DISSERTATIONES
PHILOLOGIAE
ESTONICAE
UNIVERSITATIS
TARTUENSIS
53

LIIS ERMUS

The phonetic variation of plosives in Estonian

DISSERTATIONES PHILOLOGIAE ESTONICAE UNIVERSITATIS TARTUENSIS

53

DISSERTATIONES PHILOLOGIAE ESTONICAE UNIVERSITATIS TARTUENSIS
53

LIIS ERMUS

The phonetic variation of plosives in Estonian

UNIVERSITY OF TARTU
Press

University of Tartu, Institute of Estonian and General Linguistics

Dissertation accepted for the defence of the degree of Doctor of Philosophy on June 11, 2024 by the Committee of the Institute of Estonian and General Linguistics, Faculty of Arts and Humanities, University of Tartu.

Supervisors: Associate Professor Pire Teras (University of Tartu)

PhD Meelis Mihkla (Institute of the Estonian Language)

Opponent: Assistant Professor, PhD Eleanor Chodroff (University of Zürich)

Commencement: August 26, 2024 at 14.15 at Ülikooli 18–139, Tartu

This study has been supported by the Graduate School of Linguistics, Philosophy and Semiotics; funded by the European Regional Development Fund (University of Tartu ASTRA Project PER ASPERA, Institute of the Estonian Language ASTRA Project EKI-ASTRA).

European Union
European Regional
Development Fund

Investing
in your future

ISSN 1406-1325 (print)
ISBN 978-9916-27-608-2 (print)
ISSN 2806-2450 (pdf)
ISBN 978-9916-27-609-9 (pdf)

University of Tartu Press
www.tyk.ee

ACKNOWLEDGEMENTS

Life's a rocky road. I am phenomenally good at wandering off, taking wrong turns and getting lost between shortcuts. This path, too, has become longer and harder than I planned. But fortunately, there have been many, thanks to whom I did not give up at various points.

First, I want to thank my supervisors, Dr. Meelis Mihkla from the Institute of the Estonian Language, who proposed the doctoral studies to me and Dr. Pire Teras from the University of Tartu, who agreed to supervise me for the third time in a row since my BA. They trusted me to proceed on my own but did not let me fall off the cliffs.

I am grateful to Dr. Pärtel Lippus from the University of Tartu, Dr. Einar Meister from Tallinn University of Technology and Dr. Eleanor Chodroff from University of Zürich for agreeing to review my thesis and giving their valuable feedback.

My scattered style of writing was upgraded by many anonymous and non-anonymous readers, reviewers and editors during the years. Special thanks go to Writing group no 7 and Reedeklubi who always managed to come up with good advice, despite 'not understanding a thing'.

Great thanks to my colleagues at IEL. Tiina Laansalu has been the best boss anyone could wish for, thank you for always being supportive for everything I do. I am grateful to Dr. Liisi Piits for accepting me as a full member of the speech research group and all the members for fruitful work and great company at conferences. I wanna thank all fellow doctoral candidates from IEL over the years, especially Kristina, Marit and others who have shared the burden of seminar organising.

Special thanks go to dr. Mari Uusküla and prof. Reili Argus from Tallinn University for giving me an opportunity to share my phonetic knowledge as a teacher and annoy several sets of bachelor students with differences in IPA and Estonian plosive spelling.

I have been supported by various projects during the studies: Institutional research project "Speech styles, sentence prosody, phonological variation: description, theory and modelling" (IUT35-1, 2015–2020), "Spontaneous speech synthesis" (EAG144, 2022), "Speech studies at the service of speech technology" (EKI-BAAS-2023/2, 2023–2024), "Centre of Excellence in Estonian Studies" (TK145, 2015−2023), Doctoral School of Linguistics, Philosophy and Semiotics in University of Tartu and in Institute of the Estonian Language. I am happy to have been able to attend various conferences, winter, and summer schools, seminars and workshops and meet all the lovely people there.

I could not have done it without friends and family. Thank you, Sven, Eva-Liis, Mihkel and everyone else in Omaklubi for providing an endless supply of kettlebells and good company. Thank you Ilmar, Tõnis, Maria, Chris, Evelyn and occasional members of Viimsi Koob for awesome Monday evenings in Uus Laine. Thank you Siim, for table game evenings and keeping the amount of sport questions under control. Thank you, Esta and Gerli for lunch breaks and occasional crosswords. Thank you, dear Kristina, for being my second sis.

I want to thank mom Linda and my sister Maarja who have always had their hearts and doors (including the ones of greenhouse and pantry) open for me. My husband Ilmar has firmly stood beside me all this time. Thank you for your unconditional love. Thank you for putting up with my moods and forcing me to go to bed at normal time.

All remaining mistakes are my own.

CONTENTS

LIST OF PUBLICATIONS

P1 Ermus, Liis. 2017. Eesti keele lühikeste klusiilide hääldusе variatsioon ja seda mõjutavad tegurid. *Mäetagused* 68. 27–52. https://doi.org/10.7592/MT2017.68.ermus.

P2 Ermus, Liis & Meelis Mihkla. 2019. Predictability of plosive reduction from written text in Estonian. In Sasha Calhoun, Paola Escudero, Marija Tabain & Paul Warren (eds.), *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019*, 2635–2639. Canberra, Australasia: Australasian Speech Science and Technology Association Inc.

P3 Ermus, Liis. 2019. Estonian geminate plosives: some durational characteristics. *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics* 10(1). 35–52. https://doi.org/10.12697/jeful.2019.10.1.02.

P4 Ermus, Liis. 2023. Acoustic characteristics of Estonian plosive bursts. In Oliver Niebuhr (ed.), *Proceedings of the 13th International Conference of Nordic Prosody*, 94–105. Sciendo. https://doi.org/10.2478/9788366675728-007.

1. INTRODUCTION

1.1. Objectives

Plosives belong to the sound class of obstruents, together with affricates and fricatives. Plosive sounds differ from other sounds in articulatory composition and acoustics (Johnson 2003: 135). Articulatorily, the plosive comprises three phases: a closure is formed in the vocal tract, obstructing the airflow until the pressure in the tract forces the closure to open. Acoustically, plosives are more often divided into two phases: a closure and a burst. These phases can be characterised based on duration and acoustic features like voicing during the closure or the spectrum of the burst. Aside from the distinctions based on place of articulation, contrastive categories may be formed based on voicing or length. The phonological distinctions can occur in either word-initial, word-medial or word-final position or in a combination of those, and the allophonic variation in a language is dependent on that; it is believed that fewer distinctions allow more variation, including reduction inside one distinctive category (Keating, Linker & Huffman 1983).

The present thesis concerns the acoustic variation of plosive sounds in Estonian connected speech. Estonian is a Finno-Ugric language with four plosive phonemes in all main articulation places: bilabial /p/, alveolar /t/, palatalised alveolar /tʲ/, and velar /k/. Estonian does not use the voicing distinction in obstruents. Instead, Estonian has a quantity distinction with three length categories – short, long, and overlong – that can be applied to both vowels and consonants. The quantity system is complex and includes suprasegmental factors besides sound durations, like syllable duration ratios and pitch movement (Lippus et al. 2013). Besides quantity, lenis and fortis distinction is sometimes used for obstruents as well (Metslang et al. 2023: 57–58).

The thesis aims to describe the acoustic characteristics of plosives in Estonian connected speech from a broader typological perspective and to find out the main factors influencing the variation in connected speech. The aim is motivated by the need to fill gaps in the acoustic description of plosives in Estonian. Sporadic studies of plosive acoustics of Estonian have been conducted throughout the 20th and 21st centuries, but apart from the first one (Ariste 1933a; 1933b) that covered three quantity degrees and word-initial position, all the acoustic studies have been concerned with acoustics of singleton plosives (Raasik 2010; Suomi & Meister 2012). The quantity in Estonian has mostly been studied on a suprasegmental level; in geminates, durational aspects of segments have been covered (Lehiste 1966; Eek 1974; Türk 2018; Türk 2019a). Additionally, Türk, Lippus & Simko (2017) have studied plosive gemination from an articulatory perspective. The word-initial position has mostly been researched from the temporal aspects under different stress conditions (Eek & Meister 2003; Suomi & Meister 2012; Suomi et al. 2013). Voicing in the closure phase has been studied by Ariste (1933a; 1933b) Spectra of burst phases have been researched only in word-initial plosives (Eek & Meister 1996a; 1996b; 1996c). Acoustic characteristics of the burst phase have not been studied concerning gemination or quantity. Most studies have been conducted on small data sets and lab speech. The current study will contribute to the existing body of research in allophonic and durational variation of plosives using corpus-based data.

In the thesis, I seek answers to the following research questions concerning the allophonic and durational variation of the plosive allophones and the acoustic behaviour of the burst phase.

RQ1: What kind of allophonic variants emerge in plosives in Estonian connected speech?

H1 There occurs a large variety of plosive allophones in Estonian (Keating, Linker & Huffman 1983; Raasik 2010).

RQ2: How do segmental and suprasegmental features (stress, quantity) influence the allophonic variation of plosives?

H1 There are differences in the allophonic variation according to the place of articulation (Raasik 2010; Recasens 2016).

H2 The vowel context influences the allophonic variation (Raasik 2010).

H3 The stressed position of the plosive carrying word influences the allophonic variation.

RQ3: How does the speech register influence the allophonic variation of plosives?

H1 There are more reduced allophones in spontaneous speech compared to read speech (Raasik 2010; Suomi & Meister 2012; Ernestus, Hanique & Verboom 2015).

RQ4: How do segmental and suprasegmental features (syllable structure, stress, quantity) influence the durational variation of plosives?

H1 Durational patterns in singletons and geminates are similar according to the place of articulation and allophones.

H2 There are significant differences in the segment durations between quantity degrees.

H3 Syllable structure, i.e. the phonological length of the preceding vowel influences the duration of the geminate plosives (Eek 1974).

H4 The vowel context influences the durations in singleton plosives (Raasik 2010).

H5 The stressed position of the plosive carrying word does not influence the duration of the plosive (Suomi & Meister 2012).

RQ5: How do segmental and suprasegmental features influence the acoustics of the burst phase?

H1 There are no differences in burst duration between quantity degrees or between word-initial and word-medial positions (Esposito & Di Benedetto 1999; Doty, Idemaru & Guion 2007).

H2 The burst intensity is higher in geminates, similar to Finnish (Doty, Idemaru & Guion 2007).

H3 The burst intensity is not higher in the word-initial, i.e. primary-stressed position, compared to the word-medial non-stressed position (Cho & Keating 2009; Cho & McQueen 2005).

H4 Geminates show a higher centre of gravity, higher spectral standard deviance and lower skewness compared to the singletons (Al-Tamimi & Khattab 2011).

H5 There are differences in the burst spectra between word-initial and word-medial singletons (Tabain et al. 2016).

1.2. Structure of the dissertation

The thesis comprises a summary overview and four publications from 2017–2023. The introductory part is divided into six chapters. Chapter 1 introduces the thesis, the objectives, and the main research questions. It also provides an overview of the publications and the author's contributions to the co-authored publications. Chapter 2 explains a general theoretical background: articulatory and acoustic aspects in plosive production, phonological categorisation systems of plosives in different languages, and coarticulation phenomena in plosive production. Chapter 3 describes the data and methods used in the thesis. Chapter 4 discusses the results of the publications according to the research questions. Chapter 5 gives conclusions. The summary in Estonian is provided in chapter 6. This is followed by the references cited in the introductory part. The second part of the dissertation consists of four publications.

