation of word-final voiced plosives in the text passage task would depend on the informant group at the chosen significance level, for the calculated p-value of 0,09 is greater than 0,05 (Table 3).

Variables (Controlled x measured)	p-values from the χ^2-test ($\alpha=5e-02$)
Group x Pronunciation of word-final voiced plosives (LIST)	4,38e-15
Group x Pronunciation of word-final voiced plosives (LIST: LENIS)	3,20e-08
Group x Pronunciation of word-final voiced plosives (TEXT)	9,32e-02
Group x Pronunciation of word-final voiced plosives (TOTAL: LENIS)	1,06e-06
FIN: Voicing x Place of articulation (LIST)	77,03e-02
FIN: Voicing x Place of articulation (LIST: LENIS)	95,45e-02
FIN: Voicing x Place of articulation (TEXT)	59,18e-02
AUT: Voicing x Place of articulation (LIST)	77,64e-02
AUT: Voicing x Place of articulation (LIST: LENIS)	96,87e-02
AUT: Voicing x Place of articulation (TEXT)	6,39e-02

Table 3: p-values from the χ^2 -test for the word-final voiced plosives /b/, /d/ and /g/.

In addition to the general voicing of the word-final plosives, possible connections between the pronunciation and the place of articulation were investigated. As illustrated in Figure 49, the differences in percentages are relatively small for Group 1 (max. 10%), and the p-values from the χ^2 -test indicate no statistically significant correlation between Group 1 pronunciation and the place of articulation (Table 3). Similarly, the word list samples of Group 2 show no connection of between voicing and the place of articulation. The results from the analysis of Group 2 text passage samples are more varied than those from Group 1 and the differences between the three sounds are greater. The majority of the analysed word-final voiced bilabial plosives (/b/) were produced with voicing by the Austrian informants, whereas only around 45% of the voiced velar stops (/g/) were pronounced as voiced. Of all the p-values describing the relationship between place of articulation and pronunciation, the p-value for the Austrian text passage samples is the lowest implying a statistically significant correlation at the 10% significance level, but there is not enough data to show that there is connection at the chosen significance level of 5% (Table 3).

The word list word least often pronounced in a standard-like manner by Group 1 informants was the word *cab*. 18 out of 25 informants pronounced the word either with visible voicing or with a longer preceding vowel. In the samples of the seven remaining Finnish informants, the pronunciation of the word *cab* (Figure 50) closely resembled the pronunciation of the voiceless counterpart *cap* (Figure 51). An example of a voiced bilabial plosive at the end of the word *rib* is presented in Figure 52. In comparison to the pronunciation of *rib*, the vowel in Figure 53 is shorter and there is no trace of periodicity in the waveform of the word *rip*. The most ‘difficult’ word in the word list task for the

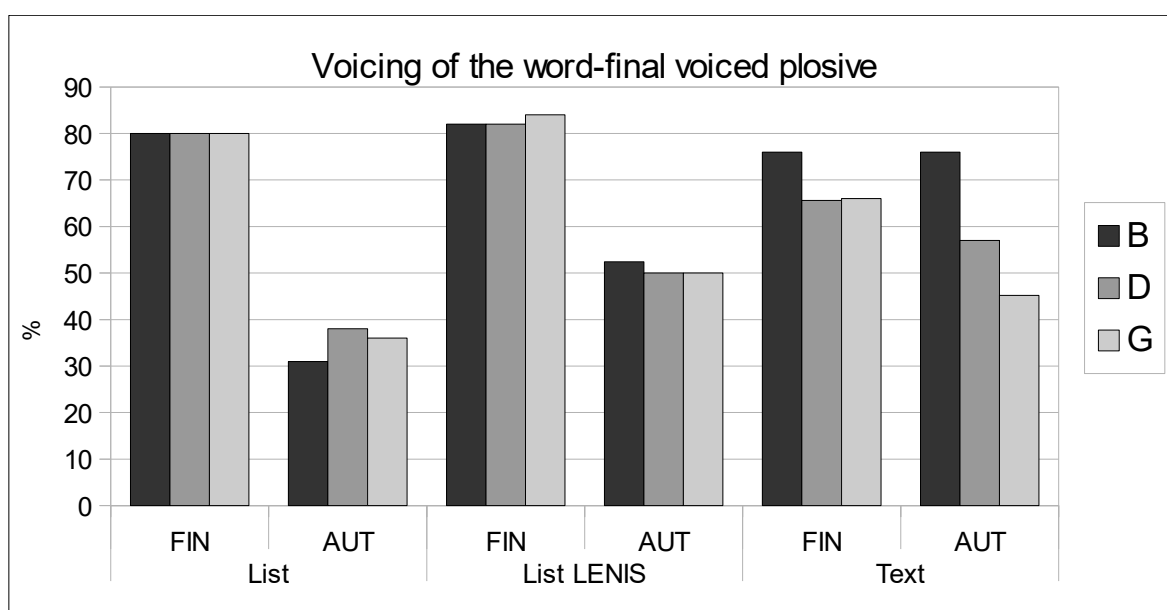


Figure 49: Voicing the word-final voiced plosive. Individual percentages for each of the three places of articulation in both pronunciation tasks: word list (2 sets of bars on the left), word list including lenis character of the preceding vowel (2 sets of bars in the middle) and text passage (2 sets of bars on the right).

Austrian informants was the word *ipad*: only 6 out of 21 informants produced a voiced plosive and/or a longer preceding vowel when pronouncing the word. The rest of Group 2 informants pronounced the word frequently with aspiration. In many cases, like in Figure 54, the spectrogram of the last syllable of the word *ipad* was identical with the spectrogram of *pat* in Figure 55. The standard-like pronunciations of the word-final alveolar plosives are presented in Figures 56 and 57: here the word *said* was pronounced both with a longer vowel and with voicing during the closure, and the /t/ in *set* was produced with aspiration and preceded by a shorter vowel sound. Figures 58 and 59 illustrate the pronunciation of the words *log* and *lock* by an Austrian informant. These two samples were included to demonstrate the lenis-pronunciation of a voiced plosive. Both waveforms show no sign of periodic waves during the closure, but the duration of the vowel in Figure 58 is markedly longer than the vowel in Figure 59. Additionally, the word-final plosive in Figure 59 was pronounced with aspiration. In contrast to Figure 58, voicing is clearly visible as periodicity during the compression phase in Figure 60, which presents the waveform and the spectrogram of the word *wag* as pronounced by a Finnish informant. Additionally, the duration of the vowel sound is 1,6 times longer than the vowel in the word *whack* (Figure 61).

The Group 1 percentage of standard-like pronunciation was lower in the less formal task. This was partially because many of the words with a word-final voiced plosive were

followed by words that begin with a vowel. When a non-stressed word with a word-final voiced alveolar plosive was followed by a vowel, the /d/ was often pronounced as a flap [ɾ] like in Figure 62. The Group 2 samples from the text passage task contained more instances of voicing the word-final voiced plosives. This could at least partially be explained by the sounds following the word-final voiced stops. Similar to /d/, the word-final stops that were followed by a vowel were often like word-medial plosives and were thus commonly pronounced with voicing. For example, the word *Meg* in string 23 was generally pronounced together with the following word *and* forming something like [megən(d)]. In comparison, the *Meg* in string 15 is the last word of the sentence i.e. it is followed by a pause. The results of the analysis support this theory: the *Meg* in sentence 23 was pronounced with voicing by 21 Finnish and 14 Austrian informants, and the same word in sentence 15 was voiced by 12 Finns and 5 Austrians. An example of the devoicing of the word-final voiced velar plosive is illustrated in Figure 63. The standard-like pronunciation of the only instance of the word-final (although followed by a vowel) voiced bilabial plosive in the text passage can also be seen in Figure 63.

As the above presented results suggest, the differences between the two groups were greater in the more formal pronunciation task. The performance statistics of individual informants supports these results: in the word list task an average Finnish informant produced a standard-like sound in five out of six occurrences whereas an Austrian informant pronounced the word-final voiced plosive with voicing and/or with a longer preceding vowel on average in three out of six instances. The Finnish word list data is slightly skewed to the right for both the median and the mode are greater than the arithmetic mean (Appendix G). The majority of Group 1, i.e. 18 out of 25 informants, produced a voiced plosive in at least five of the six instances. Four out of 21 Group 2 informants pronounced five or more of the word-final plosives in a standard-like manner. In the text passage task, the difference in arithmetic mean of individual performance between the two groups is smaller: on average, Group 1 informants produced a voiced plosive in 5,5 and Group 2 informants in 4,6 out of eight occurrences. Similar to the word list data, there were more Finns (9 out of 25) who pronounced the word-final plosive with voicing in at least seven out of eight instances. In comparison, only one Austrian produced a voiced plosive in at least seven occurrences in the text passage task.

F17_w123

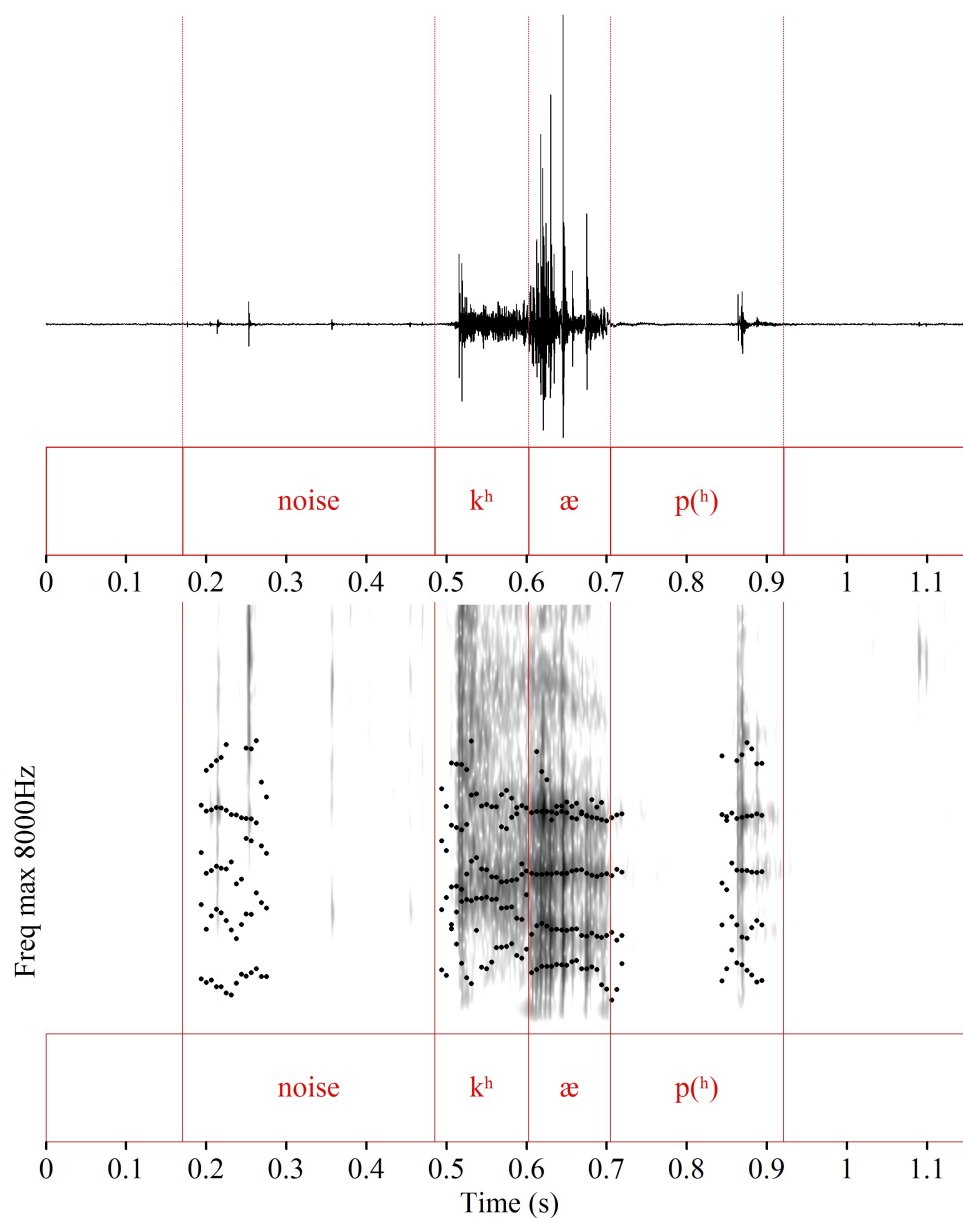


Figure 50: The waveform (above) and the spectrogram (below) of word 23 of the word list as pronounced by informant F17. The word *cab* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

F17_w112

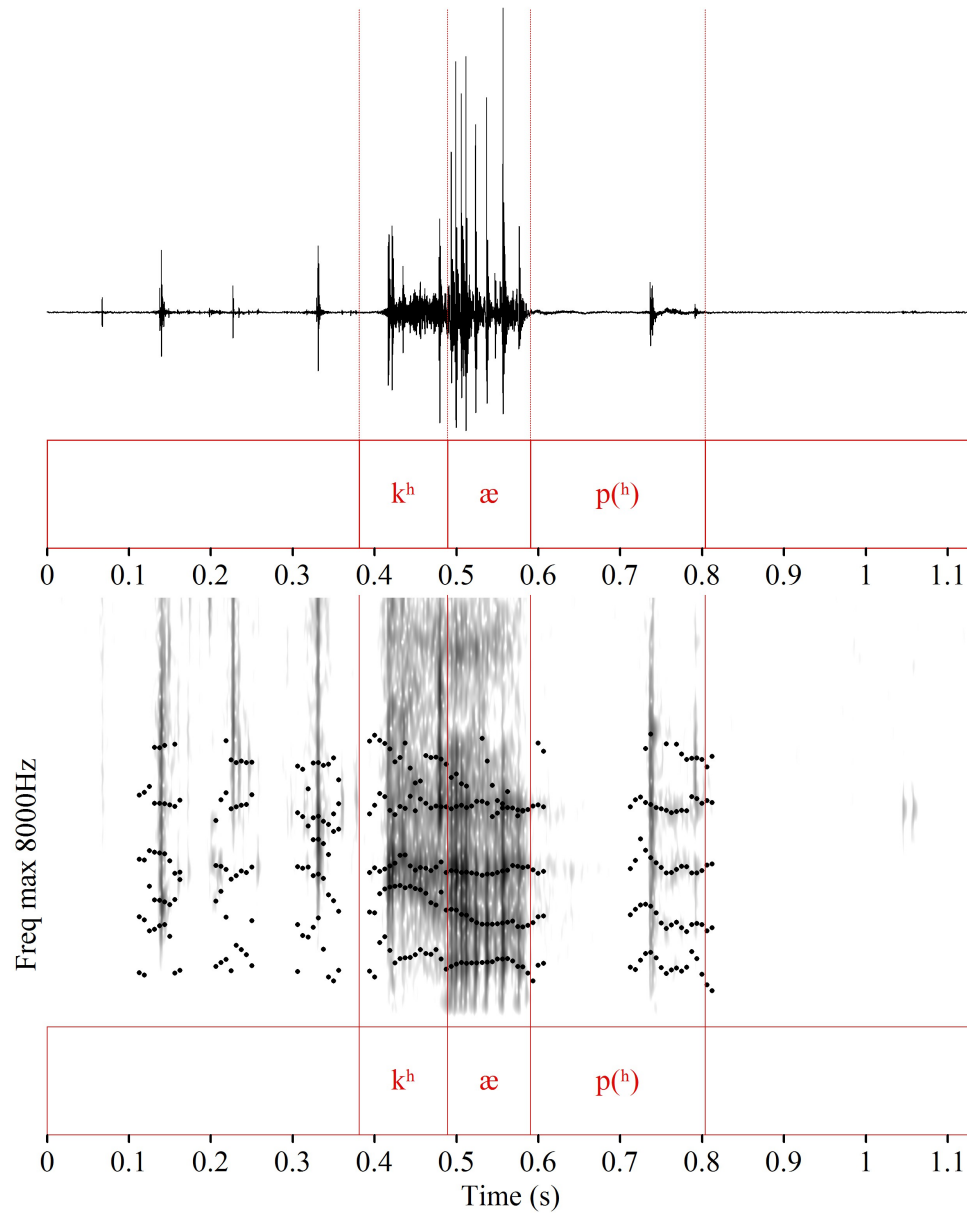


Figure 51: The waveform (above) and the spectrogram (below) of word 12 of the word list as pronounced by informant F17. The word *cap* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