1.3. Overview of the publications and author's contributions

The four publications of the thesis explore different aspects of Estonian plosive acoustics. The articles represent a corpus-linguistic approach to the subject. This approach was chosen because of the wish to capture the phenomena from natural fluent speech rather than from specially crafted lab material. An overview of the main topics of the publications is given below. For the co-authored publications, the authors' contributions are described.

P1 studies the acoustics of short plosives in spontaneous speech and factors influencing the variation. Allophonic variation of intervocalic short plosives in most frequent content words containing intervocalic short plosives on the boundary of the 1st and the 2nd syllable (VCV) were studied. Effects of vowel context and stressed position of the word carrying the plosive were sought.

P2 presents the results of a pilot study of the connections between allophones of short plosives and textual features. The purpose was to find the main factors influencing the allophonic variation of the short plosives that are obtainable from written text (e.g., adjacent letters, position of the letter in a syllable or a word, etc). Liis Ermus annotated the material and prepared the data set. Statistical analyses and interpretation of the results were done jointly. Liis Ermus wrote the main part of the introductory chapter and made the figures; other parts of the publication were written jointly.

P3 focuses on the durational characteristics of geminate plosives. Intervocalic geminate plosives (sequences VCCV, VC:CV, VVCCV, VVC:CV) in semi-spontaneous and read speech were studied. Duration of closure, burst phase, and the voiced transition in partly voiced closures were measured. The study sought the influences of quantity and syllable structure (length of the preceding vowel) on the durational behaviour of the plosives.

P4 focuses on the duration and acoustic characteristics of the burst phase in plosives in word-initial and word-medial positions. In the article, intervocalic word-initial and word-medial plosives in semi-spontaneous speech were studied. VOT, intensity and spectral measures (centre of gravity, standard deviation, skewness) of the burst were measured. Influences of the place of articulation and length category of the plosive were sought.

2. THEORETICAL BACKGROUND

This chapter explains the key terms and concepts and gives a theoretical overview of the characteristics of plosives. First, basic aspects of plosive articulation, acoustics, and main categorising systems are introduced in Chapter 2.1. and 2.2. respectively. Chapter 2.3 describes the main factors influencing the coarticulation of plosives in connected speech. Chapter 2.4. gives the overview of the Estonian plosive system and the previous research on the background of the information provided in previous chapters.

2.1. Articulatory and acoustic characteristics of plosives

Plosives occur in all known languages (Maddieson 2013). The most common places of articulation (POA) of plosives are bilabial [p], alveolar [t], and velar [k]. Voiceless velar [k] is the most common plosive, being represented in 90% of the 2186 languages in the Phoible phonetic database (Moran & McCloy 2019).

Plosives (along with affricates and fricatives) belong to the sound class of **obstruents** – the sounds where airflow is obstructed during articulation. The plosives differ from other consonants in that acoustically, they are dynamic, while the acoustics of other consonants can be described as stable or monolithic (Johnson 2003: 135; Clark, Yallop & Fletcher 2007: 43). During the articulation of the plosive, three phases can be distinguished (Ladefoged & Johnson 2011: 14). First, total closure is formed in the articulatory tract during the implosion or closing phase. The second phase is **occlusion**, the holding of the closure. Air gets trapped and raises the pressure in the tract. Finally, the pressure forces the closure to open with a **burst** or **explosion**. The forming and holding of the closure may or may not be distinguishable acoustically. So, in acoustic phonetic studies, the plosive is mostly divided into distinctive phases: the closure and the burst phase, of which the first can be voiceless, partly or fully voiced (see Fig 2–4 in Chapter 4.1.1., p. 29–30).

The burst phase is often described using **voice onset time** (**VOT**) (Lisker & Abramson 1964). VOT is the time from the release of the closure to the beginning of the regular wave or voicing of the following sound segment (Cho & Ladefoged 1999), usually measured in milliseconds (ms). VOT is positive when the regular wave starts after the release of the closure and negative when the regular wave (voicing) starts before the release of the closure. In the latter case, the time before the opening is measured. Positive VOT often coincides with the duration of the burst[1] but may outlast it. VOT was first used as a measure for voicing distinction but is now frequently reported as a burst phase measure also in languages with

[1] As Estonian does not use any categorisation based on VOT and as the burst duration was reported in most of the publications, I henceforth use the term burst duration to refer to Estonian.

no voicing contrast, where it roughly describes the duration of the burst phase (e.g., Arvaniti 1999; Doty, Idemaru & Guion 2007; Genee & Li 2023). Three categories are commonly used when grouping the VOT durations: lead/negative VOT, short-lag VOT and long-lag VOT. Lisker and Abramson (1964) studied VOT in 11 languages and found the following VOT ranges for each: lead –125 – –75 ms, short-lag 0–25 ms, and long-lag 60–100 ms, with median values of –100 ms, 10 ms and 75 ms, respectively. Chodroff et al. (2019) calculated the grand average from over 100 languages: <0 ms for lead, 0–35 ms for short-lag and >35 ms for long-lag VOT.

The **place of articulation (POA)** is the primary influencer of the durations of closure and burst phase/VOT. Cho and Ladefoged (1999: 209–213) have listed the main aerodynamic factors that affect VOT durations, but as the article concerns only voiceless plosives, these factors can be generalised to the burst phase. Different articulation places differ in the size of the supraglottal cavity on both sides of the constriction: the larger the cavity behind the constriction, the longer the closure; the larger the cavity in front of the closure, the longer the burst phase. The speed of articulators and the contact between the articulators influence the speed of release of the closure: the faster the articulators are, the shorter the burst phase; the larger the contact between the articulators, the longer the burst phase. Finally, the size of the glottal opening has an influence: a smaller glottal opening causes longer burst phase. So, it can be said that the durations of closure and burst move in reverse direction: the longer the closure, the shorter the burst phase. Therefore, the bilabial /p/ usually has the shortest and velar /k/ the longest burst phase.

The amplitude and **intensity** or of the burst characterise the power of the articulation (Hayward 2000: 43–44). These are measured in decibels (dB). Intensity is usually measured relative to the intensity of the following segment by subtraction (e.g., Vicenik 2010; Nicolaidis et al. 2019) or dividing (e.g., Ridouane 2007) the two adjacent intensities.

The spectrum of a burst phase is often characterised using measures that are called **spectral moments** (Forrest et al. 1988). Spectral moments describe the shape and composition of the burst spectrum. There are four spectral moments commonly used: centre of gravity (the first moment), standard deviation of the burst (the second moment), skewness (the third moment) and kurtosis (the fourth moment). **Centre of gravity** (COG) shows the highest energy concentration in the spectrum frequencies; **standard deviation** shows the range of the energy in the spectrum (Forrest et al. 1988; Jongman, Wayland & Wong 2000; Vicenik 2010; Tabain & Butcher 2015). COG and SD are measured in Hz.

Skewness and kurtosis are unitless measures that describe the spectrum's shape compared to the normal distribution. **Skewness** describes the asymmetry of the energy in the spectrum: a positive value indicates energy concentration on lower frequencies and vice versa (the value is zero for normal distribution) (Forrest et al. 1988; Jongman, Wayland & Wong 2000; Tabain & Butcher 2015; Guthrie 2020). **Kurtosis** value describes the spectrum's shape: a positive value indicates a flat spectrum with no clear energy peaks, and a negative value indicates a clear peak in the spectrum (the value is zero for normal distribution). (Forrest et al. 1988; Tabain & Butcher 2015; Guthrie 2020)

A more frontal place of articulation gives higher COG and lower skewness (Fant 1960; Harrington 2010). Based on previous studies (e.g., Forrest et al. 1988; Cole et al. 2007; Chodroff & Wilson 2014; Nicolaidis et al. 2019), spectra and spectral moments of different places of articulation can be described as follows: labials have diffuse and falling spectra that are more intense on lower frequencies; alveolars have diffuse and rising spectra, which are more intense on higher frequencies; velars have a compact spectrum that is most intense in mid-frequencies. Labials tend to have the highest SD of the spectrum. Alveolars tend to have the highest COG and the lowest skewness. Velars tend to have the lowest COG and SD.

2.2. Plosive categorisation systems

Many languages have some distinctive categories for plosives besides place of articulation. This chapter overviews the primary phonological systems for categorising plosives: **voice**, **fortis**, and **length**. None of these systems is used exclusively for plosives. Length can be used for all sounds, gemination for all consonants, voicing and fortis for obstruents, including affricates and fricatives.

2.2.1. Voice

The most common categorical distinction used in obstruents is the voice distinction. Maddieson (2013) lists 567 languages, of which 347 (61%) have voicing distinction either in plosives or all obstruents. Phonetically, voicing manifests in weak glottal pulsing during the closure phase.

Voicing distinction usually includes the voiceless unaspirated category and the voiced category and/or voiceless aspirated category (Keating, Linker & Huffman 1983). However, some languages also have a voiced aspirated category, such as Hindi (Singh & Tiwari 2016). Classical voicing languages, like French (Fougeron & Smith 1993) and Japanese (Okada 1991), have voiceless unaspirated and voiced categories where the difference between the categories is in the voicing of closure.

In languages with voicing contrast in word-initial or syllable-initial position, the distinction is often based on VOT values; the voiced category has lead/negative VOT, and the voiceless category has short lag VOT. Some languages, like English (Roach 2004; Cox & Palethorpe 2007) and Swedish (Engstrand 1990), also have an aspirated category in a stressed position. The aspirated category usually has VOT over 60 ms (Lisker & Abramson 1964).

Voiceless obstruents (incl. plosives) occur more frequently than voiced ones (Keating, Linker & Huffman 1983; Ohala 2013: 666). Ohala (1983) has listed physiological constraints that obstruct the voicing in plosives: the longer the closure is held, the more likely the plosive is voiceless (or voiced geminate plosive gets devoiced) because of the air pressure rising and blocking glottal movements; the more backwards the closure is formed, the more likely the plosive is voiceless because the cavity behind the closure is smaller and therefore air pressure needed to stop the glottal movements rises faster.

2.2.2. Fortis and lenis

Lenis and fortis distinction is the most abstract phonological distinction used for obstruents. The distinction between lenis and fortis is based on articulatory force or effort (Jaeger 1983; Kohler 1984). There is no single acoustic correspondent feature for effort. According to Jaeger (1983), the fortis consonants are claimed to be produced with greater pulmonic force, greater articulatory force, longer durations, and differences in glottal factors. Burroni et al. (2021) have listed acoustical correlates for the effort force in fortis plosives: longer duration, higher intensity, differences in VOT, lower spectral tilt, higher f0, and higher articulatory resistance (more voiceless). The most frequent correlates in plosives are duration, voicing and differences in the burst. Fortis category is longer in duration, voiceless, and has longer or louder bursts. Lenis is pronounced weaker – with less force, so it is shorter, more often voiced and reduced. The measurable features for fortis and lenis overlap with features of voice and length categories. Lenis category frequently shows reduction patterns similar to the voiced category. Fortis plosives usually have longer closures, like geminates.

Plosives, for example, in Standard German (Jessen 1998), Korean (Cho, Jun & Ladefoged 2002), and Mixtecan Trique (Dicanio 2012) are described using fortis and lenis contrast. Sometimes, two systems are used parallelly in phonological descriptions; this is the case for Swiss German (Ladd & Schmid 2018 for fortis; Kraehenmann 2001 for length).