F07_w121

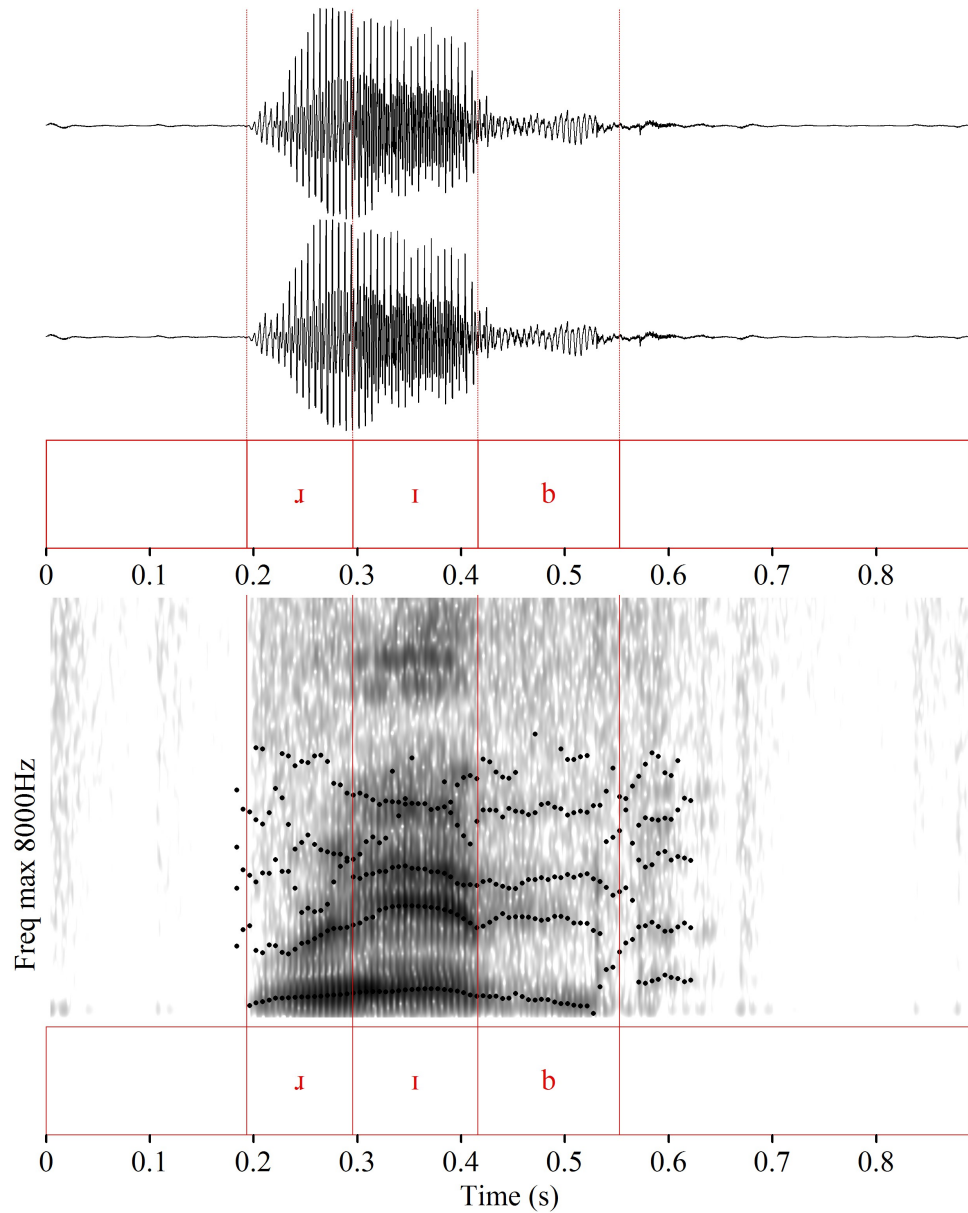


Figure 52: The waveform (above) and the spectrogram (below) of word 21 of the word list as pronounced by informant F07. The word *rib* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

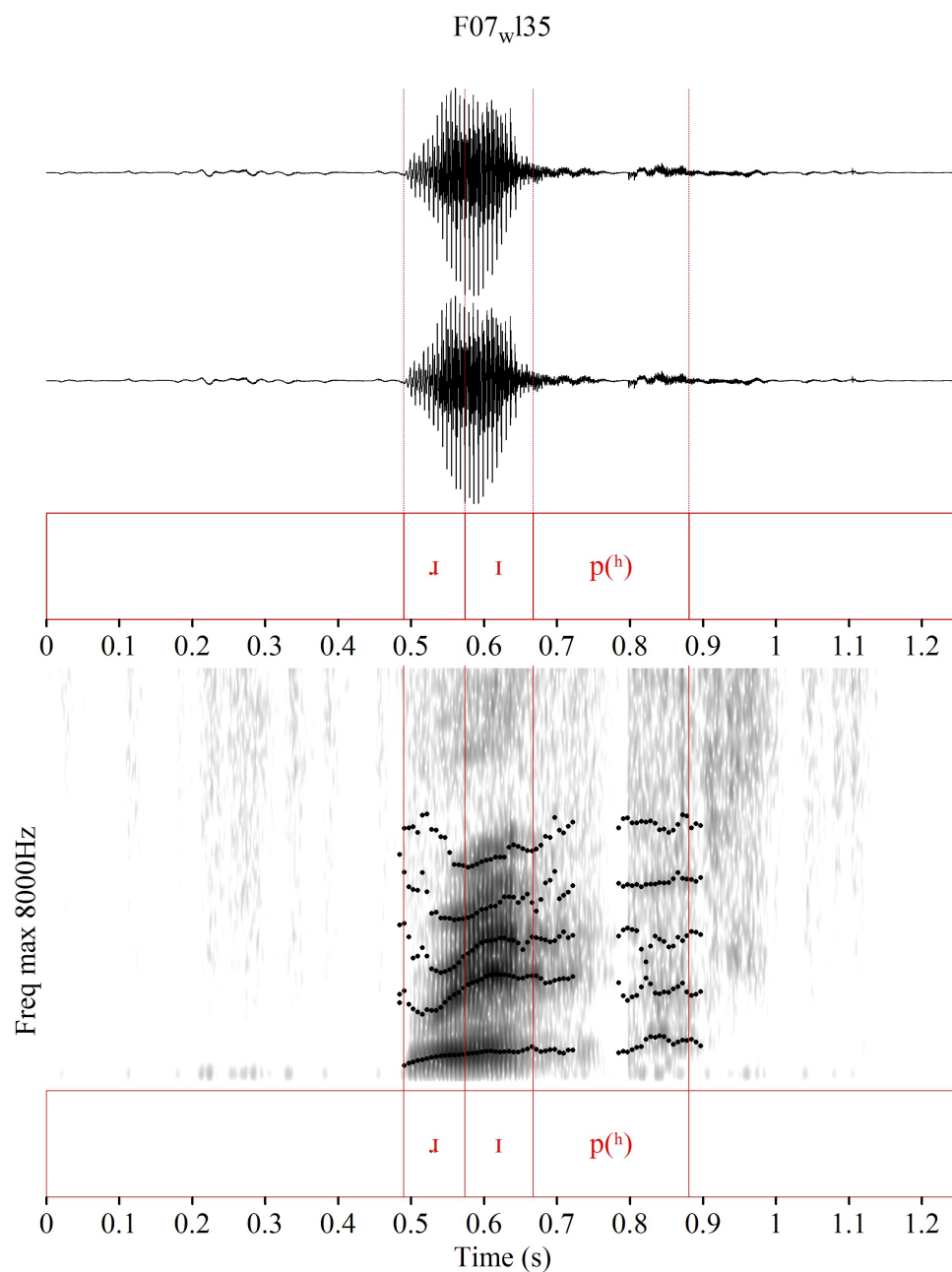


Figure 53: The waveform (above) and the spectrogram (below) of word 35 of the word list as pronounced by informant F07. The word *rip* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

A01_w103

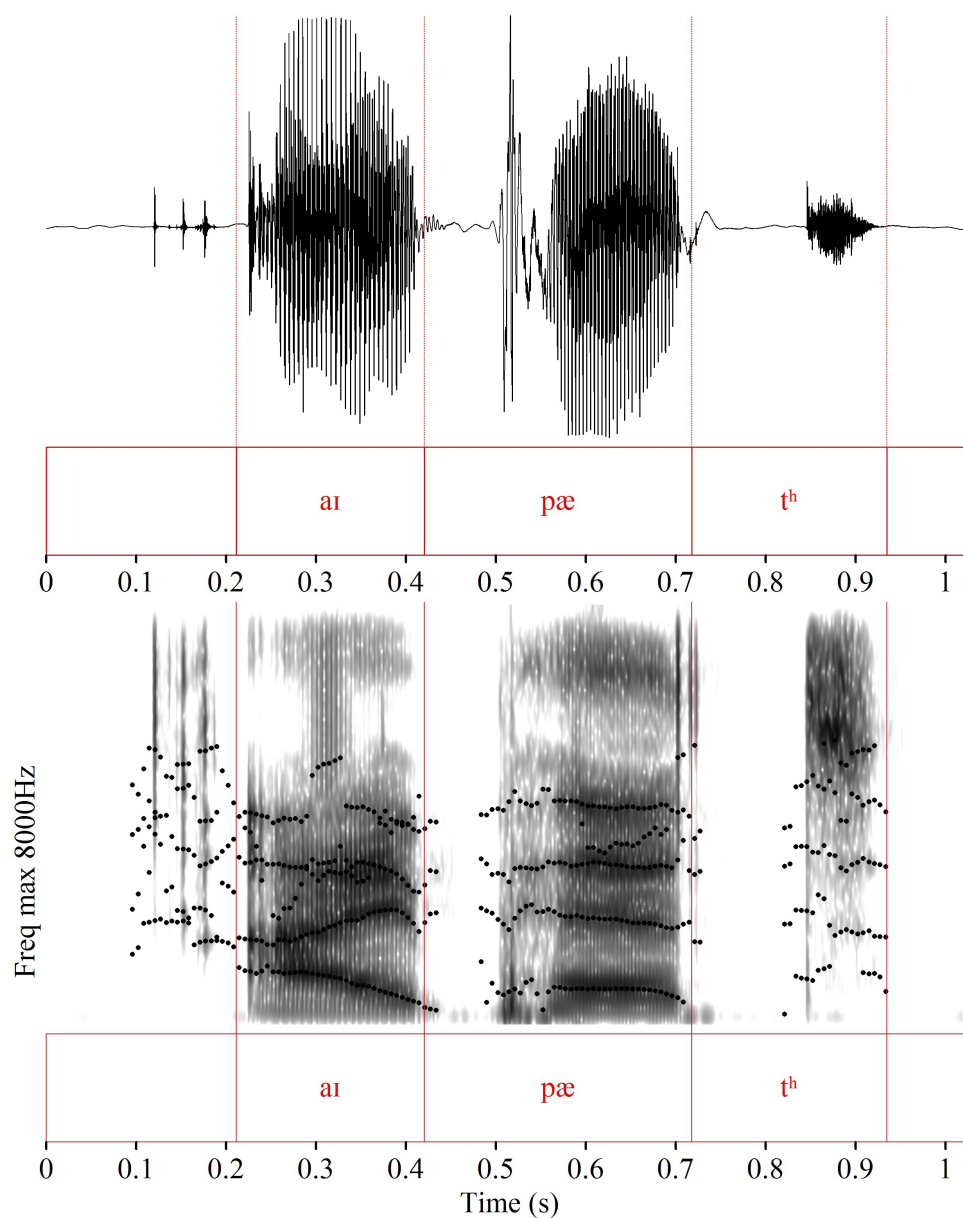


Figure 54: The waveform (above) and the spectrogram (below) of word 3 of the word list as pronounced by informant A01. The word *ipad* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

A01_w117

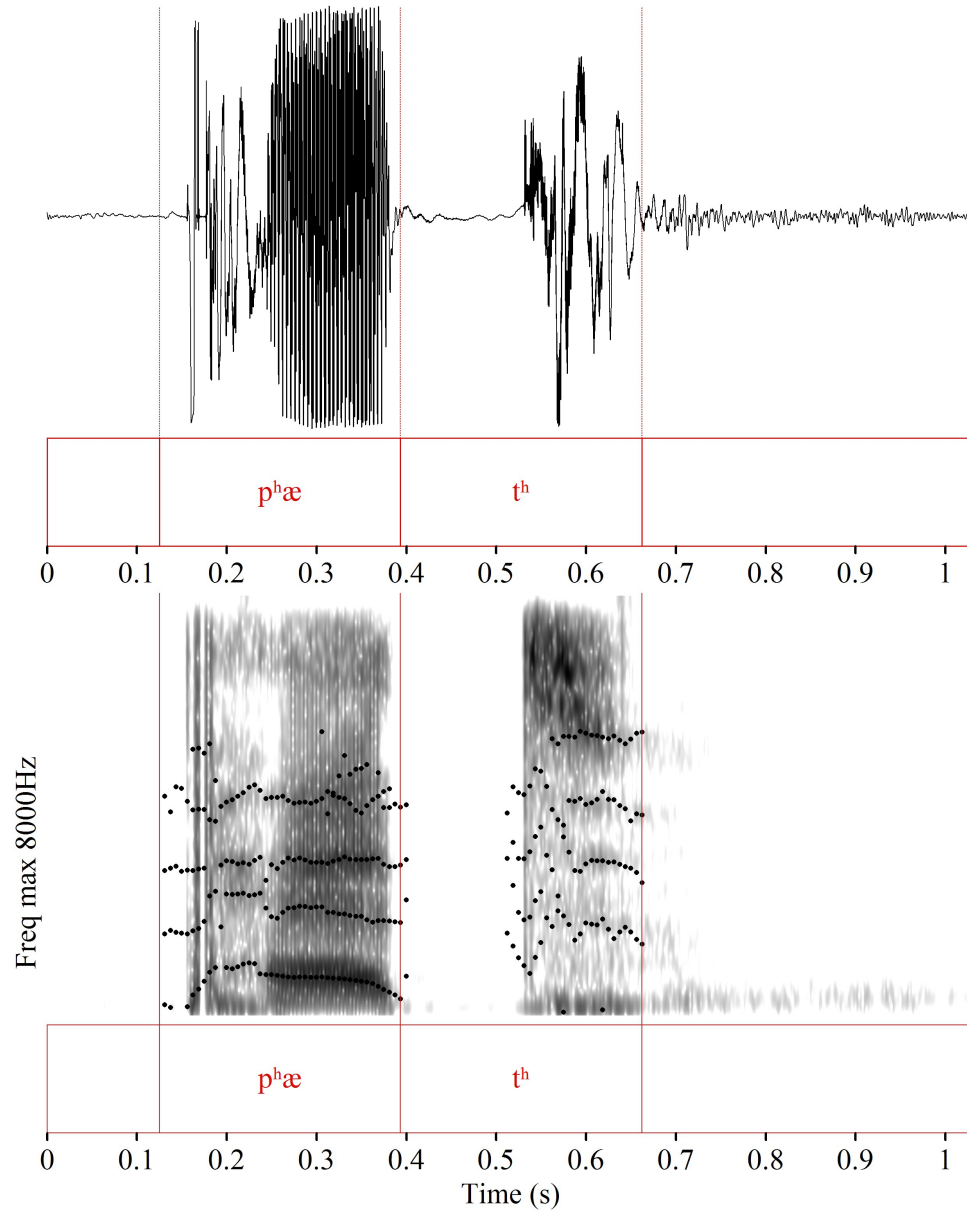


Figure 55: The waveform (above) and the spectrogram (below) of word 17 of the word list as pronounced by informant A01. The word *pat* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