2.2.3. Length

In length distinction, consonants can be viewed as short and long or singletons and geminates. Ladefoged and Johnson (2011: 251) have defined geminate as a long consonant that can be analysed as a sequence of two consonants of the same nature. However, the phonological nature of the geminates is controversial because it is difficult to map the distinction to the phonetic realisation (Lehiste 1970: 44–45; Davis 2011). Geminates occur in a large variety of languages in many language families: Finno-Ugric languages Estonian (Asu & Teras 2009), Finnish (Suomi, Toivanen & Ylitalo 2008), Livonian (Lehiste et al. 2008), Ingrian (Markus et al. 2013) and Saami languages (Türk et al. 2019); Indo-European languages Bengali (Kotzor, Wetterlin & Lahiri 2017), Cypriot Greek (Arvaniti & Tserdanelis 2000), Italian (Esposito & Di Benedetto 1999; Payne 2005), Swiss German (Kraehenmann 2001); Semitic Maltese (Galea 2016); Austronesian Kelantan Malay (Hamzah, Fletcher & Hajek 2016), Pattani Malay (Abramson 1986); Afro-Asiatic Tashlhiyt Berber (Ridouane 2007); Japanese (Kawahara 2005), etc.

Usually, the distinction between short and long consonants or singletons and geminates is binary, but a few languages also have a ternary contrast (e.g., Estonian, Saami, Livonian). In most languages, geminates occur only word-medially, but some languages also have word-final (e.g., Bengali, Estonian) and even word-initial geminates (e.g., Malay, Maltese, Swiss German). In most quantity languages studied, the distinction between singleton and geminate or short and long categories is based on closure duration (Lehiste 1970: 33–35; Kubozono 2017: 2). Geminate consonants have about 1.5 to 3 times longer durations than singletons (Ladefoged & Maddieson 1996: 33; Hamzah, Fletcher & Hajek 2016: 137). Differences in adjacent segmental context and non-durational features, such as the length or quality of adjacent vowels, are also observed during gemination. For example, in Italian, vowels preceding geminates are shorter than vowels preceding singletons (Esposito & Di Benedetto 1999), but vice versa in Japanese (Idemaru & Guion 2008).

Although there are aerodynamic constraints for voiced geminate plosives, there still are many languages that contain these; they include, for example, Bengali (Reetz, Mikuteit & Lahiri 2019), Italian (Esposito & Di Benedetto 1999), Lebanese Arabic (Al-Tamimi & Khattab 2018), Maltese (Galea 2016), Pattani Malay (Abramson 1986) and Tashlhyit Berber (Ridouane 2007).

Singleton and geminate plosives may show differences in VOT or burst intensity along with the closure duration. However, the patterns differ between languages. For example, the VOT in geminates was longer than in singletons in Cypriot Greek in Arvaniti (2001), but the opposite pattern was found in Turkish by Lahiri and Hankamer (1988). Higher burst intensities in geminates compared to singletons have been found in Finnish (Doty, Idemaru & Guion 2007) and Tashlhiyt Berber (Ridouane 2007). No studies have compared bursts between singletons and geminates in Estonian.

2.3. Plosives in connected speech

This chapter first introduces the main concept of coarticulation, followed by a brief cross-linguistic overview of allophonic variation in plosives.

2.3.1. Coarticulation

In connected speech, sounds of speech are not realised completely because the movements of different articulators overlap and interact. **Coarticulation** is a term that refers to the mutual influence of adjacent or nearby articulatory segments, so they become like each other (Kühnert & Nolan 2006: 7). The need for the concept of coarticulation came from the need to explain the transformation of the discreet phonological segments to continuous speech. The term was first used in 1933 by Menzerath and de Lacerda (1933, cited by Farnetani & Recasens 2013). Coarticulation is a universal phenomenon observed in all languages. At the same time, co-articulatory processes are, to some extent, language-specific.

Coarticulation mostly manifests in speech in two ways: in reduction and assimilation. **Reduction** in phonetic level refers to shortening and less clear articulation or even absence of segments compared to their canonical forms under the influence of segmental and prosodic context (Harrington 2013: 91; Zellers, Schuppler & Clayards 2018: 1).

Assimilation is often used synonymously with coarticulation. Farnetani and Recasens (2013: 320) define assimilation as "contextual variability of speech sounds, by which one or more of their phonetic properties are modified and become like those of the adjacent segment". The distinction between coarticulation and assimilation can be made based on the level of the speech process. Ladefoged and Johnson (2011) have defined coarticulation on the articulation level as "[t]he overlapping of adjacent articulations" (p. 306) and assimilation on the phonological/perceptive level as "[t]he change of one sound into another, making it more similar to a neighboring sound" (p. 305). Sometimes, assimilation is used to refer to a context-sensitive coarticulation that would be marked in phonetic transcription (Volenec 2015: 74–75), but the division is rather vague. Three main types of assimilation are place assimilation, manner assimilation and voice assimilation (Clark, Yallop & Fletcher 2007: 89). In **place assimilation**, one segment influences the articulation place of the other (alveolar /n/ becomes a velar [ŋ] in front of /k/, like in Estonian *kang* [kɑŋk] 'bar'). In **manner assimilation,** one segment influences the articulation manner of the other (in Estonian, the sequence [hts] in the word *lihtsalt* 'simply' assimilates to [sːs] lisːsɑlt). In **voice assimilation,** one segment influences the voicing of the other (in Estonian, word-final voiced consonants become voiceless after voiceless obstruents, *lehm* [lehːm̥] 'a cow').

The coarticulatory influence of adjacent segments can be two-way. In the case of the left-to-right or **carryover** coarticulation, the segment's feature spreads to the next segment. In the case of right-to-left or **anticipatory** coarticulation, the segment's feature spreads to the preceding segment(s). (Clark, Yallop & Fletcher 2007: 86). Carryover coarticulation is believed to happen due to the inertia in the articulators. A common carryover coarticulation phenomenon is carryover voicing, often observed in obstruents. Anticipatory coarticulation happens because of the simultaneous adjustments in the articulators for several segments. Classic examples of anticipatory coarticulation are palatalisation and labial coarticulation. The extent of coarticulation depends on many factors, like the articulatory properties of the sounds, the phonological composition, the phonetic distinctions possible in the language, prosodic structures, speech rate, speaking style, etc. More reduction is observed in spontaneous speech and faster speech.

2.3.2. The allophonic variation of plosives

In plosives, the coarticulation most often manifests in (partial) voicing in the closure phase and durational shortening in both the closure and the burst phase or VOT. Carry-over voicing seems to be more common in plosives. Davidson (2016; 2018) has distinguished four types of voicing in obstruents in American English according to which part of the closure becomes voiced:

(a) **bleed** – voicing from the previous segment carries on to the beginning of the closure;

(b) **through** – voicing carries throughout the closure;

(c) **negative VOT** – voicing occurs before the burst;

(d) **hump** – voicing occurs and then ceases somewhere during the closure.

In American English, voiced plosives showed at least partial voicing in over 70% of the tokens, and most of the tokens fell into bleed and through types (Davidson 2016). Also, about half of the observed voiceless plosive tokens in Davidson (2018) were at least partly voiced, and over 80% of the partly voiced tokens fell into the bleed type.

Due to the weakening of articulation, the burst phase can shorten, become voiced in voiceless plosives and, in the last stage, reduce to complete loss. This coarticulatory behaviour of the burst phase has often been found in both voiced and voiceless plosives in Romance languages, like French and Spanish (Duez 1995; Torreira & Ernestus 2011), and also in the voiceless plosives in Estonian (Raasik 2010; Suomi & Meister 2012).

It has been found that different articulation places show different resistance to reduction. The effects of adjacent context on the pronunciation and vice versa depend on the mobility of the articulator and its participation in the pronunciation of the surrounding segments. According to the gestural models (Browman & Goldstein 1992; Fowler & Saltzman 1993), segments that share more articulators with adjacent segments are more influenced by coarticulation and assimilate more. So, among the plosives, bilabial [p] should be the least influenced by adjacent vowels because lips are only shared articulators and velar [k] the most influenced by the adjacent vowels because it uses the tongue body that is also used in the articulation of vowels (Fowler & Saltzman 1993). However, according to model of degree of articulatory constraint (Recasens, Pallarès & Fontdevila 1997; Recasens & Espinosa 2009), velar [k] should be the most resistant to coarticulation and influence other adjacent segments most.

The reduction rate is influenced by the speech register, being more extensive in spontaneous speech (Lindblom 1990; Moon & Lindblom 1994) and can result in the loss of whole syllables because of several simultaneous segmental deletions (Schuppler et al. 2011).

Durational and assimilatory patterns are influenced by suprasegmental and prosodic features like stress (word and sentence stress) and prosodic boundaries (phrase, utterance), durations being longer and the assimilation weaker in more prominent positions (Keating et al. 2004). In plosives, the extension of the overall and burst phase duration and VOT and the higher amplitude of the burst have been found in the stressed position, with contrastive stress having a more significant effect, while non-contrastive sentence stress has little or no effect. For example, in English, the characteristics of the voicing distinction (duration of the closure, VOT, burst amplitude) were more extreme in the contrastively stressed position, while noncontrastive stress only influenced the VOT (Cole et al. 2007; Cho & Keating 2009). In German, which uses lenis and fortis distinction, Kuzla and Ernestus (2011) found that the categories behaved differently in different stress conditions: the closure duration was prolonged under contrastive stress in both lenis and fortis; differences in VOT occurred only in fortis that showed longer VOT in stressed positions.

2.4. Estonian plosive system

Estonian has four plosive allophones in three places of articulation: bilabial /p/, alveolar /t/, palatalised alveolar /tʲ/[2] and velar /k/. All plosives are voiceless and unaspirated.

Estonian falls into the rare set of languages, along with Livonian, Ingrian, and some Saami languages, which have a ternary quantity system (see Türk 2019b). For consonants, the first quantity degree (Q1) corresponds to the singleton, the second quantity degree (Q2) to the short geminate, and the third quantity degree (Q3) to the long geminate. The duration ratios of the plosive segments are 1.7–2.1 for Q2/Q1, 1.2–1.4 for Q3/Q2 and 2.1–2.6 for Q3/Q1, as measured by Türk (2018; 2019a). However, the quantity is a suprasegmental phenomenon. It includes the duration ratio of stressed and unstressed syllables in a disyllabic foot and pitch contour in addition to the duration of the segment. The duration ratios of the syllables are roughly 2/3 for Q1, 3/2 for Q2 and 2/1 for Q3, not including the syllable onset consonants that do not participate in length distinction (Lehiste 1960). The syllable rates found in different studies are 0.6–0.9 for Q1, 1.4–2.0 for Q2 and 2.0-3.9 for Q3 (overview in Meister 2011: 29). The word-initial position does not participate in quantity distinction, but it carries primary stress (except in some onomatopoetic and loanwords) (Asu et al. 2016: 127). Voiced consonants can occur as singletons after short and long vowels and as geminates only after short vowels. Voiceless obstruents, on the other hand, can occur as singletons and geminates both after short and long vowels and diphthongs. Singletons can occur in all three quantity degrees in combination with vowels. So, plosives can occur intervocalically in seven possible structures, as shown in Table 1 (examples derived from Prillop 2015: 184).

[2] As the phonetic outcome of the palatalisation is very varied, especially in spontaneous speech, the palatalised phoneme is not distinguished from the alveolar in this thesis.

Table 1. Possible segment structures in three quantity degrees in Estonian plosives.

Quantity 1 (Q1):	short V1 singleton C2 (VCV)	*ude* [ute] 'a vellus hair, nom.sg'
Quantity 2 (Q2):	long V1 singleton C2 (VVCV)	*uude* [u:te] 'innovation, nom.sg'
	short V1 geminate C2 (VVCV)	*ute* [utte] 'an ewe, gen.sg'
	long V1 geminate C2 (VVCCV)	*uute* [u:tte] 'new, gen.pl
Quantity 3 (Q3):	long V1 singleton C2 (VV:CV)	*uude* [u::te] 'new, ill.sg'
	short V1 geminate C2 (VC:CV)	*utte* [ut:te] 'an ewe, par.sg'
	long V1 geminate C2 (VVC:CV)	*uute* [u:t:te] 'innovation, gen.sg'

Word-initial singletons have been studied from durational aspects, mostly in quantity studies (e.g., Suomi et al. 2013). The durations of word-initial singletons usually overlap with the word-medial singletons, except in the case of contrastive sentence stress, when the word-initial segment undergoes more significant elongation compared to the word-medial segment (Eek & Meister 2003; Suomi & Meister 2012; Suomi et al. 2013).