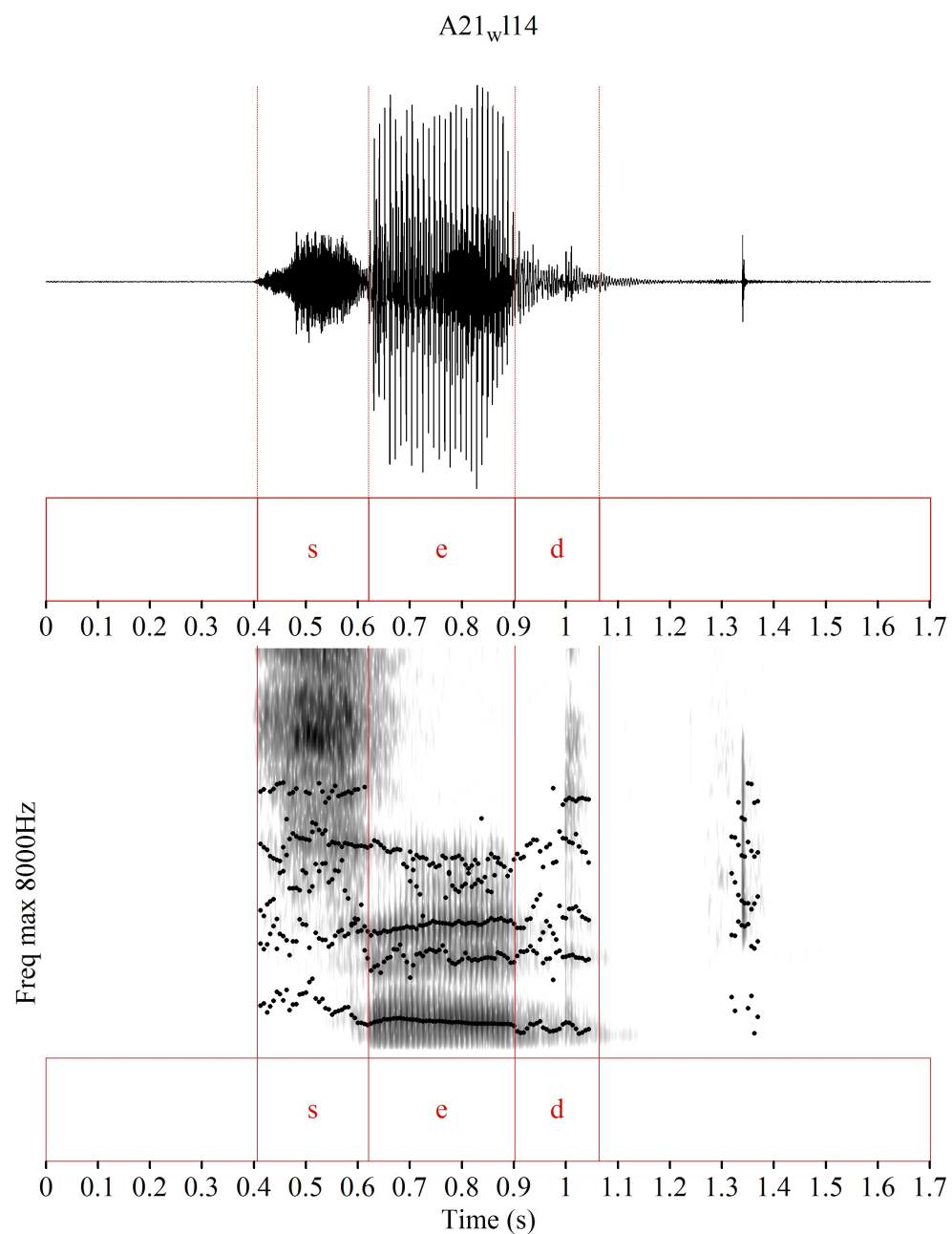


Figure 56: The waveform (above) and the spectrogram (below) of word 14 of the word list as pronounced by informant A21. The word *said* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

A21_w105

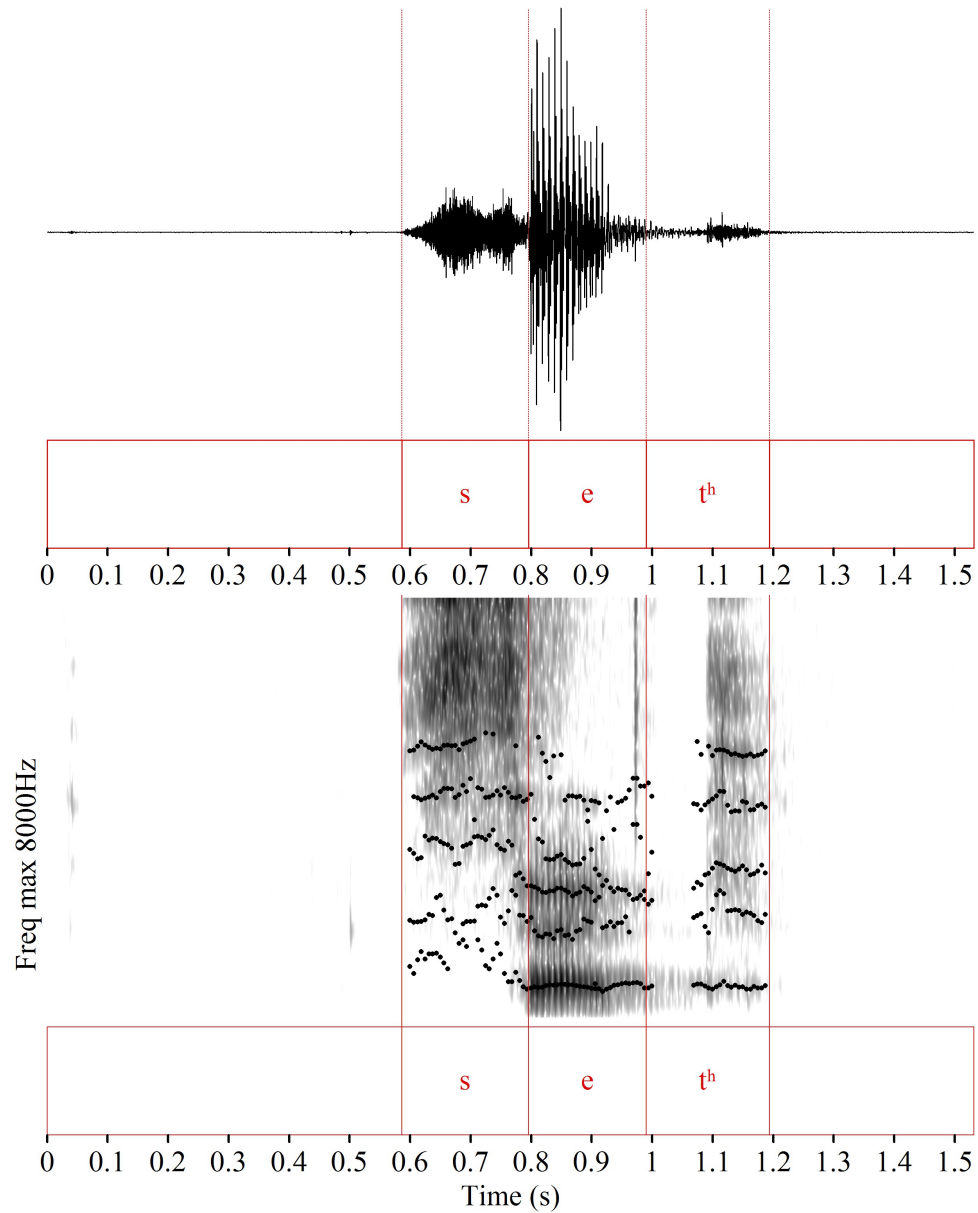


Figure 57: The waveform (above) and the spectrogram (below) of word 5 of the word list as pronounced by informant A21. The word *set* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

A16_wl13

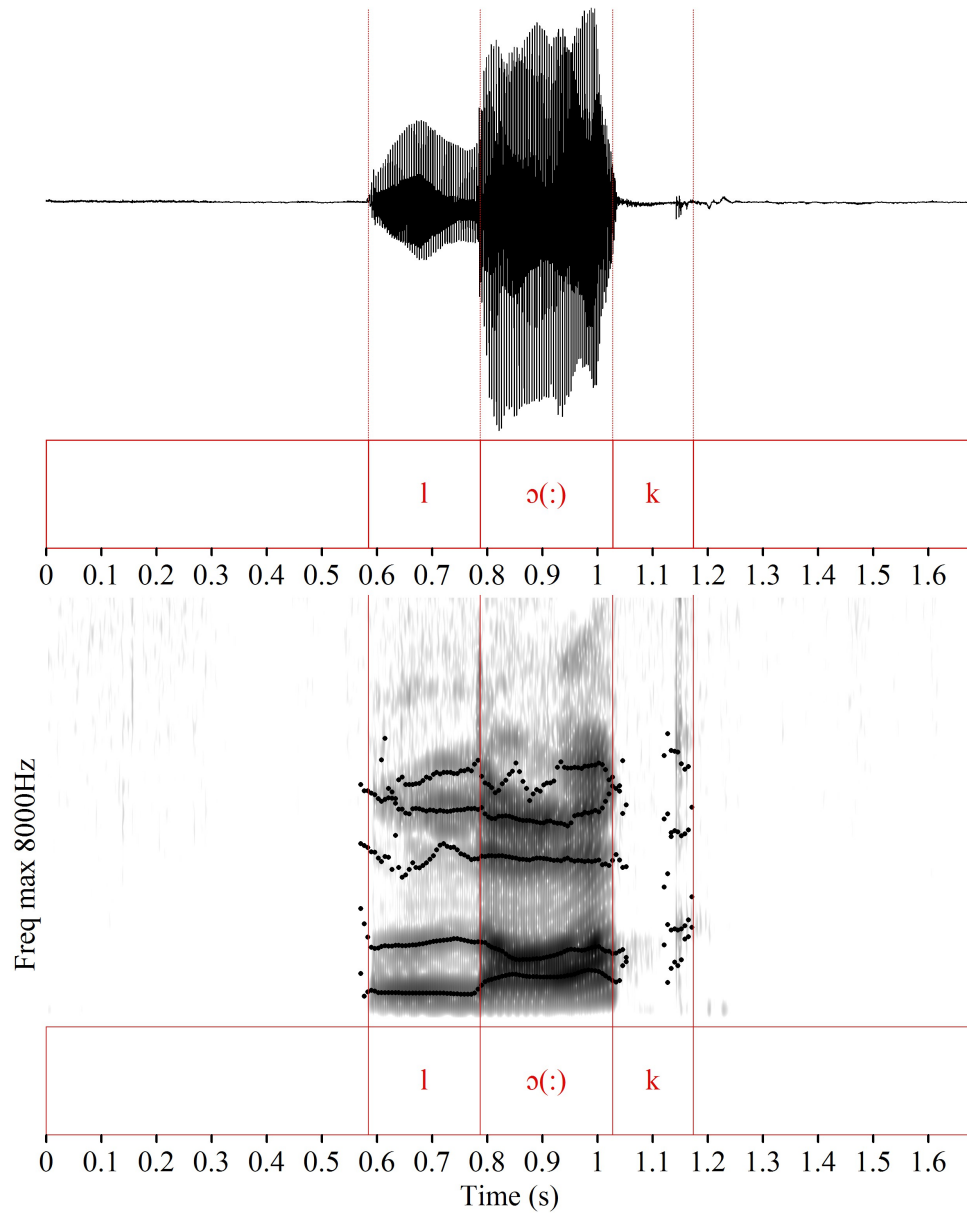


Figure 58: The waveform (above) and the spectrogram (below) of word 13 of the word list as pronounced by informant A16. The word *log* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

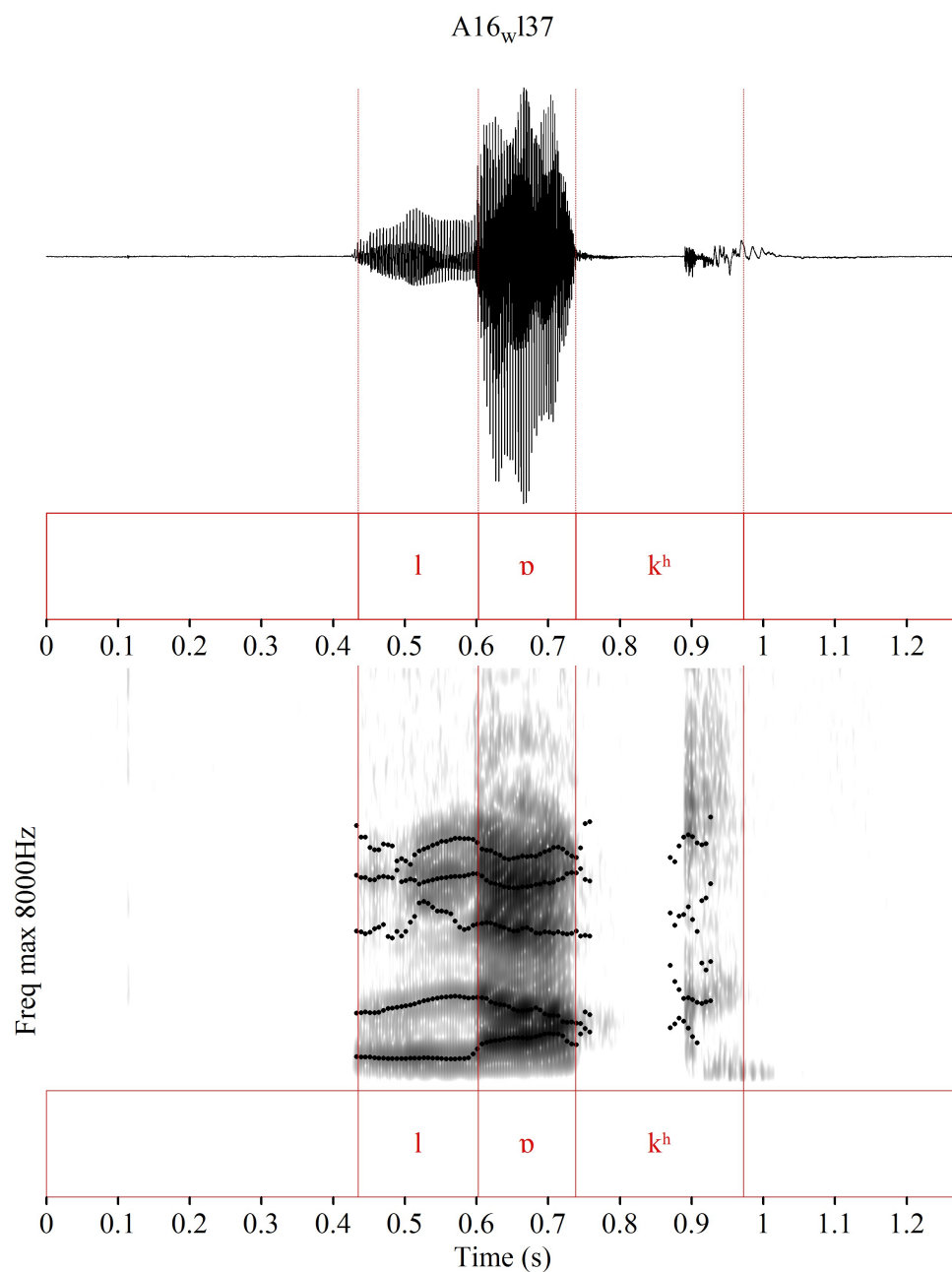


Figure 59: The waveform (above) and the spectrogram (below) of word 37 of the word list as pronounced by informant A16. The word *lock* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

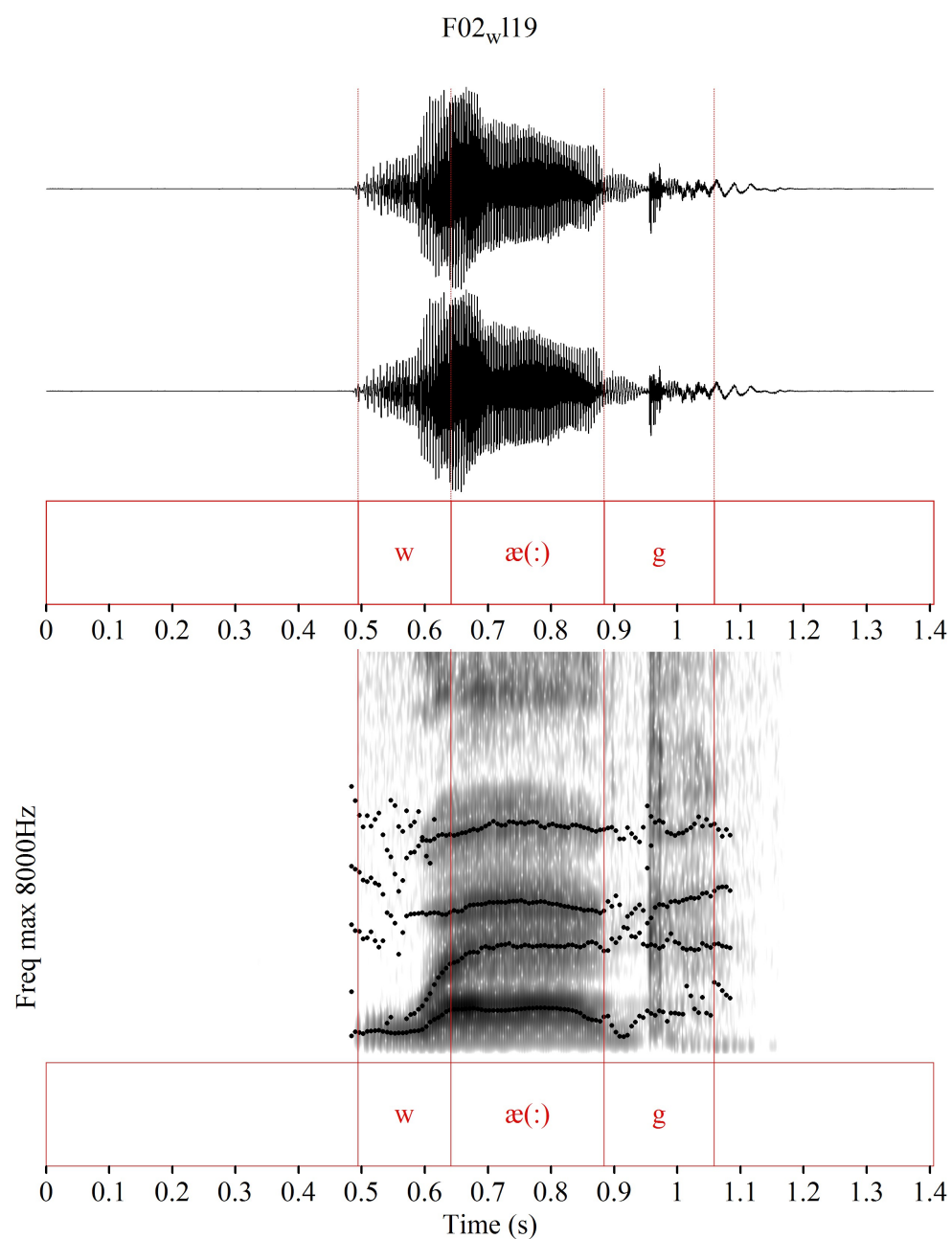


Figure 60: The waveform (above) and the spectrogram (below) of word 19 of the word list as pronounced by informant F02. The word *wag* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

F02_w139

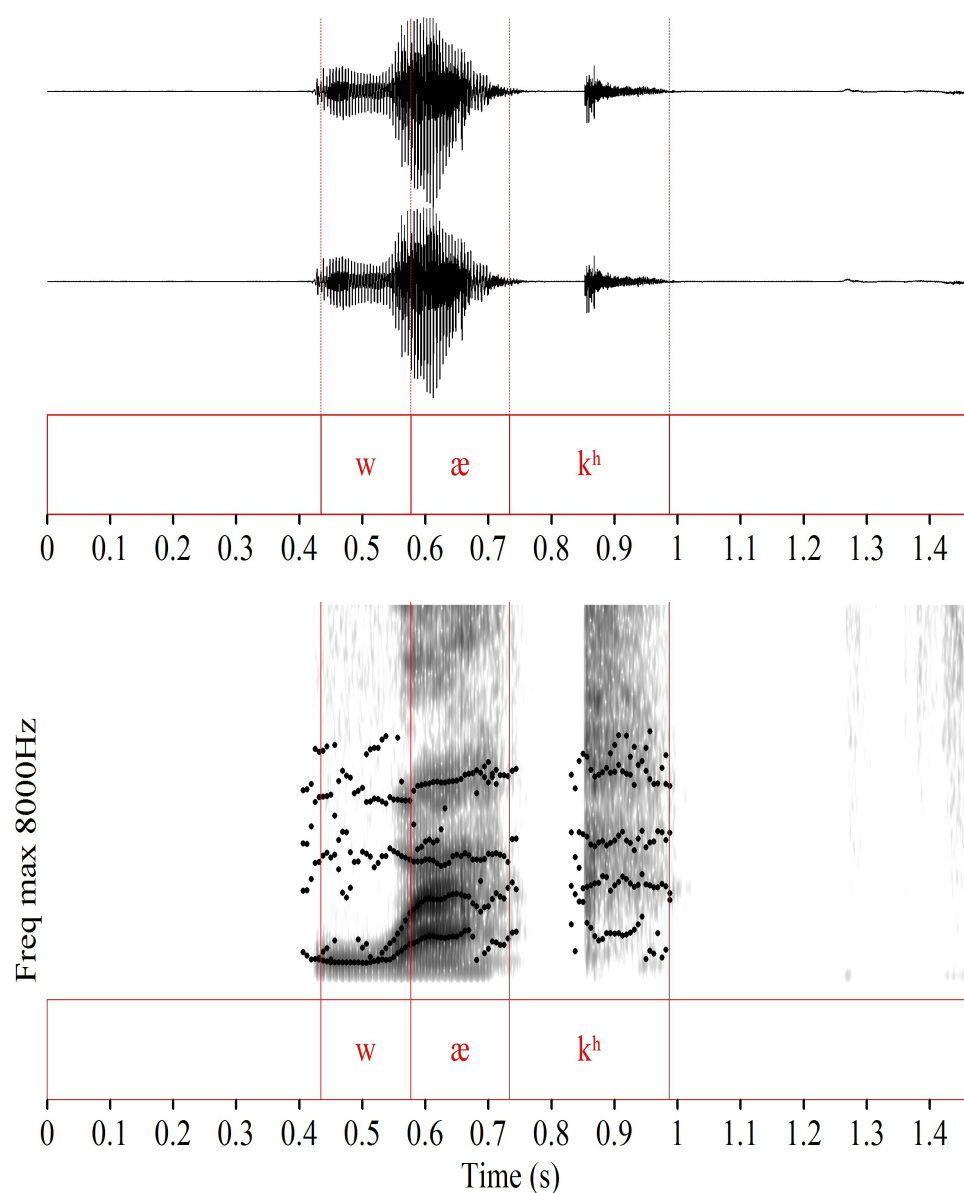


Figure 61: The waveform (above) and the spectrogram (below) of word 39 of the word list as pronounced by informant F02. The word *whack* is marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

A08_t05

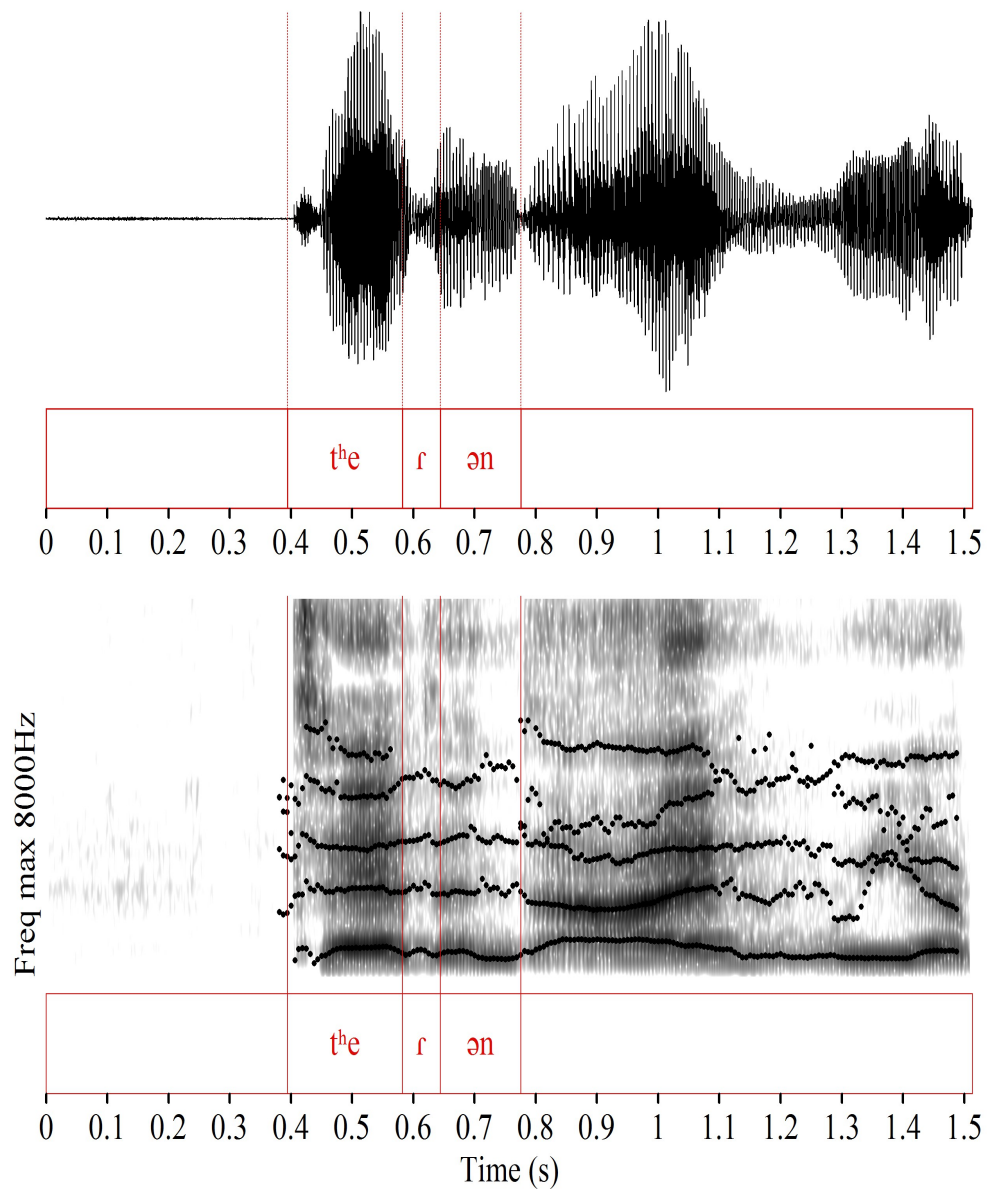


Figure 62: The waveform (above) and the spectrogram (below) of string 5 of the text passage as pronounced by informant A08. The words *Ted* and *and* are marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

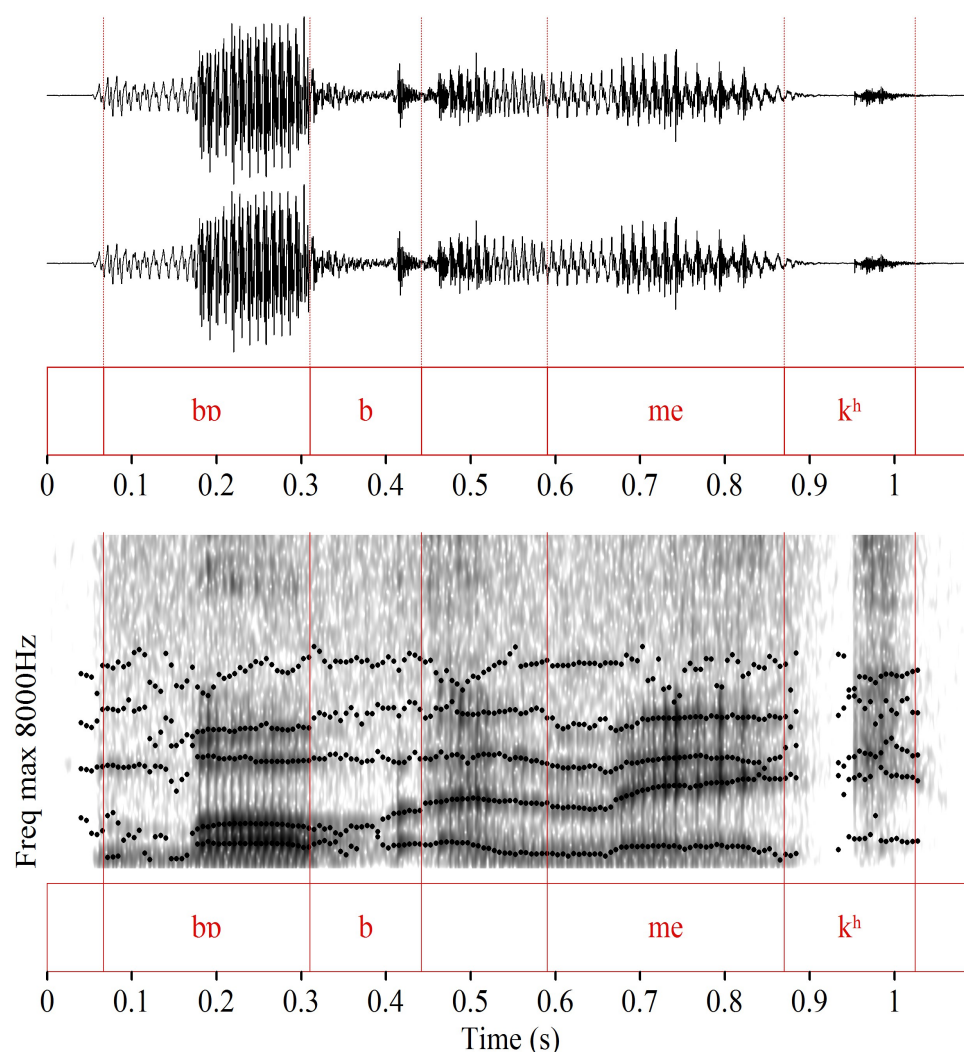


Figure 63: The waveform (above) and the spectrogram (below) of string 15 of the text passage pronounced by informant F03. The words *Bob* and *Meg* are marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

4.4 Voiceless alveolar plosive /t/ in word-medial prevocalic position

A flap is produced by quickly tapping the alveolar ridge with the tongue. The average duration of a flap is clearly shorter than the duration of the aspirated [t^h]. Due to its short duration, there is no visible closure (i.e. silence) and the release burst is often very weak or non-existent. Figure 64 presents four different ways of pronouncing the voiceless alveolar plosive /t/ in word-medial prevocalic position. Both the lack of closure before the flap sound and its short duration are visible in the spectrogram of [sɪrɪ] (second from the left).