Besides quantity, lenis and fortis distinction has been used for Estonian (Ariste 1933a; 1939; 1968; Metslang et al. 2023: 57–58). It is said that geminates are fortis, and word-medial and word-final singletons in voiced context are lenis. The word-initial position is either classified as fortis (Ariste 1933a; Metslang et al. 2023) or neutralised (Eek & Meister 1996a: 67). The view may be supported by the orthography of Estonian that uses the same graphemes – *p*, *t*, *k* – for marking word-initial singletons and geminates and *b*, *d*, *g* for marking word-medial singletons. Ariste (1933a) found that the fortis distinction should be used, as Estonian does not have voicing distinction. Later, Ariste (1968: 40–41) listed all the constraints for fortis and lenis in Estonian and concluded that fortis could only occur in an absolute initial and final position; in other positions, it is either lenis or geminate. No experimental studies have been conducted so far to find differences in articulatory effort or corresponding acoustic features. The current thesis provides a short discussion of Estonian plosive acoustics' compatibility for the distinction.

The burst phase of Estonian plosives has only been studied scarcely. Durational aspects of bursts in singleton plosives have been studied in read (Suomi & Meister 2012) and spontaneous speech (Raasik 2010). The relationships in burst phase durations have been in accordance with the universal tendencies, being the shortest in bilabials and the longest in velars. Overall, the burst phases were shorter in read speech. The burst spectra in the word-initial position have been described by Eek and Meister (1996c). The spectrum of /p/ was diffuse and weakening in upper frequencies; the strongest spectral peak was under 500 Hz. The spectrum of /t/ was diffuse and strengthening in upper frequencies, having a decline in intensity from 4000 Hz to 5000 Hz. The spectrum of /k/ had the strongest peak between 700 and 2900 Hz. Spectral peaks depended on the following vowel, especially in /k/. Bursts of geminates have not been studied, nor have any studies compared burst phases in word-initial and word-medial positions. Publications

P1 and P2 will investigate the durational aspects of singletons and geminates, respectively. The P4 will add new knowledge about Estonian plosive bursts in aspects of duration, spectral composition, and intensity that may give information to support or reject the suitability of the fortis distinction in Estonian.

There have been few studies on the coarticulatory behaviour of Estonian plosives. Closure voicing has been found in singleton plosives both in read (Ariste 1933b; Suomi & Meister 2012) and spontaneous speech (Raasik 2010). All the studies found voicing in the closure phase. Two latter studies also observed burst phase reduction. The proportion of reduced tokens was dependent on the speech register, with a lower proportion in read (/p/ 1.1%, /t/ 3.3%, /k/ 25.9%) and higher in spontaneous speech (/p/ 15%, /t/ 22%, /k/ 48%). Velar plosive /k/ has been standing out with weak articulation, showing the greatest reduction. The closure voicing has also been found in geminates by Ariste (1933b; 1933a). Coarticulatory patterns in geminates have not been researched further. The publications P1, P2 and P3 will further study the allophonic variation and reduction in singleton and geminate plosives in Estonian, adding a comparative view of allophonic variation in two speech registers.

3. MATERIAL AND METHODS

In this chapter, materials and analysis methods used in publications are presented.

3.1. Acoustic material

The speech material for most of the articles (P1, P3, P4) came from the University of Tartu Phonetic Corpus of Estonian Spontaneous Speech (PCESS) (Lippus 2018; Lippus et al. 2021; Lippus et al. 2023). The read-aloud material for P2 and P3 came from the Speech Synthesis Training Corpus (SSTC) of the Institute of the Estonian Language (Piits 2016).

The PCESS consists of several sub-corpora: spontaneous dialogues recorded in a sound studio, monologues recorded in a lecture hall using a head-mounted microphone, and dialogues recorded in quiet conditions during fieldwork. The sound files are in 16-bit 44.1 kHz PCM wave format. Speakers are chosen to represent all age groups and geographical areas of Estonia. Sound files are annotated[3] using Praat software (Boersma & Weenink 2023) on several tiers: word, SAMPA (Speech Assessment Methods Phonetic Alphabet) (Wells 2005), syllables, metric feet, morphological annotation, speech quality, intonation phrases and interpausal units. The current version of the corpus (v1.3) contains about 135 hours of speech from 207 speakers.

The SSTC is compiled of read-aloud sentences from literature and news texts. The readers are the voice donors for the text-to-speech synthesis. Most of the readers are professional news readers and actors. The corpus of one voice includes a specific set of sentences that covers all possible sound combinations in the language and additional material from literature and news texts. The speech material is annotated automatically on word and phoneme levels using the web-based annotation service WebMAUS (Kisler, Reichel & Schiel 2017). The current version of the corpus contains about 152.5 hours of speech from15 speakers, including two children and three dialect speakers[4].

3 Overview of annotation principles https://foneetikakorpus.ut.ee/ekskfk_info_eng.html#Annotation; Corpus annotation guide (in Estonian) https://foneetikakorpus.ut.ee/ekskfk_margendamise_juhend.html

4 https://koneveeb.ee/korpused/ (10.06.2024)

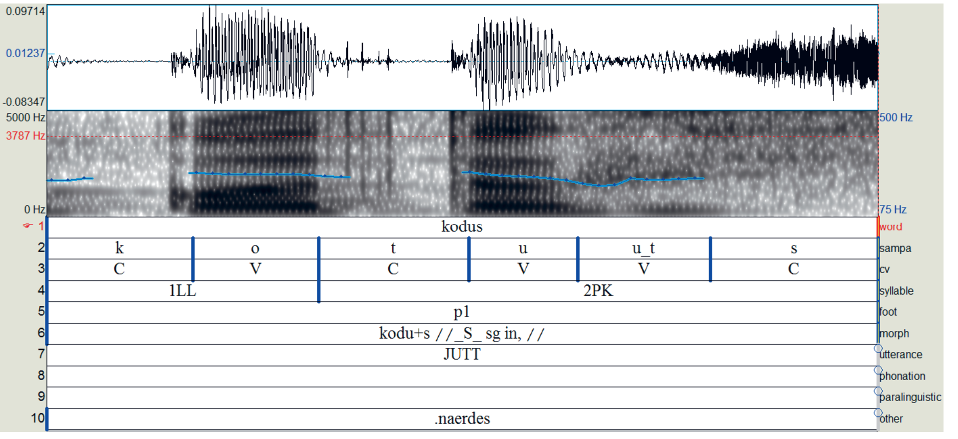

Figure 1. Annotation tiers in PCESS. Word *kodus* [kotus] 'at home'. (P1: 32)

The material for P1 was gathered using the public search engine of the PCESS[5]. The search engine enables search from different layers of the text grid. The search returns a search token with 2 seconds of context around it (sound file and text grid annotation, see Fig. 1). No metadata about the speaker is added. The data set was compiled of frequent content words (substantives, verbs, adjectives, numerals, ordinals) that consisted of a short intervocalic plosive on the boundary of the 1st and the 2nd syllable (sequence VCV, C marks the short plosive, V marks the short vowel). The words were chosen when lemmas representing them occurred at least 50 times in the lemma list of PCESS (Lippus 2019). The final set of words was under 50 for some lemmas because of the stem changes in case paradigms where intervocalic position (e.g., *sõber* : *sõbra* 'a friend nom.sg : gen.sg') or the plosive token was lost (e.g., *sada* : *saja* 'a hundred nom.sg : gen.sg'). The data set contained 1491 tokens from 19 lemmas, as presented in Table 2.

Table 2. Analysed lemmas in P1 (adapted from Ermus (2017: 33))

Plosive							Total
/p/	*luba* [lupɑ]	*paber* [paper]	*sobi* [sopi]	*sõber* [sɤper]	*tuba* [tupɑ]	*vaba* [vapɑ]	
Tokens	36	23	34	14	26	39	172
/t/	*kodu* [kotu]	*kadu* [katu]	*nädal* [nætal]	*pida* [pitɑ]	*pidi* [piti]	*sada* [satɑ]	
Tokens	108	26	134	80	95	45	488
/k/	*luge* [luke]	*lugu* [luku]	*maga* [makɑ]	*nägi* [næki]	*tege* [teke]	*tegi* [teki]	
Tokens	127	55	60	90	235	264	831
Total							**1491**

5 https://foneetikakorpus.ut.ee/

The data set for P2 was taken from SSTC. The material comprised short news stories read by two professional news readers, one male and one female. The data set consisted of short plosive tokens in the following sequences: VCV, SCV, VCS, SCS, #CV, VC#, #CS, SC#, where C marks a short plosive, V marks a vowel, S marks a voiced consonant and # marks a pause. 512 tokens (262 from the male and 250 from the female speaker) were analysed. The distribution of the tokens is presented in Table 3.

Table 3. Distribution of analysed tokens by the position in the word in P2. Abbreviations used: WIS – word-initial singleton, WMS – word-medial singleton, WFS – word-final singleton.

Plosive	WIS	WMS	WFS	Total
/p/	51	24	4	79
/t/	62	109	57	228
/k/	104	97	4	205
Total	217	230	65	**512**

The data set for P3 was gathered from the monologue sub-corpus of PCESS for the semi-spontaneous spoken data and from SSTC for the read-aloud data. The data set consisted of word-medial intervocalic geminate plosives on the boundary of the 1st and the 2nd syllable (sequences VCCV, VC:CV, VVCCV, VVC:CV, where V marks a short and VV a long vowel, CC a short geminate in Q2 words, and C:C a long geminate in Q3 words). Material was obtained from the speech of five male and seven female speakers from PCESS for semi-spontaneous speech and four male and three female speakers from SSTC for read speech. The data set contained 1795 tokens: 511 tokens of /k/, 374 tokens of /p/ and 910 tokens of /t/. The distribution is shown in Table 4.

Table 4. The distribution of analysed tokens in P3.

Plosive	Q2	Q3	Total
/p/	205	169	374
/t/	283	627	910
/k/	177	334	511
Total	665	1130	**1795**

The material for P4 was gathered from the monologue sub-corpus of PCESS. The material from the speech of three male and three female speakers was used. The data set consisted of 2346 tokens of intervocalic plosives with burst in word-initial and word-medial in three quantities (sequences V|CV(V), VCV, VCCV, VC:CV, VVCV, VV:CV VVCCV, VVC:CV, where C marks singleton plosive, CC short geminate plosive, C:C long geminate plosive, V marks a short vowel, VV a long vowel, VV: an overlong vowel and | marks the word boundary). The distribution of the tokens in the data set is shown in Table 5.

Table 5. The distribution of analysed tokens in P4. Abbreviations used: WIS – word-initial singleton, WMS – word-medial singleton, SG – short geminate, LG – long geminate.

Plosive	WIS	WMS	SG	LG	Total
/p/	109	32	62	39	242
/t/	187	281	473	204	1145
/k/	356	305	171	127	959
Total	652	637	709	369	**2346**

3.2. Annotation and acoustic analyses

In addition to the existing corpus annotation, voiced and voiceless parts of the closure and the burst phase or the phase of constriction in the case of change in the articulation manner were manually annotated based on auditory and visual inspection for all studies (see Fig 2–8 in Chapter 4.1). The closure phase was divided into voiced and voiceless parts rather than implosion and occlusion phases. The reason for this was that it is a common procedure in acoustic studies (e.g., Kuzla & Ernestus 2011; Davidson 2018). Also, the author of this thesis finds that it is not possible, especially in the case of singletons, to reliably distinguish between implosion and occlusion phases based just on voicing.