Like described in section 2.2.4, the non-stressed word-medial prevocalic /t/ is often pronounced as a flap especially by speakers of American-style English. The results of the

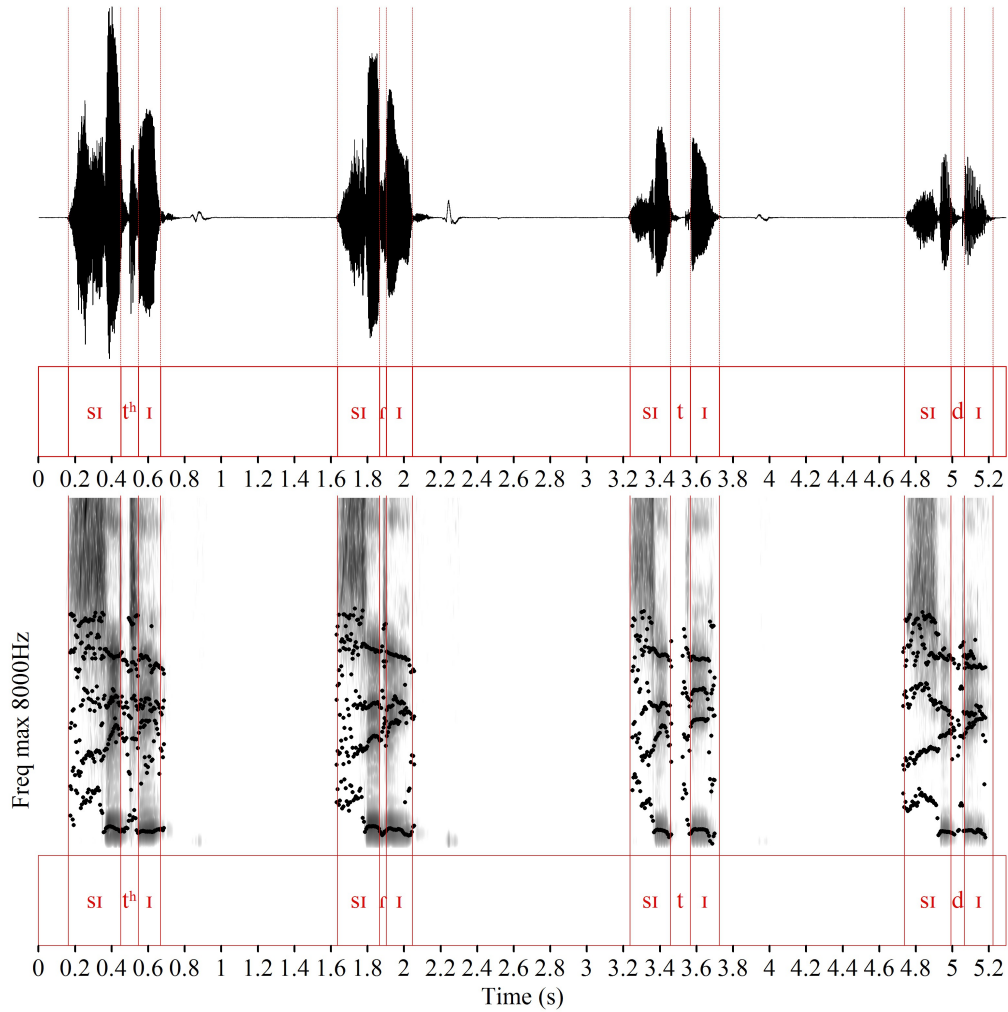


Figure 64: The waveforms (above) and the spectrograms (below) of the word *city* as pronounced by the author with four different allophones of /t/: [sɪtʰɪ], [sɪɾɪ], [sɪtɪ] and [sɪdɪ]. The red transcriptions mark the locations of each sound. Formants for vowels and other sonorants are marked with black dots.

analysis indicate that the American-style flapping is more common among the Finnish informants. As illustrated in Figure 65, a third of the occurrences of word-medial prevocalic /t/ in the word list samples of Group 1 were flapped, and almost 50% of all analysed occurrences in the text passage were pronounced with a [ɾ] by Group 1 informants. The majority of the Austrian informants pronounced the /t/ as a [tʰ] in both tasks, although the amount of flapped /t/ sounds increased noticeably in the text passage samples (15,2% → 34%). The results indicate that the pronunciation of the word-medial prevocalic /t/ depends on the informant group. This correlation is statistically significant for all of the p-values are markedly smaller than the chosen significance level α (Table 4). The category ‘other’ in Figure 65 refers to occurrences that were not flapped nor aspirated.

These allophones of word-medial prevocalic /t/ include mostly instances of [t̬] (Figure 28). The word list word *thirty* (word 8) was most frequently flapped by both groups. The spectrograms of this word in Figures 27, 28 and 29 include the three most commonly produced allophones of word-medial prevocalic /t/. In the text passage samples, Group 1 informants also flapped the word *thirty* most frequently (76%), whereas 13 Group 2 informants (62%) produced a flap in between the words *that* and *I* (Figure 66). Other examples of flapping can be seen in Figures 9, 12, 42 and 44.

As the analysis of the informant groups' pronunciation shows, the inter-medial prevocalic /t/ was flapped more frequently in the less formal task i.e. the text passage. Also the average amount of flapped sounds per informant increased as the task formality decreased. On average, a Group 1 informant pronounced 36% of the word list and 39% of the text passage occurrences with a flap. These percentages for a typical Group 2 informant are 16% and 29%. Because the distributions of flapping are positively skewed and the amount of informants who consistently flapped all or most instances of inter-medial prevocalic /t/ was very low, the mean value is not necessarily the best way to describe the individual performance score.

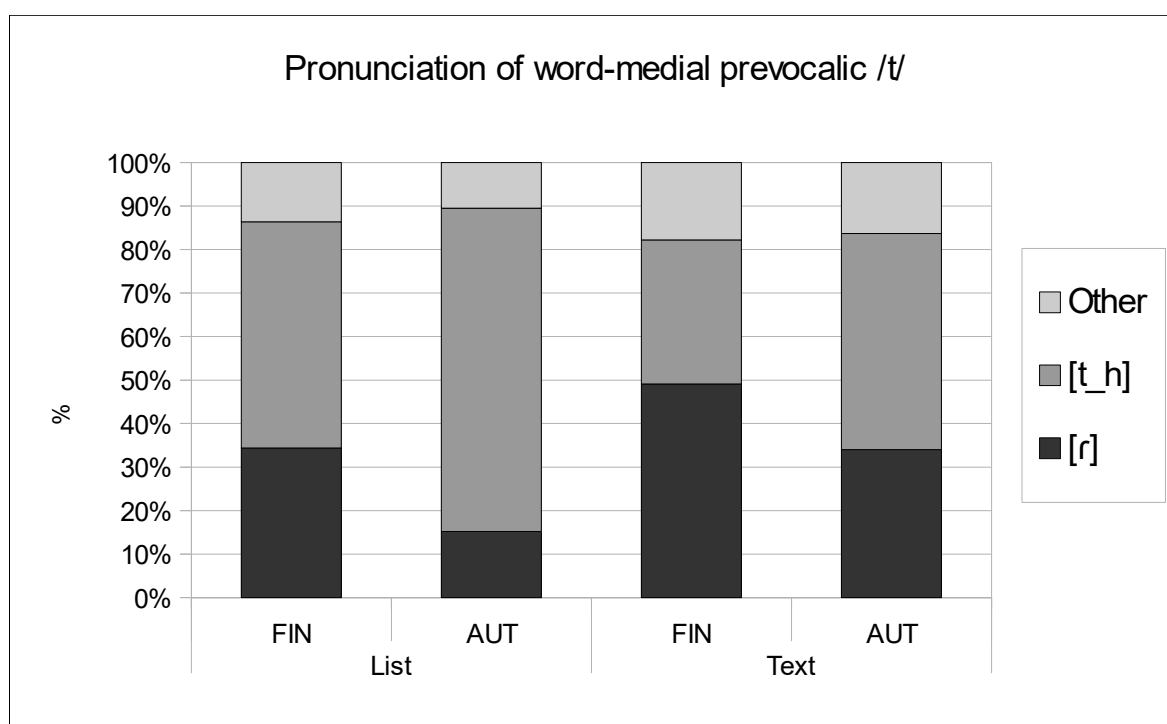


Figure 65: Different ways of pronouncing the word-medial prevocalic /t/.

Variables (Controlled x Measured)	p-values from the χ^2 -test ($\alpha=5e-02$)
Group x Pronunciation of word-medial prevocalic /t/ (LIST)	1,56e-03
Group x Pronunciation of word-medial prevocalic /t/ (TEXT)	8,68e-03
Group x Pronunciation of word-medial prevocalic /t/ (TOTAL)	5,67e-05

Table 4: p-values from the χ^2 -test for the word-medial prevocalic /t/.

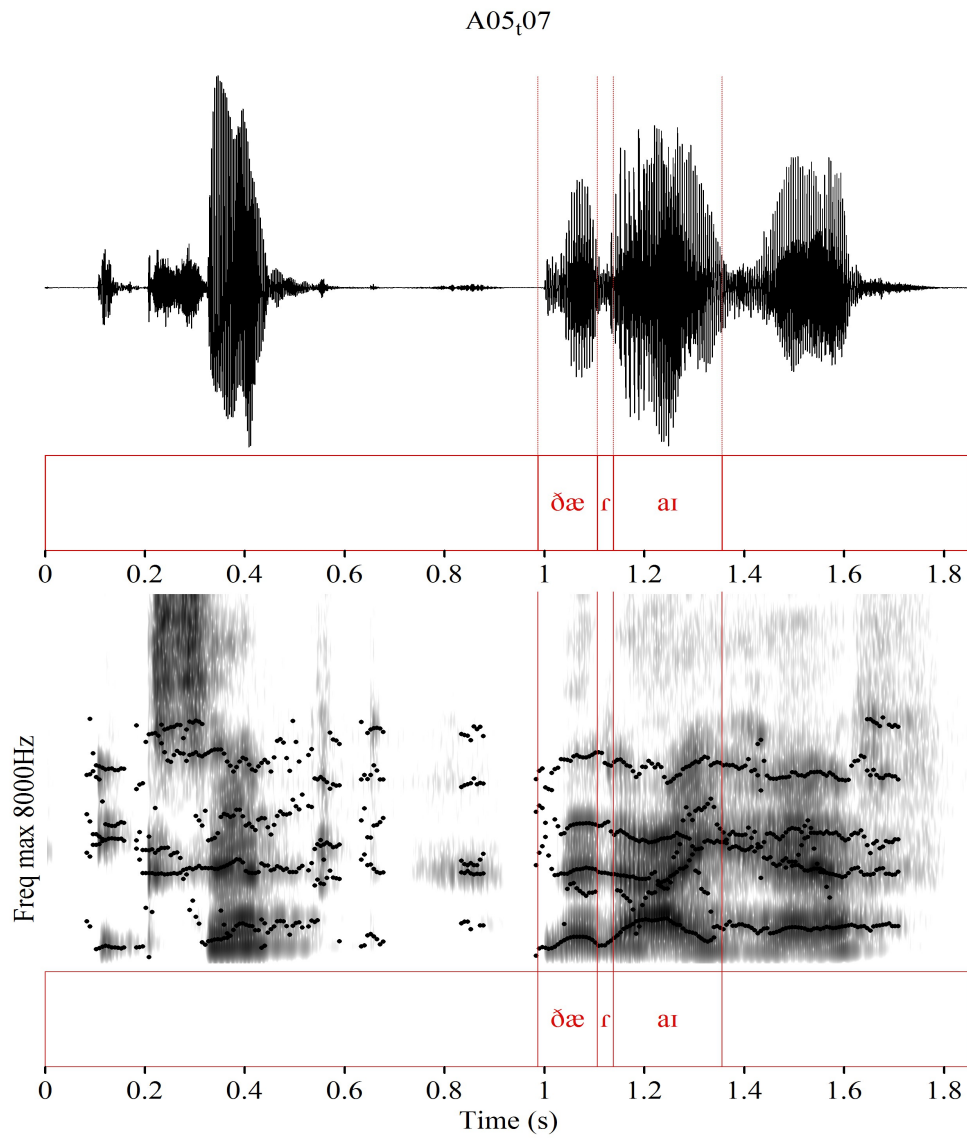


Figure 66: The waveform and the spectrogram of string 7 of the text passage as pronounced by informant A05. The words *that* and *I* are marked and transcribed in red. Formants for vowels and sonorants are marked in black dots.

5 Discussion

In section 4, the results of the pronunciation analysis were presented in detail. The Finnish and Austrian informants' pronunciation of the investigated sounds varied from informant to informant, and in the case of many of the researched sounds, a statistically significant correlation between informant group (i.e. informant's L1) and pronunciation of a particular sound could be established. The possible reasons behind the observed variation in pronunciation between individual informants, groups and pronunciation tasks are discussed in the following paragraphs.

The first step in learning the phonology of a new language is to become aware of the sounds of the particular language and to store them in one's mind as idealised articulatory targets. If these new sounds are not perceived correctly (e.g. due to problems in hearing etc.) or if the main source of input (e.g. teacher's accent) differs from the standard, the targets stored in the learners' mind might be non-standard-like from the beginning on thus resulting in a foreign accent. These learners may even be under the impression that their pronunciation of these sounds is standard-like. Another possible source of variation is the interrelationship between written and spoken language. Instead of imitating spoken language, the informants were asked to read both the word list and the text passage aloud i.e. to transform written text into sound. Some of the samples of non-standard-like pronunciation may have resulted from misguided interpretation of the written words. These types of errors were taken into account in the composition of both pronunciation tasks. In order to minimise the influence of mispronunciations, only relatively simple words and everyday language were included in both the word list and the text passage. Nevertheless, some misinterpretations of the written language could be observed: the word *log* (word 13 in the word list), for example, was pronounced as [lodʒ], i.e. like the word *lodge*, by one Group 1 informant and the word *vein* (word 30) as [vi:n] by a Group 2 informant (see Figure 21).

Spelling i.e. the written form may also have affected the pronunciation of the labiovelar approximant. As mentioned earlier, the letter 'w' is commonly used in the German language, whereas in Finnish it only appears in relatively few loan words. In both languages, /w/ is pronounced labiodentally, though unlike in German, some of the Finnish 'w'-words may also be pronounced with a [w]. These differences between German and English in the sound corresponding with the letter 'w' may have influenced the Austrian informants' pronunciation of the investigated sound. Additionally, the orthography of some

common English ‘w’-words resembles or is identical with some German words. This may additionally influence how speakers of (Austrian) German pronounce the /w/ in such words as *will* and *western* (*ich will* ‘I want’ and *West* ‘west’). The effects most likely resulting from similarities in spelling between L1 and L2 could also be observed in the analysis of the pronunciation of the word-initial voiced labiodental fricative. Of all the analysed cases of possible hypercorrection, the /v/ in the word *verb* (word 25 in the word list) was least frequently pronounced as a [w]. The possible explanation behind this result may be the similarities in spelling (and meaning) between the English *verb* and the German *Verb* ([vɛʁp] or [vɛʁp]) and the Finnish *verbi* ([vɛrbi]). In addition to the similarities in spelling and differences in pronouncing the letter ‘w’, the pronunciation of the labiovelar approximant was potentially influenced by the process of assimilation. This phonological process describes how a sound is influenced by the surrounding sounds and often shares qualities with another sound in its immediate context. Even though no connection between roundedness and standard-like pronunciation of /w/ could be established in the word list data, the text passage samples showed a statistically significant correlation between the pronunciation of /w/ and the position and the roundedness of the following vowel. At least in word-initial position, the roundedness of the following vowel could have affected the pronunciation of the preceding /w/: the rounding of the lips in preparation for the rounded vowel could have begun already during the /w/ making it easier to produce a [w] with rounded lips than a labiodental sound.

Similar to the labiovelar approximant, the English (inter-)dental fricatives are not included in either the Finnish or the (Austrian) German sound systems. The letter combination ‘th’ does occur in both languages but is always pronounced either with an aspirated voiceless alveolar stop [t^h] (German) or with a non-aspirated dentalalveolar plosive [t̪] (Finnish). According to natural phonology, new sounds are, at least in the beginning of the learning process, replaced with “easier sounds” (Nathan 1982:123). The Finnish informants seemed to be influenced by the spelling, for the most frequent non-standard-like allophone for the (inter-)dental fricatives was, in addition to the dentalalveolar [t̪], the aspirated voiceless alveolar [t^h]. The Finns replaced the fricative sounds with stops, whereas the Austrian substitutions for the researched sounds were most commonly other fricatives such as the labiodental [f] and the alveolar [s]. In the text passage samples, the Austrian informants frequently pronounced the voiced (inter-)dental fricative also as a non-aspirated alveolar plosive. This could at least partially be explained by elision, for in most cases the ‘th’-sounds are preceded or followed by an alveolar plosive. In strings 7, 9, 12

and 20, for example, the (inter-)dental fricatives were deleted (e.g. in Figures 40, 41, 11 and 39). As mentioned earlier, also the native speakers pronounce the voiced (inter-)dental fricative frequently as a weak plosive, so replacing it with a short dentalveolar /t/ in non-stressed positions, as many of the informants did in the text passage task, cannot be considered as a feature of a strong foreign accent. Since the word-final voiced (inter-)dental fricative is also produced voicelessly by many native speakers of English, pronouncing the /ð/ in *smooth* and *with* as a [θ] is only a deviation from the chosen dictionary standard, not from the actual realisations produced by native speakers. Additionally, these and many other words with /ð/ in the word-final position rarely have minimal pairs with a [θ] within the same word class (noun, verb etc.), so devoicing of these sounds seldom leads to serious misunderstandings.

Unlike the pronunciation of the word-final voiced (inter-)dental fricative without voicing, devoicing the English word-final voiced plosives may often lead to misunderstandings for many words with a word-final voiced plosive have a voiceless minimal pair from the same word class (like 3 of the 6 words in the word list). Especially the Austrian informants devoiced the researched word-final sounds relatively frequently so that there was very little difference between the minimal pairs (see Figures 54 and 55). The differences between the two informant groups could partially be explained by the differences in the first languages. In German, the word-final devoicing is an active phonological process. At least in the beginning of the learning process, German-speaking learners of English often identify the word-final voiced plosives with the phonemes in their L1 and apply the process of devoicing making the English sounds resemble the aspirated voiceless plosives of their L1. In Finnish, voiced plosives (or any voiced obstruents) do not occur in the word-final position. A similar situation was investigated by Donegan and Stampe as they analysed the pronunciation of word-final obstruents by Vietnamese learners of English (1979). They found that devoicing the word-final voiced obstruents was a common phenomenon among the informants even though the L1 (Vietnamese) has no word-final voiced obstruents of any kind. This was explained by the principles of natural phonology: having never been confronted with word-final voiced obstruents, the Vietnamese learners had had no reason to suppress the process of devoicing and thus applied it to the English sounds (1979:133). In the case of the Finnish informants, the word-final voiced plosives were generally pronounced with voicing although they also had never had to suppress the process of devoicing. This could perhaps be explained by the phonemic orthography of the Finnish language: speakers of Finnish are often used to “the

situation in which phonemic distinctions are rather systematically indicated in the orthography” (Suomi, Toivanen & Ylitalo 2009:37). According to Suomi, Toivanen & Ylitalo, Finns with the knowledge of such foreign languages in which the “foreign” consonants /b/ and /g/ occur (e.g. in English) and with a high level of education often include these sounds in their inventory of Finnish consonant sounds and thus pronounce the Finnish words such as *banaani* ‘banana’ and *geeni* ‘gene’ as [bana:ni] and [ge:ni] (2009:36). It can be thus assumed that most of the Finnish informants, being highly educated and having knowledge of at least two languages in which the sounds occur (English and Swedish), have these voiced plosive sounds in their consonant inventory. Since these sounds always correspond to the graphemes ‘b’, ‘d’ and ‘g’, it is also likely that the English words spelled with these letters are pronounced with [b], [d] and [g] i.e. with voicing. This theory is supported by the findings of the word list sample analysis. By some Group 1 informants, the only difference between minimal pairs such as *cab* and *cap* was the voicing of the final plosive. In such samples, such as in Figures 67 and 68, no significant differences in the vowel length could be observed.

In the text passage samples, the word-final /d/ was often pronounced as a flap when followed by a word beginning with a vowel, like in strings 5, 12 and 30. An example of this type of pronunciation was presented earlier in Figure 62. As explained in section 4.3, the increased voicing of the word-final plosives in strings 15 and 23 of the text passage, especially by Group 2 informants, could have been induced by the following vowel sounds. In fast non-stressed speech, the words with the voiced plosives were often merged together with the following words. In such cases, the plosive occurred actually in word-medial, intervocalic instead of word-final positions. Since the German process of devoicing only applies to word-final sounds, these voiced plosives in their actual word-medial position were less frequently devoiced.

According to Nathan, the flapping of the word-medial prevocalic voiceless (and voiced) alveolar plosives results from the process of lenition: flapping is faster and easier to produce than the normal stops (1982:120). As mentioned in section 2.2.4, flapping is more common in GA than in RP. The prevalence of flapping among Finnish informants could result from the general exposure to spoken American-style English through different media. Additionally, the differences between Austrian German and Finnish plosives (mainly the absence of aspiration in Finnish) could have influenced the pronunciation of /t/ in word-medial prevocalic positions making the Austrians pronounce the sound more frequently with aspiration.

In addition to all of these possible factors influencing the pronunciation of the researched sounds, the Finnish informants' much greater exposure to spoken English from an early age on could have positively influenced their acquisition of the phonetic structures of English. As explained in section 2.1, both the CPH and the mental-category-based learning by Bialystok (1997) suggest that exposure at an early age increases the learner's chances to attain a higher-level proficiency. Thus, the differences between the groups could at least partially be explained by the different amounts of English input.

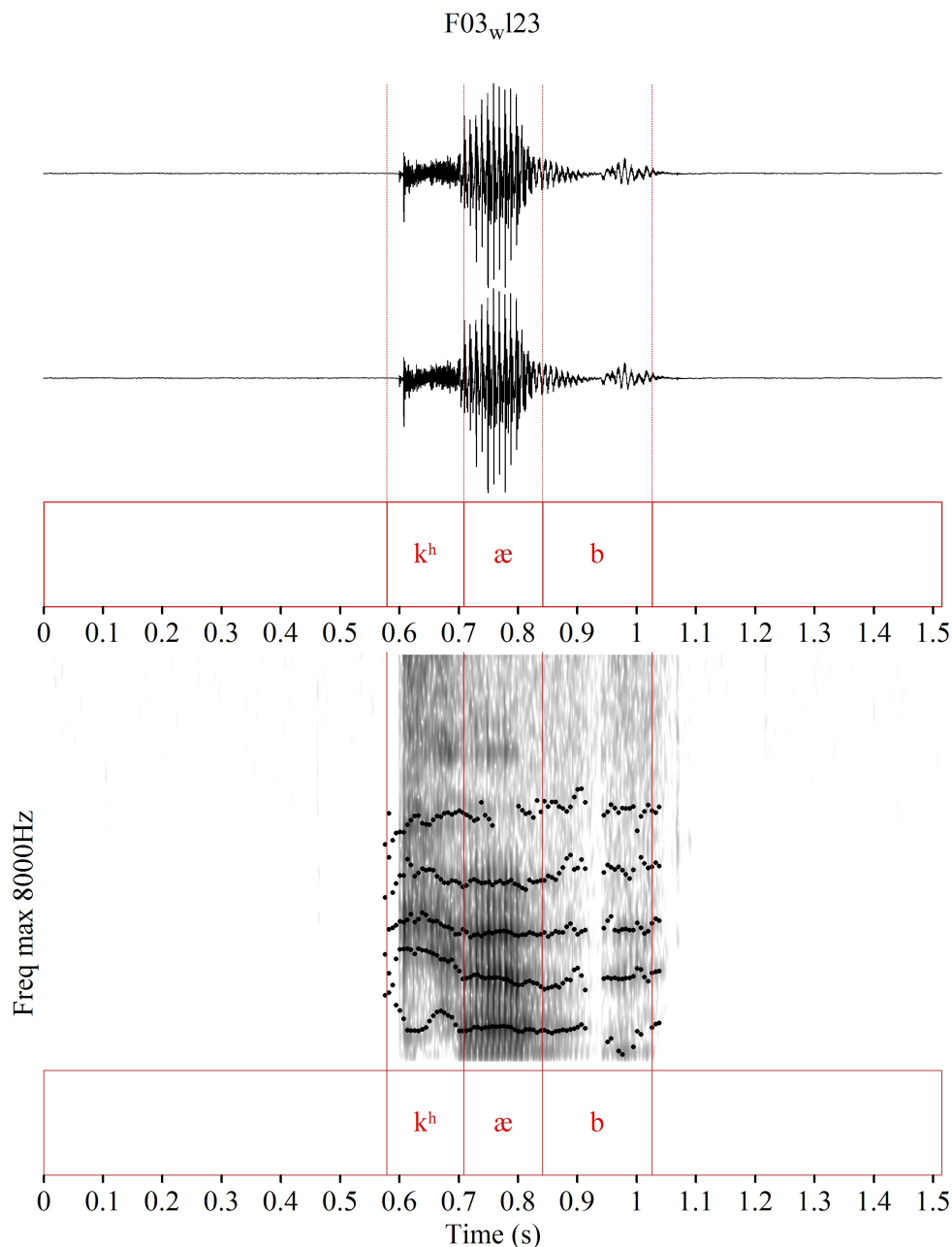


Figure 67: The waveform and the spectrogram of the word 23 of the word list as pronounced by informant F03. The word *cab* is marked and transcribed in red. Formants for vowels and sonorants are marked with black dots.

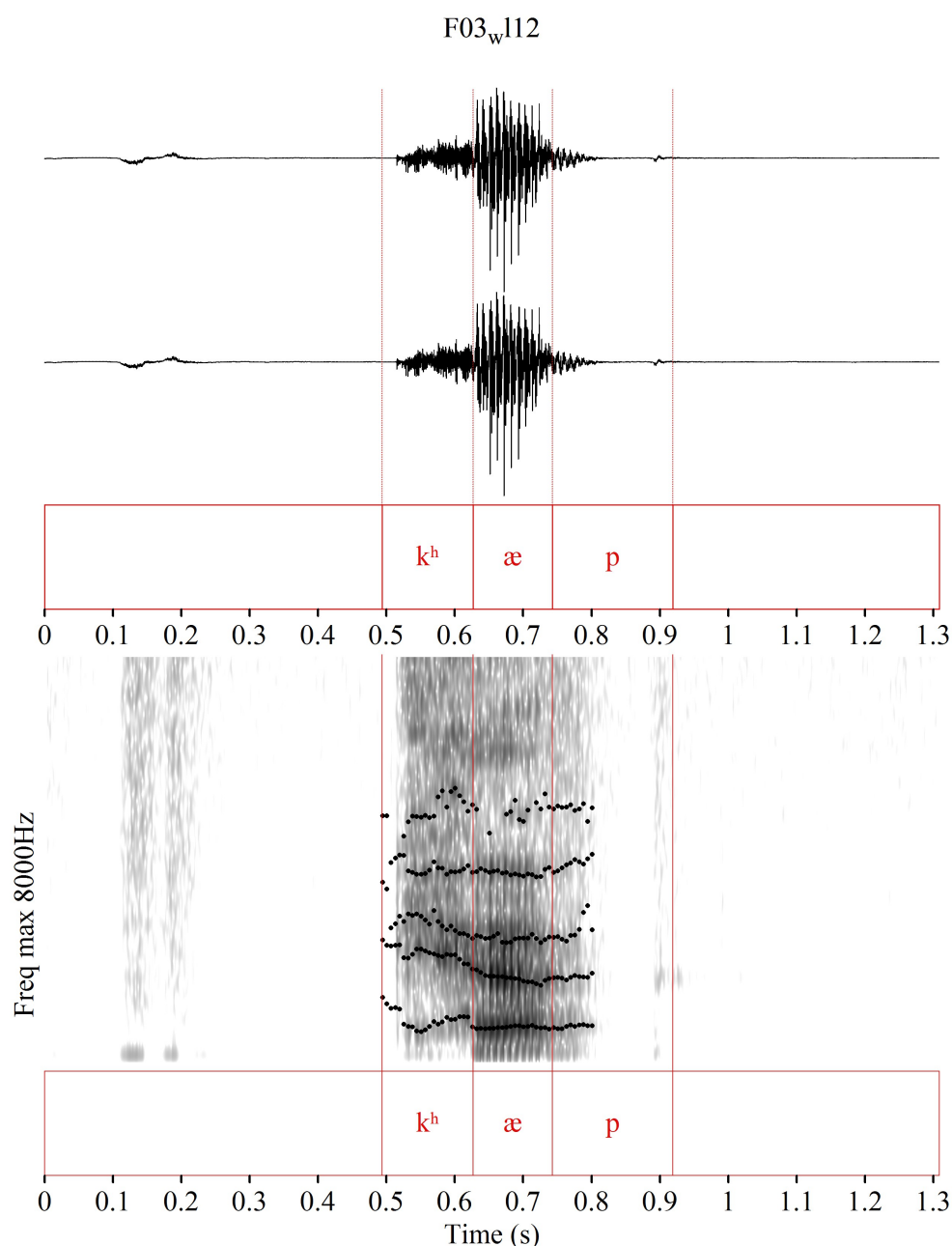


Figure 68: The waveform and the spectrogram of the word 23 of the word list as pronounced by informant F03. The word *cap* is marked and transcribed in red. Formants for vowels and sonorants are marked with black dots.