Annotation labels and durational data were extracted with Praat scripts. For P4, along with durational measures, several acoustic measures from the burst phases (intensity, centre of gravity, standard deviation, skewness) were extracted with a script (based on Brato (2015)). A Hann band pass filter between 500 Hz and 10000 Hz (smooth 30 Hz) was applied. Prior to spectral measures, the signal was pre-emphasised above 500 Hz. A 10 ms Hamming window centred at the beginning of the burst phase was analysed to calculate spectral moments (centre of gravity, standard deviation, and skewness). The mean burst intensity (dB SPL) was measured 10 ms from the onset of the burst, and the minimum pitch for intensity measures was set to 500 Hz. Relative burst intensity was calculated by subtracting the mean intensity obtained from the first 50% of the vowel following the burst from the mean intensity obtained from the first 10 ms of the plosive burst.

3.3. Statistical modelling

Statistical analyses in P1 were conducted using the Statistics module in MS Excel. ANOVA tests were used to assess the statistical significance of the differences in the durations of the allophones according to the plosive, preceding and following vowel, stress on the plosive-carrying word in an utterance.

P2 used binary logistic regression models to predict closure voicing and burst reduction in short plosives. Predictors were chosen from the features obtainable from the written text: preceding context (pause, vowel, sonorant consonant), following context (pause, vowel, sonorant consonant), position of the letter in the word (initial, medial, final), number of the letter-carrying syllable in the word from the first (1–4), position of the word in utterance (initial, medial, final), part of speech (11 levels), letter-carrying word being a foreign word (yes/no) or a compound word (yes/no). The models were fit for each plosive separately. The classification ability of the models was evaluated using linear discriminant analysis run on the same data set. Analyses were conducted using SYSTAT software.

Statistical analyses in P3 and P4 were conducted in software R (R Core team 2023) using mixed model functions (*lmer*, *glmer*) from the lme4 package (Bates et al. 2015). The mixed model approach takes into account the random effects in complex, collinear and hierarchical data and is considered suitable for potentially non-balanced data sets, which are common in linguistic corpus material. For post-hoc analyses, Tukey's Honest Significant Difference (TukeyHSD) test and in the P4 emmeans package (estimated marginal means) (Lenth 2022) were used.

In P3, linear mixed models (*lmer*) were fitted for durations of plosive closures and burst phases and generalised mixed models (*glmer*) for the occurrence of the voiced transition in closure. Fixed variables were the place of articulation (/p/, /t/, /k/), quantity of the word (Q2, Q3), allophone (voiceless, partly voiced), syllable structure (preceding vowel short/long), reading style (read, spoken). In *glmer* models fitted for the occurrence of the voiced transition in closure, the duration of the plosive was also included as a fixed variable. The speaker (19 levels) and the words carrying the plosive (858 levels) were included as random intercepts.

In P4, linear mixed models were fitted for durations of the bursts, absolute and relative intensity of the bursts, centre of gravity, standard deviation, and skewness of the burst spectra. Fixed factors were the place of articulation (/p/, /t/, /k/) and plosive category (word-initial singleton, word-medial singleton, short geminate, long geminate). Speaker (six levels) and following vowel (/ɑ e i o u/) were included as random intercepts.

4. RESULTS AND DISCUSSION

In this chapter, the results of the publications are discussed according to the research questions as follows:

RQ1: What kind of allophonic variants emerge in plosives in Estonian connected speech?

RQ2: How do segmental and suprasegmental features influence the allophonic variation of plosives?

RQ3: How does the speech register influence the allophonic variation of plosives?

RQ4: How do segmental and suprasegmental features (syllable structure, stress, quantity) influence the durational variation of plosives?

RQ5: How do segmental and suprasegmental features influence the acoustics of the burst phase?

4.1. Allophonic distribution

The results of this chapter concern the first three research question. First, general allophonic variants emerging in P1, P2 and P3 is presented, then allophonic variation in singletons in spontaneous (P1) and read speech (P2) is described comparatively, followed by allophonic variation in geminates (P3). Influences of segmental features (place of articulation of the plosive, adjacent segments) and suprasegmental features (syllable structure, stress) are discussed for singleton plosives.

4.1.1. Allophonic variants emerging in plosives in Estonian connected speech

Results revealed two types of allophones: **full** – allophones characterised by the burst phase and **reduced** – allophones without the burst phase. However, reduced allophones only occurred in singleton plosives.

Full allophones found are as follows:

1. **voiceless allophone** – canonical fully realised allophone constituting of voiceless closure phase followed by burst phase (see Fig 2);
2. **partly voiced allophone** – closure phase with a voiced transition at the beginning followed by burst phase, the **bleed** type according to Davidson (2016; 2018) (see Fig 3);
3. **voiced allophone with a burst** – fully voiced closure phase followed by burst phase, the **through** type according to Davidson (2016; 2018) (see Fig 4).

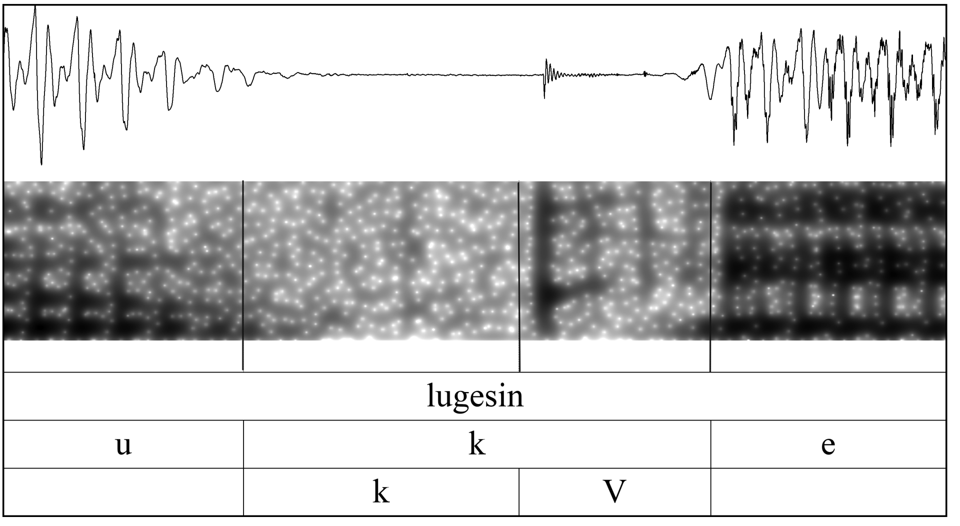

Figure 2. The voiceless allophone of singleton /k/ in the word *lugesin* ‘I read’ past simple (spontaneous speech) in SAMPA annotation. Acoustic phases on the lowest tier: ‘k’ voiceless part of the closure; ‘V’ the burst phase.

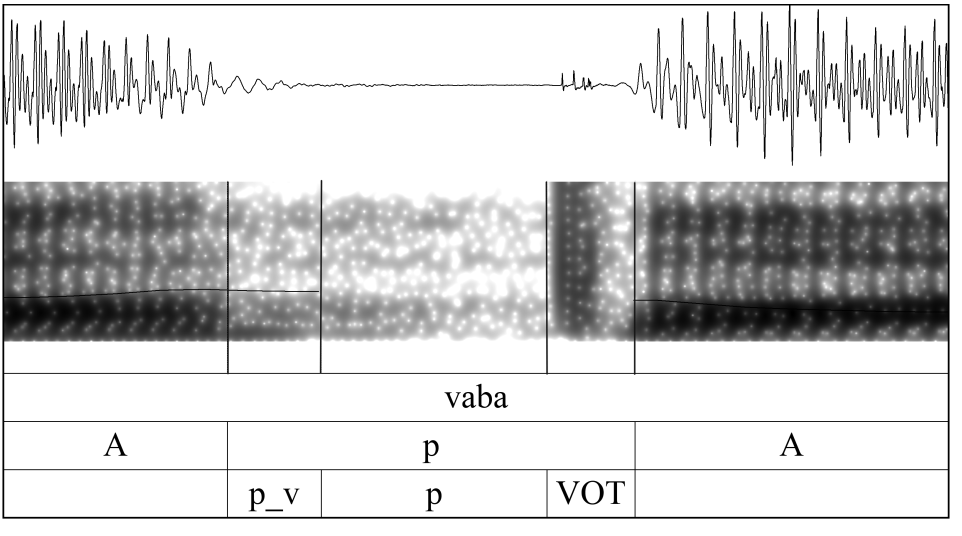

Figure 3. Partly voiced allophone of singleton /p/ in the word *vaba* ‘free’ (spontaneous speech) in SAMPA-annotation. Acoustic phases on the lowest tier: ‘p_v’ voiced part of the closure; ‘p’ voiceless part of the closure; ‘VOT’ the burst phase.

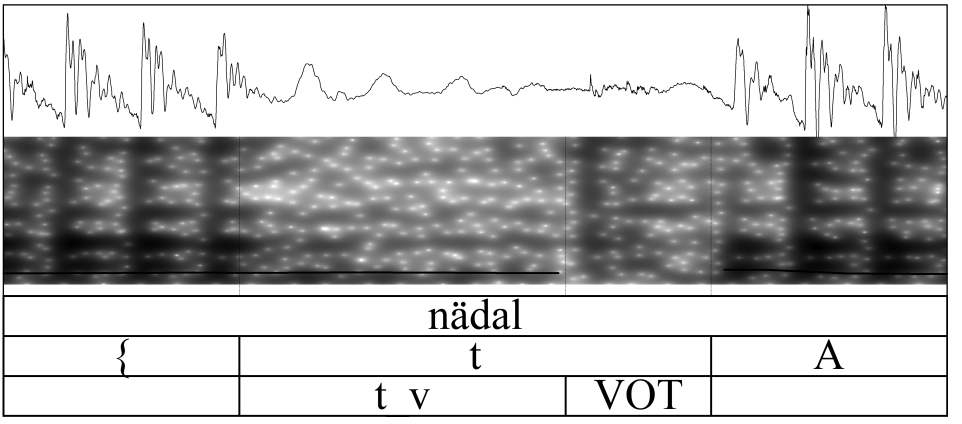

Figure 4. Voiced allophone of singleton /t/ in the word *nädal* ‘a week’ (spontaneous speech) in SAMPA annotation. Acoustic phases on the lowest tier: ‘t_v’ voiced closure, ‘VOT’ the burst phase. (from P1: 35).

The reduced allophones found are as follows:

1. **burstless voiced allophone** – fully voiced closure phase that is still distinguishable in the sound wave and spectrogram, burst phase has reduced (see Fig 5);
2. **fricative allophone** – either the closure phase or the burst has fricativised so that it is distinguishable in the sound wave and spectrogram (see Fig 6);
3. **approximant allophone** – burst has reduced totally; closure phase has reduced to approximant (see Fig 7);
4. **total loss** – the plosive segment is not distinguishable in the speech wave (see Fig 8).