6 Conclusions

In this paper, the pronunciation of certain English sounds by both native speakers of Finnish and Austrian German was analysed to ascertain how prevalent the stereotypical L1-style pronunciations are among young Finnish and Austrian adults and to identify any possible differences between the two informant groups. Six of the analysed sounds do not occur in either of the first languages. Both the voiced labiovelar approximant and the

(inter-)dental fricatives can be considered as “new sounds” for they do not occur in any of the two L1s. The three voiced plosives /b/, /d/ and /g/ do not occur in word-final position in Finnish and are always devoiced in Austrian German. The seventh sound, the word-medial prevocalic /t/, was included to assess the informants’ tendency towards American-style pronunciation. The general hypothesis was that the Finnish informants’ pronunciation, due to greater exposure to spoken (American) English through various media, would be more standard-like and include more instances of American-style flapping.

The results of this study are based on the audio data from two pronunciation tasks: a word list consisting of 39 words and a short text passage. The data consist of 25 Finnish and 21 Austrian pronunciation samples. All informants had to fulfil specific criteria related to age and education: suitable informants were born between 1979 and 1987 and had at least a bachelor’s degree from a regular university or from a university of applied sciences. These criteria were chosen to minimise the differences between the two groups to two main factors: informants’ L1 and exposure to spoken English. The analysis was mainly conducted by examining the visualisations of the audio data, though a more subjective method based on the hearing of the author was applied in order to differentiate between voiceless labiodental and (inter-)dental fricatives.

The pronunciation of the researched sounds varied from informant to informant and some sounds were more frequently pronounced in a stereotypical way than others. For example, [s], the stereotypically Austrian way to pronounce a /θ/, was relatively rare, whereas the devoicing of word-final voiced plosives was a more commonly applied process among Austrian informants. The total results indicate a statistically significant correlation between the informant group and the pronunciation i.e. the standard-like pronunciation of the researched sounds seems to depend on the L1 of the informant (Table 5). These results seem to support the original hypothesis of Finnish informants’ having a more standard-like pronunciation. The flapping also occurred more frequently in the samples of the Finnish informants. As discussed in the previous section, various factors could have caused these differences between the groups ranging from problems in perception and false articulatory targets to misguided interpretation of the pronunciation of written words. In addition, the phonological processes or their suppression may be stronger in one L1 than in the other i.e. learning to pronounce certain English sounds in a standard-like manner may be more difficult for German than Finnish speakers or vice versa. The findings of this study could also indicate that the greater amount of input through exposure to spoken English in non-institutional settings (mainly films and TV) from an early age on

Variables (Controlled x Measured)	p-values from the χ^2-test ($\alpha=5e-02$)
Group x Pronunciation of the /w/ (TOTAL)	3,8e-4
Group x Pronunciation of /θ/ (TOTAL)	8,89e-3
Group x Pronunciation of /ð/ (TOTAL)	1,89e-4
Group x Pronunciation word-final voiced plosives (lenis) (TOTAL)	1,06e-6
Group x Pronunciation of word-medial prevocalic /t/ (TOTAL)	5,67e-5

Table 5: p-values from the χ^2 -test for the occurrences of all researched sounds in both pronunciation tasks.

leads to more standard-like pronunciation of the researched sounds. It is possible that the Finnish informants, thanks to the significantly greater input of English at an early age, created a new mental category for the English phonemes instead of extending the boundaries of the already existing phoneme categories of the L1. Or, in terms of natural phonology, the Group 1 informants have learned to suppress the right phonological processes and to apply others. In comparison, the results suggest that many Austrian informants have either never learned not to apply the L1 processes, or they have included the new English sounds into the existing categories of the first language by expanding the boundaries instead of creating new mental categories. How much these observed differences result from factors related to individual informants such as perception and knowledge about how words are pronounced cannot be estimated, though the relatively high group size (>20) should reduce their effect on the total results. The degree of influence resulting from the different levels of constraint induced by the speakers L1 could only be assessed by comparing the samples from two large groups of informants which differ from each other solely by the informants' L1. Since it is possible that the differences between the two informant groups of this study result, at least partially, from the varying constraints the two L1s might impose on the learning process of L2 sounds, the exact degree of the effects of exposure cannot be determined from the results of the pronunciation analysis. Yet in my opinion, the strength of the correlation between group and pronunciation (as indicated by the small p-values in Table 5) cannot be explained only through the effects and constraints of different L1s. The results of the analysis indicate that the greater amount of input, i.e. exposure to spoken English through media, must have positively influenced the Finnish informants' acquisition of English sounds. Thus the results of this thesis can be seen to support the conclusion of the study by Safar et al. stating that "subtitling helps to improve the mastery of foreign languages" (2011:26).

7 References

- 123apps LCC. 2016. "Voice Recorder online". [Online software] <http://online-voice-recorder.com> (24 Feb. 2016)
- Ammon, Ulrich; Bickel, Hans; Ebner, Jakob; Esterhammer, Ruth; Gasser, Markus; Hofer, Lorenz; Kellermeier-Rehbein, Birte; Löffler, Heinrich; Mangott, Doris; Moser, Hans; Schläpfer, Robert; Schloßmacher, Michael; Schmidlin, Regula; Vallaster, Günter (eds.). 2004. *Variantenwörterbuch des Deutschen. Die Standardsprache in Österreich, der Schweiz und Deutschland sowie in Liechtenstein, Luxemburg, Ostbelgien und Südtirol*. Berlin: Walter de Gruyter.
- Ashby, Michael; Maidment, John. 2005. *Introducing phonetic science*. Cambridge: Cambridge University Press.
- Bialystok, Ellen. 1997. "The structure of age: in search of barriers to second language acquisition". *Second language research* 13(2), 116-137.
- Biersack, Sonja. 2002. "Systematische Aussprachefehler deutscher Muttersprachler im Englischen – Eine phonetisch-phonologische Bestandsaufnahme". *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München* 39, 37-130.
- Boersma, Paul; Weenik, David. 2015. *Praat: doing phonetics by computer*. (Version 5.4.22). [Computer programme] <http://www.praat.org/> (27 Oct. 2015).
- Bongaerts, Theo; Planken, Brigitte; Schils, Erik. 1995. "Can late starters attain a native accent in a foreign language? A test of the critical period hypothesis" In: Singleton, D. & Lengyel, Z. (eds.). *The age factor in second language acquisition*. Avon: Multilingual Matters Ltd, 30-50.
- Dalton-Puffer, Christiane; Kaltenböck, Gunther; Smit, Ute. 1997. "Learner attitudes and L2 pronunciation in Austria". *World Englishes* 16(1), 115-128.
- de Cillia, Rudolf; Krumm, Hans-Jürgen. 2010. "Fremdsprachenunterricht in Österreich". *Sociolinguistica* 24, 153-169.
- Donegan, Patricia J.; Stampe, David. 1979. "The study of natural phonology". In: Dinnsen, David A. *Current approaches to phonological theory*. Bloomington: Indiana University Press, 126-173.
- Ehrlich, Karoline. 2009. "Die Aussprache des österreichischen Standarddeutsch – umfassende Sprech- und Sprachstandserhebung der österreichischen Orthoepie". PhD thesis, Department of German Philology, University of Vienna.

- Flege, James E.; Munro, Murray J.; McKay, Ian R. A.. 1995. "Factors affecting the strength of perceived foreign accent in a second language." *Journal of the Acoustical Society of America* 97(5), 3125-3134.
- Flege, James E.; Yeni-Komshian, Grace H.; Liu, Serena. 1999. "Age constraints on second language acquisition." *Journal of Memory and Language* 41, 78-104.
- Hebenstreit, Andrea. 1997. "Language attitudes of Austrian pupils towards British and American English". Diploma Thesis, University of Vienna.
- Ioup, Georgette; Boustagui, Elizabeth; El Tigi, Manal; Moselle, Martha. 1994. "Reexamining the critical period hypothesis. A case study of successful adult SLA in a Naturalistic Environment". *Studies in Second Language Acquisition* 16, 73-98.
- Johnson, Keith. 2003. *Acoustic and auditory phonetics*. (2nd edition). Oxford: Blackwell Publishing Ltd.
- Jongman, Allard; Wayland, Ratre; Wong, Serena. 2000. "Acoustic characteristics of English fricatives." *Journal of the acoustical society of America* 108(3), 1252-1263.
- Kresta, Ronald. 2015. "Aussprachefehler von deutschen und deutschsprachigen Studierenden technischer Studiengänge in englisch-sprachigen Fachvorträgen – Eine empirische Untersuchung". In Busch-Lauer, Ines-Andrea (ed.). *Facetten der Fachsprachenvermittlung Englisch – Hands on ESP teaching*. Berlin: Frank & Timme GmbH, 113-136.
- Korpela, Jukka K. 2015. *Handbook of Finnish*. (Kindle Version). Retrieved from Amazon.de.
- Könnölä, Jouko; Lahtinen, Matti; Lankinen, Timo; Lindström, Aslak; Numminen, Ulla; Kangasvieri, Anneli; Souri, Liisa; Blom, Heikki; Kinnunen, Elisabet. 2001. *Nuorisosteen koulutuskokeilun arviointiryhmän muistio 2001*.
http://www.minedu.fi/export/sites/default/OPM/Julkaisut/2001/liitteet/opm_530_nukomuistio.pdf?lang=en
- Labov, William. 1972. "Some principles of linguistic methodology." *Language in Society* 1(1), pp. 97-120.
- Labov, William. 1984. "Field methods of the project on linguistic change and variation". In: Baugh, J.; Sherzer, J. (eds.). *Language in Use: Readings in Sociolinguistics*. Englewood Cliffs: Prentice-Hall, pp. 28-53.
- Ladefoged, Peter. 1993. *A course in phonetics*. Fort-Worth: Harcourt Brace College Publ.

- Ladefoged, Peter; Maddieson, Ian. 1996. *The sounds of the world's languages*. Oxford: Blackwell Publishers Ltd.
- Ladefoged, Peter; Disner, Sandra Ferrari . 2012. *Vowels and consonants*. (3rd edition). Oxford: Wiley-Blackwell.
- Leppänen, Sirpa; Pitkänen-Huhta, Anne; Nikula, Tarja; Kytölä, Samu; Törmäkangas, Timo; Nissinen, Kari; Kääntä, Leila; Virkkula, Tiina; Laitinen, Mikko; Pahta, Päivi; Koskela, Heidi; Lähdesmäki, Salla; Jousmäki, Henna. 2009. *Kansallinen kyselytutkimus englannin kielestä Suomessa. Käyttö, merkitys ja asenteet*. Jyväskylä: Jyväskylä University Printing House.
- Lightbown, Patsy M.; Spada, Nina. 2006. *How languages are learned*. Oxford: Oxford University Press.
- Muhr, Rudolf. 2000. *Österreichisches Sprachdiplom Deutsch. Lernzielkataloge zu Basisformulierungen Lexik – Sprechhandlungen, Höflichkeitskonventionen, Diskurs und Diskursstrukturen, Deutsch als plurizentrische Sprache*. Wien: övb & hpt.
- Nathan, Geoffrey S. 1982. "Natural phonology interference in second language acquisition". In: Boyd, John R.; Haskell, John F. (eds.). *Selected papers from the tenth Illinois TESOL/BE annual convention*. Chicago: TESOL, 118-125.
- Patkowski, Mark. 1980. "The sensitive period for the acquisition of syntax in a second language." *Language Learning* 30/2, pp. 449-472.
- Plag, Ingo; Braun, Maria; Lappe, Sabine; Schramm, Mareile. 2009. *Introduction to English linguistics*. (2nd edition). Berlin: Mouton de Gruyter.
- Piske, Thorsten; McKay, Ian R. A.; Flege, James E.. 2001. "Factors affecting degree of foreign accent in an L2: a review." *Journal of Phonetics* 29, 191-215.
- Ranta, Elina. 2004. "International English – a future possibility in the Finnish EFL classroom?" M.A. Thesis, University of Tampere.
- Roach, Peter. 2009. *English phonetics and phonology. A practical course*. (4th edition). Cambridge: Cambridge University Press.
- Ruhanen, Pasi. 2009. *Kolmoskanava 1986-1993*. <http://hapo.kapsi.fi/kolmoskanava/> (20 March 2016)

- Safar, Hayssam; Modot, Alain; Angrisani, Silvia; Gambier, Yves; Eugeni, Carlo; Fontanel, Héloïse; Hamaoui, Najwa; Verstrepen, Xavier. 2011. *Study on the use of subtitling. The potential of subtitling to encourage foreign language learning and improve the mastery of foreign languages*.
http://eacea.ec.europa.eu/llp/studies/documents/study_on_the_use_of_subtitling/rapport_final-en.pdf (20 March 2016)
- Scovel, T. 1988. *A time to speak: A psycholinguistic inquiry into the critical period for human speech*. Rowley, MA: Newbury House.
- Stampe, David. 1979. "Dissertation on natural phonology". PhD Thesis, Ohio State University
- Suomi, Kari; Toivanen, Juhani; Ylitalo, Riikka. 2008. *Finnish sound structure*. Oulu: Oulu University Press.
- Suomi, Kari. 2009. "Durational elasticity for accentual purposes in Northern Finland". *Journal of phonetics* 37(4), 397-416.
- Tilastokeskus. 2005. *Televisiokanavien ohjelmisto alkuperämaittain 1997 – 2003*.
<http://tilastokeskus.fi/til/jvie/tau.html> (29 May 2013)
- Turnbull, Joanna; Lea, Diana; Parkinson, Dilys; Phillips, Patrick; Francis, Ben; Webb, Suzanne; Bull, Victoria; Ashby, Michael (eds.). 2011. *Oxford advanced learner's dictionary of current English*. (8th edition). Oxford: Oxford University Press.
- Tyndall, Marina. 2013. *Finnish accent worksheet. Scandinavian accents panel VASTA 2013*. <http://www.yorku.ca/earmstro/scandinavia/Finnish/Finnish%20Accent%20Worksheet.pdf> (22 March 2016)
- Wood, Sidney. 2005. "Beginner's guide to Praat. What are formants?".
<http://person2.sol.lu.se/SidneyWood/praaate/whatform.html> (22 April 2016)
- Yule, George. 2010. *The Study of Language*. (4th edition). Cambridge: Cambridge University Press.