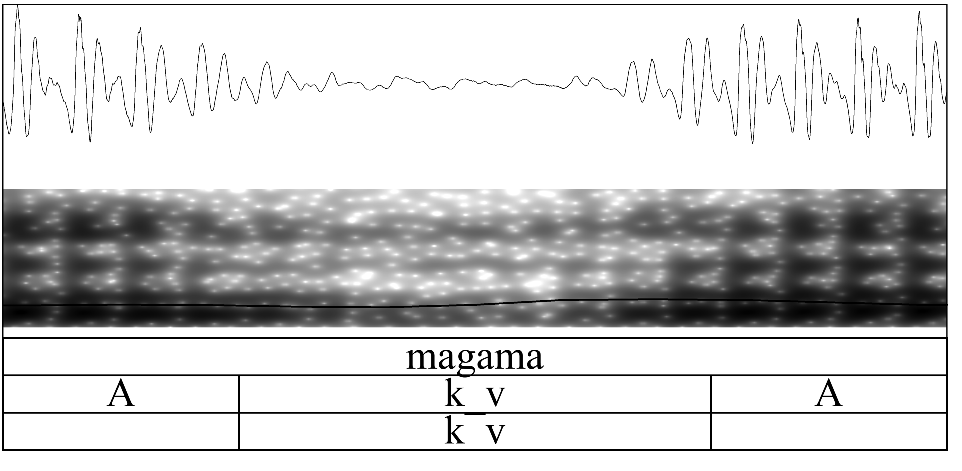

Figure 5. Burstless voiced allophone of singleton /k/ in the word *magama* ‘sleep, ma-infinitive’ (spontaneous speech) in SAMPA annotation. Acoustic phases on the lowest tier: ‘k_v’ voiced closure (From P1: 35).

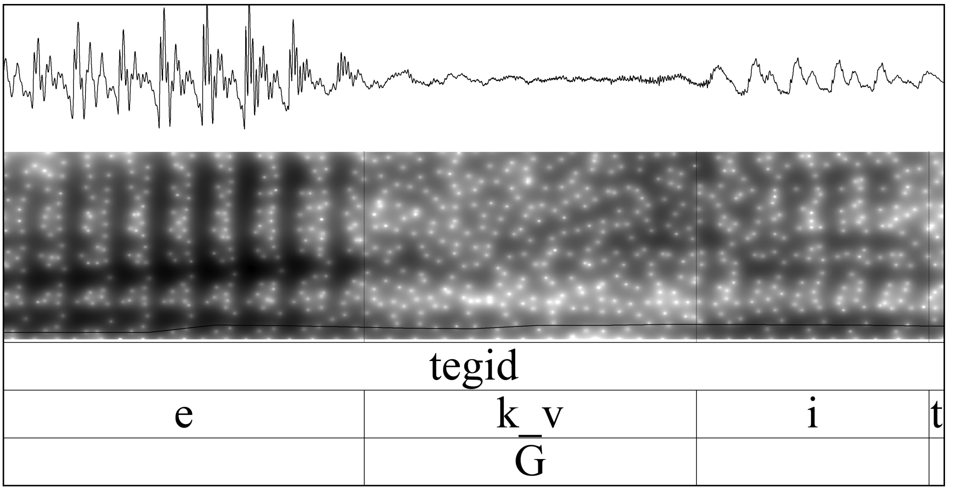

Figure 6. Fricative allophone of singleton /k/ in the word *tegid* ‘you did’ (spontaneous speech) in SAMPA annotation. Acoustic phases on the lowest tier: ‘G’ voiced velar fricative.

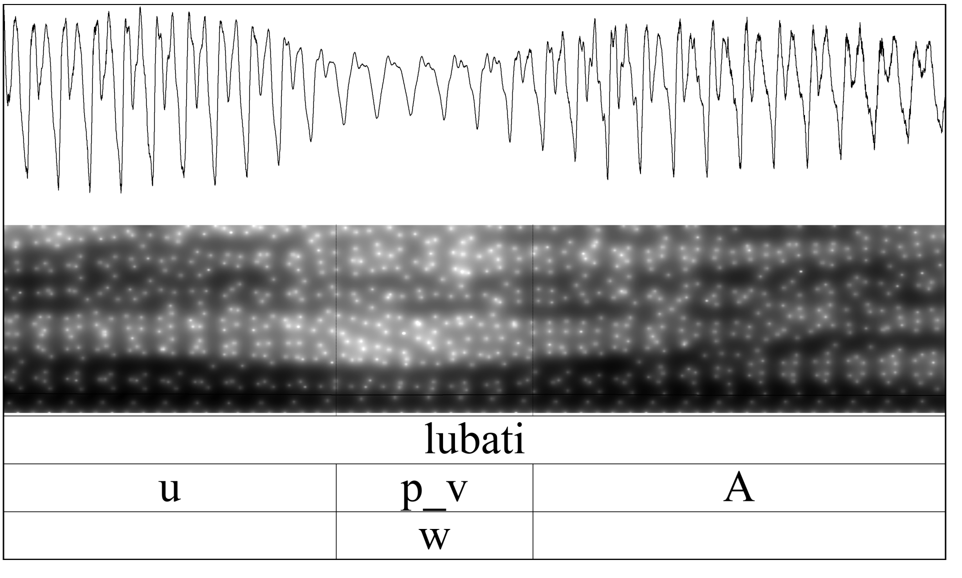

Figure 7. Approximant allophone of singleton /p/ in the word *lubati* ‘to let, past passive voice’ (spontaneous speech). Acoustic phases on the lowest tier: ‘w’ bilabial approximant.

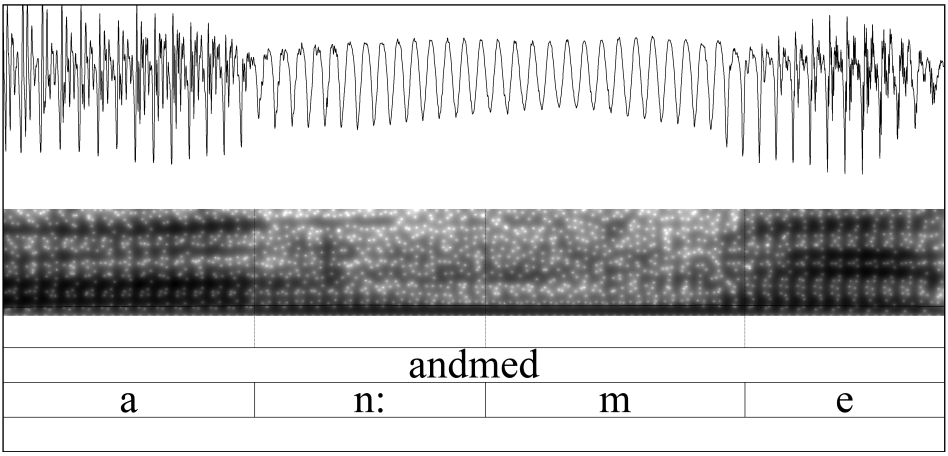

Figure 8. Total loss of singleton /t/ in the word *andmed* 'the data, nom.pl' (read speech).

Additionally, some occasional allophones occurred only a few times each. Those included, for example, nasalised tokens, voiceless closures without bursts and others. The following chapter presents these allophones in the category 'Other'.

4.1.2. Allophonic variation in singleton plosives

Most of the allophonic variation occurred in singleton plosives. For short word-medial plosives, a comparison can be made between two speech registers, spontaneous (P1) and read speech[6] (P2).

The allophonic distribution of word-medial singleton plosives according to POA in read and spontaneous speech is given in Figure 9. The most frequent allophones in both registers were voiceless, partly voiced, voiced, and burstless voiced. Fricatives and total loss occurred in read speech in small amounts, and other reduced allophones occurred only in spontaneous speech.

[6] In P2, word-medial tokens were not discussed separately, as tokens from all the positions in a word were pooled together in analyses. Here, for the sake of balance in data sets, only word-medial tokens are taken into account in comparison.

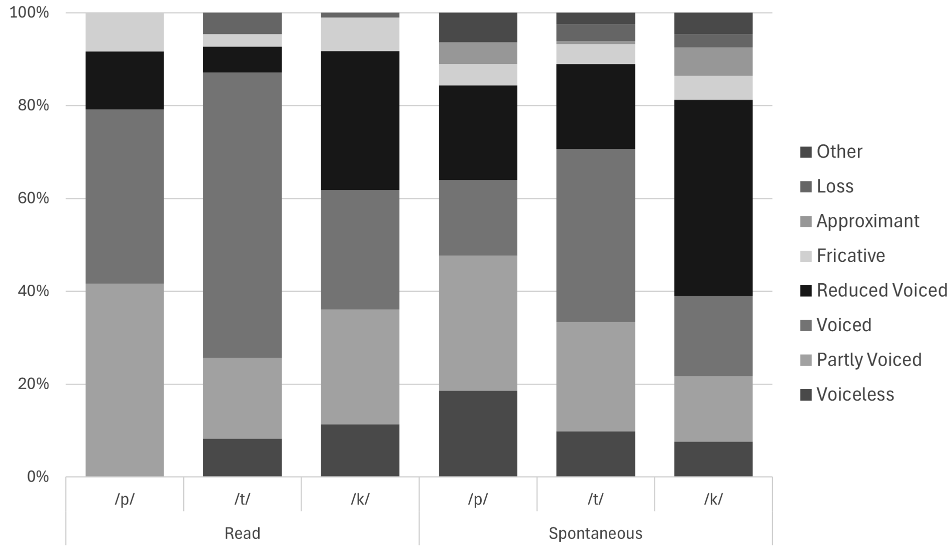

Figure 9. Allophonic distribution of word-medial singleton plosives in read (P2) and spontaneous speech (P1).

Full allophones constituted the majority in **read speech** for all POA – about 80% of tokens of /p/, 87% of /t/ and 62% of /k/ represented allophones with burst. Voiceless allophone occurred in 8% of /t/ and 11% of /k/ but was absent in /p/. Partly voiced and voiced allophones occurred about equal amounts in /p/ (42% and 38%) and /k/ (25% and 26%). In /t/, the voiced allophone was the most frequent, with 61% of all occurrences, but partly voiced allophone constituted only 17%. From reduced allophones, the burstless voiced allophone was the most frequent allophone and comprised 12% of /p/, 5.5% of /t/, but 30% of /k/ (the most frequent allophone). Fricative allophone occurred in 8% of /p/, 3% of /t/ and 7% of /k/. The total loss occurred in 5% of tokens of /t/ and 1% of /k/.

In **spontaneous speech**, full allophones were altogether more frequent than reduced ones in /p/ (64%) and /t/ (71%), but in /k/, they constituted under half of all tokens (41%). Voiceless allophone occurred about the same amount as in read speech, around 10%, in /t/ and /k/. The voiceless allophone of /p/ occurred in 19% of tokens. Partly voiced allophone occurred less in spontaneous than in read speech for /p/ (29%) and /k/ (14%) but more in /t/ (24%). Voiced allophone occurred less in spontaneous than in read speech; 16% of tokens of /p/, 37% of /t/ and 17% of /k/ were voiced allophones. Reduced allophones, expectedly, occurred more in spontaneous speech. Burstless voiced allophone was the most frequent reduced allophone, constituting about 20% of /p/ and /t/ and 42% of /k/. Around 4–5% of the tokens were fricativised in all POAs. 5% of /p/, 6% of /k/ and under 1% of /t/ were reduced to an approximant. Total loss occurred only in /t/ and /k/ (4% and 3%). Other allophones constituted about 2.5% of /t/, 5% of /k/ and over 6% of /p/.

P2 briefly covered allophonic variation of word-initial and word-final singleton plosives in read speech. Word-initial singletons were mainly realised: 37%

were voiceless, around 30% were partly voiced and around 20% voiced. Reduced allophones constituted 13% of all word-initial tokens. Many voiceless tokens were utterance-initial, but they occurred in all positions in utterance. About 40% of the word-final singletons were voiced, voiceless and partly voiced tokens both constituted about a quarter of tokens. Only under 10% of tokens were reduced allophones.

In the main part, the differences between the two registers were as expected: more fully realised tokens in read speech and more reduced tokens in spontaneous speech. The distribution of allophones between POAs was similar in both registers. Deviations from the expected patterns were the absence of a voiceless allophone of /p/ in read speech and a more frequent total loss in /t/ in read speech than in spontaneous speech (although the difference was small). The absence of voiceless allophone on /p/ in read speech can be explained by the /p/ tokens in read speech frequently occurring in non-content words (like *läbi* 'through') and possibly idiolect of the readers (there were also cases of utterance-initial voiced tokens in the speech of one reader). Total loss in /t/ tokens in read speech occurred in words where the plosive token was located between nasals (e.g. in word *andmed* 'data').

The effect of adjacent segments on the allophonic distribution of singleton plosives was studied in P1. In spontaneous speech, the vowel context had a minor influence on the allophonic distribution. Different POAs occurred in slightly different environments (see Fig 10, 11). Among the six preceding vowels (V1), only /ɑ/ occurred adjacent to all the POAs; /e/ and /u/ occurred adjacent to /p/ and /k/; /æ/ adjacent to /t/ and /k/; /i/ and /o/ occurred only adjacent to /t/. Four vowels occurred in the position following the plosives (V2). Both /ɑ/ and /i/ occurred after all POAs; /e/ occurred after /p/ and /k/; /u/ occurred after /t/ and /k/.

According to V1, allophones were rather even for all POAs; the main differences were in the proportion of reduced allophones. There were more reduced tokens, especially fricatives, in /p/ when V1 was /e/ and more reduced tokens in /k/ when V1 was /æ/. When V1 was labial vowel /u/, both /p/ and /k/ showed more approximant allophones. For /t/, there were fewer reduced tokens when V1 was labial vowel /o/. V2 seemed to have a bigger influence. For /p/, there were less reduced allophones when V2 was /ɑ/ compared to /e/ and /i/; there were also larger amounts of fricative when V2 was /i/. For /t/, there were fewer reduced tokens when V2 was labial /u/ and more than when V2 was /a/. For /k/, there were more approximant allophone when V2 was /u/. There tended to be more fricative allophone adjacent to front vowels in all POAs and more approximant allophone adjacent to labial vowels for /p/ and /k/.

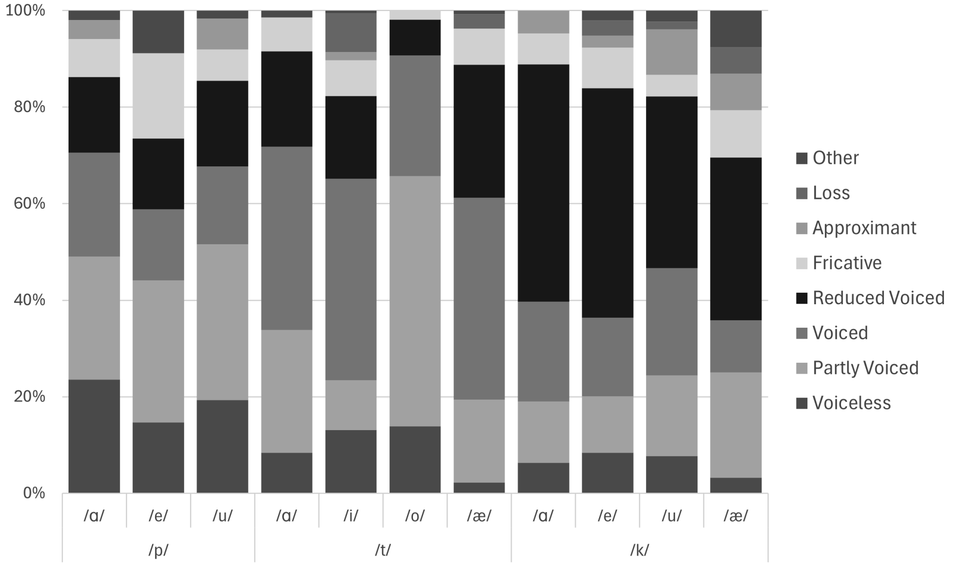

Figure 10. The allophonic distribution of intervocalic singleton plosives in P1 according to V1.

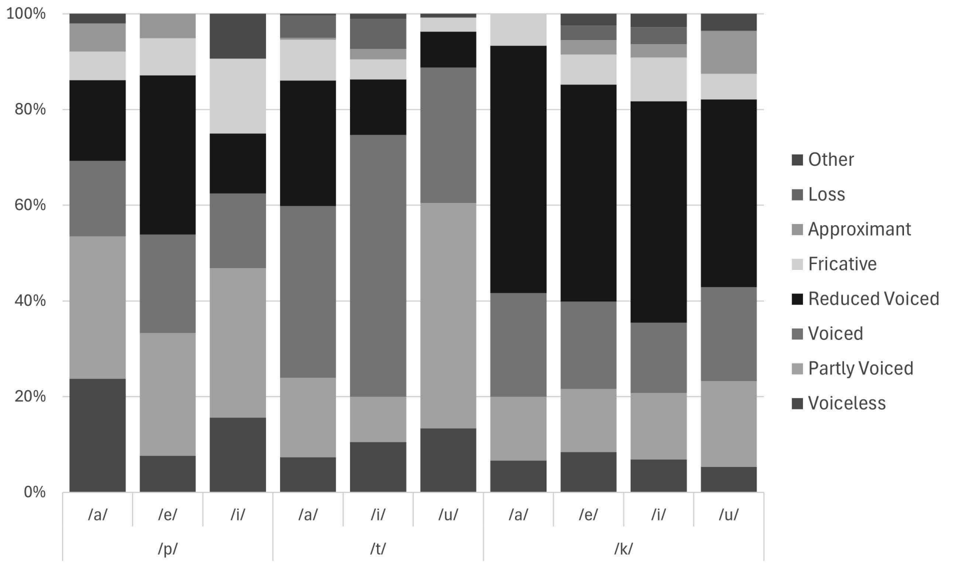

Figure 11. The allophonic distribution of intervocalic singleton plosives in P1 according to V2.

The influence of sentence stress was studied in P1. There were few cases with contrastive stress, and in many cases, there was not enough context to decide if the stress on the word was contrastive because the output of the search engine used for gathering data only gives 2 seconds of context; all the stressed words were put into one category.

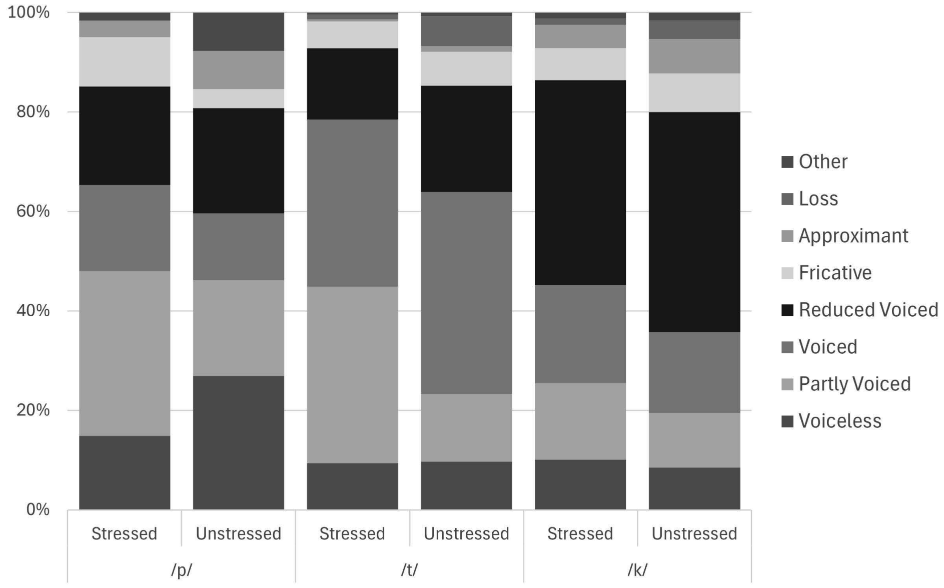

Figure 12. The allophonic distribution of intervocalic singleton plosives in P1 in stressed and unstressed positions (after P1: 40).

A stressed position somewhat influenced the allophonic distribution (see Fig 12). There tended to be more full allophones in the stressed position and more reduced allophones in the unstressed position. The difference was also visible in the proportion of non-regular allophones (fricatives, semivowels, total loss, etc.), which occurred much more in the unstressed position. Between POAs, /t/ showed the most significant difference between the two contexts. The proportion of the voiceless allophone was around the same in both positions for /t/ and /k/ but unexpectedly larger in the unstressed position for /p/. Partly voiced allophone occurred more in a stressed position for all POAs.

4.1.3. Allophonic distribution of geminate plosives

P3 studied the allophonic distribution of geminate plosives. Two types of allophones emerged, both fully realised: voiceless and partly voiced (bleed-type). The distribution of allophones is given in Figure 13.

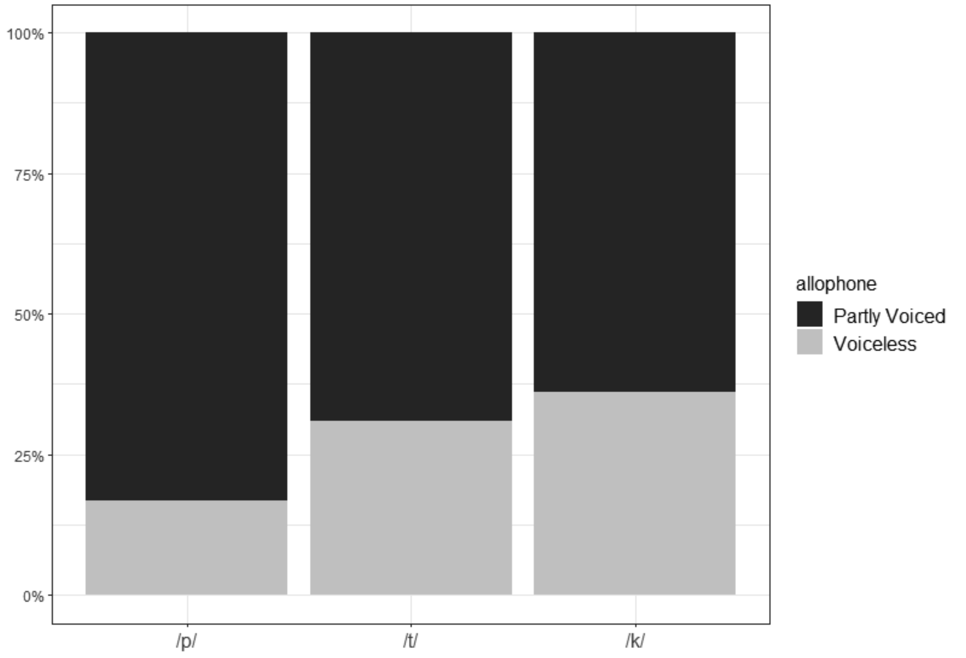

Figure 13. Allophonic distribution of inter-vocalic geminate plosives (after P3:43)

Voiceless allophone constituted the minority for all plosives. The voiceless tokens constituted almost 40% of /k/, about 30% of /t/, and about 20% of /p/. The distribution of geminate allophones was close to that of voiceless and partly voiced allophones in singletons in P1. Partly voiced allophone occurred more in geminate /p/ than other POAs in P3.

4.1.4. Discussion

Concerning allophonic variation in Estonian plosives, I hypothesised that a large allophonic variation occurs that is influenced by POA, adjacent vowel context, sentence stress, and speech register. All the hypotheses were at least weakly corroborated.