8 Appendices

8.1 Appendix A: The word list

Please, read the following words aloud:

- | | |
|--------------|---------------|
| 1. Worth | 21. Rib |
| 2. Rider | 22. Forward |
| 3. Ipad | 23. Cab |
| 4. Waiter | 24. Unwilling |
| 5. Set | 25. Verb |
| 6. Healthy | 26. Fake |
| 7. Award | 27. Fur |
| 8. Thirty | 28. Them |
| 9. Smooth | 29. With |
| 10. Tinkle | 30. Vein |
| 11. Anything | 31. Writer |
| 12. Cap | 32. Thin |
| 13. Log | 33. Otherwise |
| 14. Said | 34. Finland |
| 15. The | 35. Rip |
| 16. Water | 36. Soothing |
| 17. Pat | 37. Lock |
| 18. Moth | 38. City |
| 19. Wag | 39. Whack |
| 20. Sinful | |

8.2 Appendix B: The text passage

Please, read the following text aloud:

Greetings from South-East Asia!

I'm writing you from an award-winning hotel in Taiwan. Ted and I, we're both doing fine, except that I have the most painful sunburn on my back. I think I would have already died without the soothing after-sun lotion. We've had a great time so far, surfing and swimming in the crystal clear waters. I've promised to try wake-boarding with Ted as soon as my skin has healed.

Our next stop will be Perth in western Australia. There we'll meet our friends, Bob and Meg. We are going to rent a van and drive along the southern coastline all the way to Sydney. I would very much like to see some kangaroos and other wild animals. I have also heard that whales can often be spotted near the coast. After the road trip, we intend to enjoy the city life in Sydney for a couple of days. We'll visit some of the famous sights and celebrate Ted's birthday. He will be 30 on the 4th of March. Meg and I have planned a surprise party he'll never forget...

Looking forward to seeing you again in a couple of weeks!

Love,

Beth

8.3 Appendix C: The analysed strings from the text passage

This list consists of the analysed strings from the text passage. The individual sounds are written in italics and highlighted in light grey.

1. from South-East Asia!
2. I'm writing you
3. an award-winning hotel
4. hotel in Taiwan.
5. Ted and I,
6. we're both doing fine
7. except that I have
8. I think I would have
9. already died without the soothing
after-sun
10. surfing and swimming in the
11. the crystal clear waters
12. wake-boarding with Ted as soon
13. next stop will be
14. There we'll meet our friends,
15. friends, Bob and Meg.
16. I would very much
17. kangaroos and other wild animals
18. enjoy the city life
19. We'll visit some of the famous
20. celebrate Ted's birthday.
21. He will be 30 on
22. on the 4th of March.
23. Meg and I have
24. a surprise party he'll
25. forward to seeing you
26. Love, Beth
27. will be Perth in western
28. rent a van
29. Greetings from
30. We've had a great

Personal details and background

Gender: Male / Female (please, underline as appropriate)	
Year of Birth:	
Mother tongue:	
Education (degrees):	
At what age did you start learning English?	
For how many years did you learn English at school?	
Did you receive English tuition during your studies at university? Yes / No	
If yes, how many semesters / courses?	
How often do you use your English skills nowadays? (please, underline as appropriate)	
Writing: Daily / Weekly / Monthly / Once in 6 months / Less frequently	
Reading: Daily / Weekly / Monthly / Once in 6 months / Less frequently	
Speaking: Daily / Weekly / Monthly / Once in 6 months / Less frequently	
Listening: Daily / Weekly / Monthly / Once in 6 months / Less frequently	
Have you lived in an English-speaking* country? Yes / No	
If yes, when?	
Where?	For how long?
* e.g. United Kingdom, Ireland, South-Africa, USA, Canada, Australia, New Zealand, etc.	

With whom do you communicate in English? (please, underline as appropriate)

- **Mostly with native speakers** /**
- **Mostly with non-native speakers /**
- **Equally with both**

** e.g. British, American, Canadian, Australian, Irish, South-African, etc.

8.5 Appendix E: The Finnish informants

Informant No.	Year of birth	Highest degree	Educational institution	Age of learning Eng. (yrs)	Total amount of English tuition at school (yrs)	English during studies (No. courses)
F1	1982	Ph.D.	University	9	10	5
F2	1983	Bach.	U.A.S.	9	10	2
F3	1981	M.Sc.Econ.	University	9	10	6
F4	1983	BBA	U.A.S.	9	12	0
F5	1983	M.Sc.	University	9	10	3
F6	1983	M.A.	University	10	10	1
F7	1983	Bach.	U.A.S.	9	10	1
F8	1980	Ph.D.	University	9	10	1
F9	1982	M.Sc.	University	9	10	1
F10	1982	Bach.	U.A.S.	9	10	0
F11	1980	Bach.	U.A.S.	11	8	2
F12	1981	Bach.	U.A.S.	9	10	4
F13	1984	Bach.	U.A.S.	9	10	3
F14	1984	M.Sc.	University	9	10	0
F15	1983	LL.M.	University	11	8	0
F16	1984	M.Sc.	University	9	10	0
F17	1983	M.Sc.	University	11	8	0
F18	1985	M.Sc.	University	9	10	2
F19	1985	M.Sc.	University	9	10	0
F20	1982	Ph.D.	University	9	10	0
F21	1982	M.Sc.	University	9	10	2
F22	1981	M.A.	University	9	10	0
F24	1984	Ph.D.	University	9	10	10
F25	1981	Ph.D.	University	10	8	0
F26	1981	LL.M.	University	9	10	1
Average	1982,5			9,3	9,8	1,4

Informant No.	Living in an Eng. Speaking country?	At what age?	How long?	English use: listening*	English use: speaking*	Communication mostly with native speaker (1), non-native speaker (2), both (3)
F1	-	-	-	1	1	2
F2	-	-	-	2	2	2
F3	UK	13	1 yr	1	1	3
F4	-	-	-	3	3	2
F5	-	-	-	1	2	3
F6	-	-	-	1	1	2
F7	-	-	-	1	3	3
F8	UK	26→	9 yrs →	1	1	1
F9	UK	17	2 weeks	1	3	2
F10	-	-	-	1	2	2
F11	-	-	-	1	3,5	3
F12	-	-	-	1	3	3
F13	USA	28	1 yr	2	4	3
F14	-	-	-	1	1	2
F15	-	-	-	1	4	2
F16	-	-	-	1	1	2
F17	-	-	-	1	1	2
F18	-	-	-	1	1	3
F19	-	-	-	2	2	3
F20	Canada	28-31	3 yrs	1	1	3
F21	-	-	-	2	2	2
F22	-	-	-	2	2	2
F24	UK	31→	0,25 yrs→	1	1	3
F25	UK	32→	2 yrs→	1	1	3
F26	UK	27	1 yr	1	1	2
Average				1,3 Median = 1	1,9 Median = 2	Median = 2

8.6 Appendix F: The Austrian informants

Informant No.	Year of birth	Highest degree	Educational institution	Age of learning Eng. (yrs)	Total amount of English tuition at school (yrs)	English during studies (No. courses)
A1	1985	DI	FH	11	8	4 semesters
A2	1987	M.A.	University	10	8	-
A3	1986	LL.M.	University	10	8	-
A4	1980	Ph.D.	University	8	10	2
A5	1986	M.A.	University	11	8	-
A7	1983	M.Sc.	University	10	8	-
A8	1983	Lehramt	Pedak.	9	9	5
A9	1984	M.A.	University	10	8	-
A10	1985	M.A.	University	8	10	-
A11	1984	B.Sc.	FH	10	8	-
A12	1979	Mag.	University	6	10	-
A13	1983	M.Sc.	University	8	10	-
A14	1984	Mag.	University	9	9	6
A15	1984	Mag.	University	9	8	-
A16	1987	M.Sc.	University	7	12	6
A17	1987	M.Mag.	University	9	9	-
A18	1987	M.A.	University	11	8	-
A19	1981	Mag.	University	8	10	3
A20	1986	B.A.	University	9	9	6 semesters
A21	1980	Mag.	University	7	11	-
A22	1981	Mag.	University	8	11	3
Average	1983,9			8,9	9,1	

Informant No.	Living in an Eng. Speaking country?	At what age?	How long?	English use: listening*	English use: speaking*	Communication mostly with native speaker (1), non-native speaker (2), both (3)
A1	UK	22	1 yr	1	2	2
A2	-	-	-	3	3	3
A3	USA	16	2 weeks	2	4	3
A4	-	-	-	2	2	2
A5	-	-	-	1	5	1
A7	-	-	-	1	1	2
A8	-	-	-	2	2	2
A9	Various	29	0,5 yrs	1	2	2
A10	-	-	-	1	1	2
A11	-	-	-	1	1	2
A12	-	-	-	1	1	3
A13	-	-	-	1	1	3
A14	-	-	-	1	1	2
A15	-	-	-	1	1	2
A16	-	-	-	3	3	2
A17	-	-	-	2	2	2
A18	-	-	-	1	1	3
A19	UK	25	5 months	2	1	2
A20	USA	22	0,5 yrs	2	1	3
A21	-	-	-	2	2	3
A22	-	-	-	1	1	2
Average				1,5 Median = 1	2,3 Median = 2	Median = 2

8.7 Appendix G: Statistics of individual informant performance

WORD LIST	Mean	Median	Mode	Standard deviation	Range	Max. score	No. of informants with score \geq max. -1
/w/							
FIN	8,3	10	10	2,3	8	10	15
AUT	7,0	8	9	2,6	9	10	8
‘TH’							
FIN	7,5	9	9	3,3	11	12	4
AUT	7,2	8	10	3,2	10	12	2
/θ/							
FIN	4,0	5	6	2,2	6	6	14
AUT	3,9	5	6	2,1	6	6	11
/ð/							
FIN	2,9	3	3	1,2	5	5	7
AUT	2,7	3	3	1,3	5	5	6
Word-final voiced stops							
FIN	5,0	6	6	1,5	6	6	18
AUT	3,0	3	3	1,5	5	6	4
Flap*							
FIN	1,6	1	0	1,7	5	5	4
AUT	0,8	0	0	1,3	4	5	2

* Refers to the average amount of flapped word-medial prevocalic alveolar plosives

TEXT PASSAGE	Mean	Median	Mode	Standard deviation	Range	Max. score	No. of informants with score \geq max. -1
/w/							
FIN	13,3	14	16	3,0	10	16	12
AUT	12,2	13	14	3,1	10	16	5
‘TH’							
FIN	10,4	12	12	5,1	16	17	4
AUT	7,3	8	8	3,7	11	17	0
/θ/							
FIN	4,6	6	7	2,5	7	7	13
AUT	3,5	4	5	2,3	7	7	3
/ð/							
FIN	5,5	6	6	2,8	9	9	8
AUT	3,1	3	4	2,0	8	9	1
Word-final voiced stops							
FIN	5,5	6	5	1,9	8	8	9
AUT	4,6	5	5	1,1	4	8	1
Flap*							
FIN	2,7	2	1	2,0	7	7	2
AUT	2,0	1	0	2,3	7	7	2

* Refers to the average amount of flapped word-medial prevocalic alveolar plosives

8.8 Appendix H: Abstract

The aim of this M.A. Thesis is to assess the prevalence of stereotypical accents among Finnish and Austrian young adults by evaluating the informants' pronunciation of specific English consonant sounds that do not exist in either of the two L1s. Additionally, the possible influence of greater exposure to spoken English through media (e.g. TV) is investigated by analysing how common the American-style flapping is among the informants. The main hypothesis is that due to greater exposure to (American) English from an early age on, the stereotypical L1-style pronunciation is less prevalent and the American-style flapping is more common among the Finnish than the Austrian informants.

The hypothesis is tested by examining the audio samples from 25 Finnish and 21 Austrian informants who were asked to record their performance of two pronunciation tasks: a word list and a text passage. All informants were born between 1979 and 1987 and have at least a bachelor's degree, so ideally the only differences between the two informant groups are the L1 and the amount of exposure to spoken English. The analysis is primarily based on the visualisations of the audio data, i.e. waveforms and spectrograms. Due to the similarities in structure, the differentiation between (inter-)dental and labiodental fricatives is additionally based on the hearing of the author.

The results of the pronunciation analysis of all occurrences indicate a statistically significant correlation between nationality and the pronunciation of the investigated sounds. Only the pronunciation of the 'th'-sounds in the word list task and the word-final voiced plosives in the text passage task show no significant correlation. This overall correlation between informants' nationality and pronunciation supports the hypothesis: the Finnish informants pronounce the investigated sounds more frequently in a standard-like manner and the American-style flapping is more common among the Finnish than among the Austrian informants.

Even though the pronunciation may be influenced by many factors, such as problems in perception, misinterpretation of the written text, learned non-standard-like articulatory targets, or stronger influence of the different L1s, these results suggest that a greater exposure to spoken English is beneficial to acquiring the phonetic structures of the language.

8.9 Appendix I: Zusammenfassung

Das Ziel dieser Masterarbeit ist die Prävalenz von stereotypischen englischen Akzenten von Finnen und Österreichern zu beurteilen. Die Evaluation erfolgt durch die Auswertung der Aussprache von spezifisch englischen Konsonantlauten, die im Finnischen oder im österreichischen Deutsch nicht vorkommen. Darüber hinaus wird der mögliche Einfluss von mehr Exposition gegenüber gesprochenem Englisch durch verschiedene Medien (z.B. das Fernsehen) mittels der Analyse, wie häufig die amerikanische Aussprache (Flaps) unter den InformantInnen ist, untersucht. Die zentrale Hypothese dieser Arbeit ist, dass aufgrund der höheren Exposition mit (amerikanischem) Englisch von einem frühen Alter an, die stereotypische L1-Stil Aussprache weniger weit verbreitet ist und die amerikanischen Flaps bei den finnischen InformantInnen häufiger sind als bei den Österreichischen.

Die Hypothese wird durch die Untersuchung der Audiodaten von 25 finnischen und 21 österreichischen InformantInnen verifiziert, die ersucht wurden, ihre Aussprache mittels zweier Aufgabenstellungen aufzuzeichnen: einer Wortliste und einer Textpassage. Alle InformantInnen wurden zwischen 1979 und 1987 geboren und haben mindestens einen Bachelor-Abschluss. Im Idealfall sind die einzigen Unterschiede zwischen den beiden Gruppen daher die Muttersprache und das Ausmaß der Exposition gegenüber Englisch. Die primär verwendete Forschungsmethode basiert auf Visualisierungen der Audiodaten (Wellenformen und Spektrogramme). Aufgrund der Ähnlichkeiten der Strukturen, basiert die Unterscheidung zwischen (inter-)dentalen (/θ/ and /ð/) und labiodentalen (/f/ and /v/) Frikativlauten auf der Einschätzung des Verfassers.

Die auf diese Weise erhaltenen Ergebnisse zeigen eine statistisch signifikante Korrelation zwischen der Nationalität der InformantInnen und der Aussprache der untersuchten englischen Phoneme. Nur die Aussprache der 'th'-Laute in der Wortliste und die stimmhaften Auslautplosive in der Textpassage zeigen keine signifikante Korrelation. Diese allgemeine Korrelation zwischen der Nationalität der InformantInnen und der Aussprache unterstützt die Hypothese: die finnischen InformantInnen sprechen die untersuchten Laute häufiger gemäß der Standardaussprache aus und die im amerikanischen Englisch häufig vorkommenden Flaps treten bei den FinnInnen häufiger als bei den ÖsterreicherInnen auf.

Auch wenn die Aussprache durch viele Faktoren, wie zum Beispiel Probleme in der Wahrnehmung, falsche Interpretation des geschriebenen Textes, erlernte standardabweichende artikulatorische Ziele oder variierende Einflüsse der verschiedenen Muttersprachen, beeinflusst werden kann, legen diese Ergebnisse nahe, dass eine größere Exposition mit gesprochenem Englisch förderlich ist, um die phonologische Strukturen dieser Sprache zu erwerben.