The allophonic variation was larger in singletons than in geminates and in spontaneous speech than in read speech. The overall reduction rates in spontaneous speech were less significant in P1 (/p/ 36%, /t/ 29%, /k/ 59%) compared to Raasik (2010) (/p/ 42%, /t/ 57%, /k/ 66%) where all parts of speech were studied, so the amount of short grammatical words may have pushed up the number of reduced tokens. However, the number of reduced tokens of /k/ was on the same scale in both studies. A large proportion of burstless voiced tokens of /k/ (25.9%) was also found in Estonian read speech (Suomi & Meister 2012) whereas the proportion of reduced allophones in read speech was much smaller for /t/ (3%) and /p/ (1%).

Compared to other languages where reduction has been studied, a large proportion of tokens of reduced allophones in voiceless plosives has been found in conversational Spanish by Torreira and Ernestus (2011), where over 25% of intervocalic tokens observed in spontaneous speech were reduced, and similar to Estonian, /k/ was more reduced than other plosives. In French, a reduction has been found in voiced plosives by Duez (1995), whereas /d/ showed a larger amount of tokens with burst than reduced tokens, /b/ had more burstless tokens, and /g/ had almost the same amount of both. In voiceless plosives in French (Torreira & Ernestus 2011), the number of reduced tokens was much smaller, only 5.4% (of all tokens). Comparing the voicing patterns in Estonian word-medial singletons to word-medial voiced plosives in American English (Davidson 2016), there were fewer fully voiced tokens in American English (about 20%); in American English, the patterns showed a tendency for the voicing to decrease. However, the results in Davidson (Davidson 2016; 2018) included more segmental context than just vowels and did not include tokens of reduced allophones. When looking only at full allophones of singletons in Estonian, the proportion of voiced allophone was around 45% and partly voiced allophone around 18%, so the amount of voiced allophone was still larger in Estonian than in American English. The geminates in Estonian behaved like word-medial voiceless plosives in American English (Davidson 2018), showing more partly voiced than fully voiceless closures in both studies, although the proportion of voiceless tokens was higher in English.

High reduction rates in Estonian singleton plosives can be explained by the short segment durations. By gestural coarticulation models (e.g., Browman & Goldstein 1992), during shorter time intervals articulatory gestures reduce and overlap and it is difficult for articulators to reach articulatory targets. The vibration of the vocal folds may continue through the closure or stop for a time so short that it does not express acoustically. The pressure in the vocal tract may stay too low to form the burst phase. Velar /k/ stood out with the most extensive reduction in P1, P2 and Suomi and Meister (2012). It is logical from the gestural viewpoint, as the dorsal part of the tongue involved in the formation of velar stop is also involved in articulating vowels. Other POAs engage less. More reduced allophones occurred in /t/ and /k/ if the tongue position of the following vowel was further away from the articulation place of the plosive. For example, there were a little more reduced allophones of /k/ when V2 was /i/ and more reduced allophones of /t/ when V2 was /a/.

In geminates, there is more time to articulate the phases of plosive, so the burst phase is retained. The tendency of bilabial /p/ to show more voiced tokens is explainable with aerodynamic reasons. As the supraglottal cavity behind the bilabial closure is large, the pressure in the vocal tract rises for longer after the closing gesture, and vocal cords may resonate for a prolonged time.

A larger amount of reduction in Estonian compared to languages with voicing distinction is in accordance with the notion by Keating, Linker and Huffman (1983) that languages with fewer phonological distinctions allow more variation. As Estonian does not need to retain the perceptual differences based on voicing, there is more room to different reduction patterns. However, in spontaneous speech, great allophonic variety has been found also in languages with voicing distinction, such as French (Duez 1995).

Differences in allophonic distribution between spontaneous and read speech are in accordance with Lindblom (1990). The difference may also at least partly be influenced by speech tempo as well. In P2, the speech rate was 5.1 syl/sec for the male speaker and 4.4 syl/sec for the female speaker, and the speaker turned out to be a significant predictor in models, with the male speaker showing a greater reduction. In P1, the speech rate was not measured, but it is plausible that speech tempo is higher (although more variable as well) in spontaneous than in read speech. In P3, both semi-spontaneous and read speech were used, but the differences in speech tempo between the speakers were more significant than between speech registers.

4.2. Durational variation of plosives

This chapter discusses the answers to RQ3: how do segmental features (place of articulation of the plosive, adjacent segments) and suprasegmental features (syllable structure, stress, quantity) influence the durational variation of plosives? The durational characteristics of plosives were studied in P1 for singletons and in P3 for geminates.

4.2.1. Durational variation in singleton plosives

The durational variation of singleton plosives in spontaneous speech (P1) was large. It varied from 20 ms to 120 ms over all POAs. The durational variation is given in Figure 14. The duration was first influenced by POA and second by the allophone. Overall, /p/ had the longest mean duration (71 ms) over all tokens. Mean durations of /t/ and /k/ over all tokens were close (59 ms). It was probably due to the different allophonic distributions: a large part of tokens of /k/ represented reduced allophones. The durational pattern of most frequent allophones was the same for all POAs: partly voiced > voiceless > voiced > burstless voiced. Durations between allophones were statistically different in all POAs, except between voiceless and partly voiced allophones of /t/ and /k/[7]. Compared between POA, bilabial /p/ showed the longest mean durations for all allophones. The mean durations of /p/ and /k/ were close to each other in voiceless allophone (77 ms and 76 ms) and voiced allophone (67 ms and 66 ms), while /t/ showed systematically shorter mean durations, almost 10 ms shorter than /p/ and /k/ in voiceless (67 ms) and voiced allophone (58 ms). The differences were larger in partly voiced allophone (/p/ 88 ms, /k/ 81 ms, /t/ 74 ms) and burstless voiced allophone (/p/ 54 ms, /k/ 47 ms, /t/ 39 ms), but the pattern persisted. By allophones, the durational differences between all POAs were significant in partly voiced and

[7] For some odd reason, the statistical tests for differences in allophone durations of one plosive were not reported in P1.

burstless voiced allophones, /t/ differed from other POAs also in voiceless and voiced allophones.

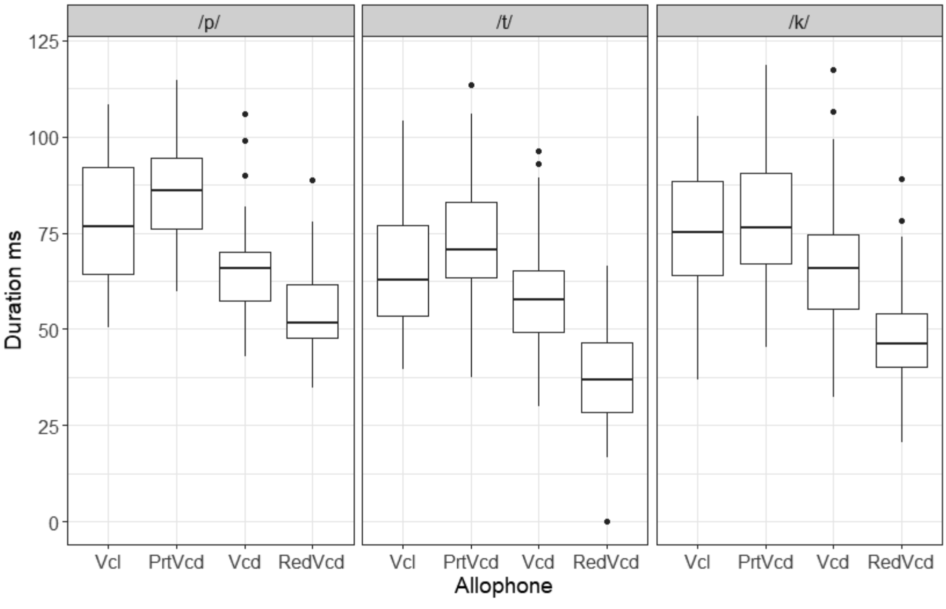

Figure 14. Durational variation of main allophones of intervocalic singleton plosives in P1. Abbreviations: Vcl – voiceless, PrtVcd – partly voiced, Vcd – voiced with burst, RedVcd – burstless voiced.

The vowel context showed some influence on the plosive duration in singletons in P1. Voiceless and partly voiced allophones showed more regular patterns; the durations of voiced allophones were often very close in different vowel contexts. From POAs, /t/ seemed to be the most influenced by adjacent vowel context in durational variation and the only one that showed statistically significant differences in the durations in different contexts. Duration differences according to the V2 (/i u a/) were significantly different in voiceless and partly voiced allophones of /t/, with /i/ showing the longest mean duration. When looking for the influence of the V1 on closure duration, then again, over all occurrences, /t/ showed the strongest influences: closure following back vowels /ɑ/ and /o/ were longer than closures following front vowels /i/ and /æ/. Other POAs did not show this influence. There was a tendency for longer mean durations near labial vowels in all POAs, with V1 /o/ in /p/ and /t/ and /u/ in /k/. **The sentence stress** had no significant effects on the durations of the same allophones of singleton plosives.

4.2.2. Durational variation in geminate plosives

The durational variation in geminate plosives (P3) is given in Figure 15. Overall variation was large, like in singletons, from around 50 ms in the shortest Q2 tokens to around 300 ms in the longest Q3 tokens. The durational pattern between POAs in Q3 tokens was similar to the pattern in singletons: /p/ (156 ms) > /t/ (143 ms) > /k/ (139 ms). In Q2 tokens, the mean durations of all POAs were very close, the ones of /t/ and /k/ being similar (119 ms) and /p/ being a few ms longer (122 ms). Partly voiced tokens were generally shorter than voiced tokens in /t/ and /k/. The duration was correlated with the syllable structure: plosive tokens were shorter after long vowels and diphthongs than after short vowels.

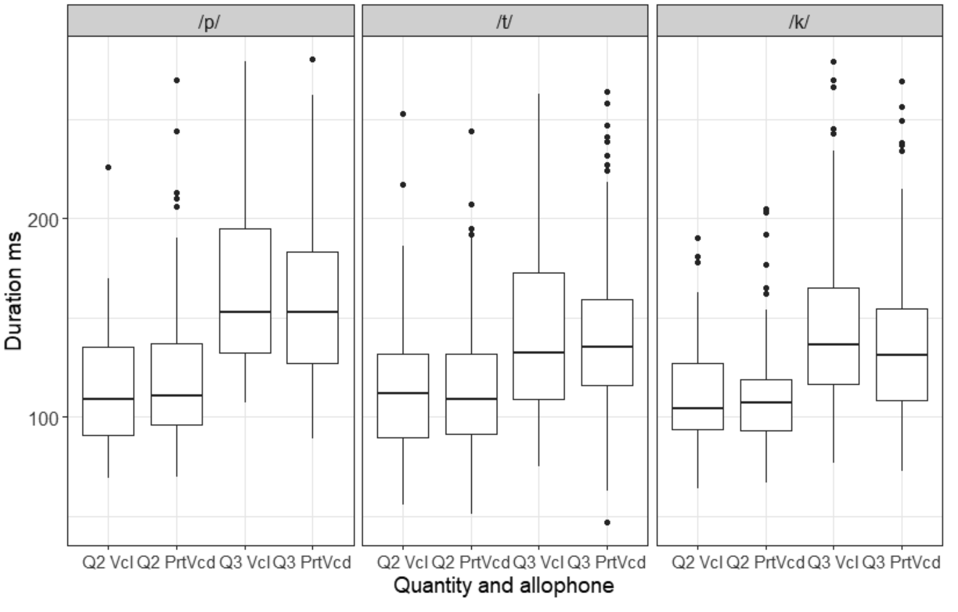

Figure 15. Durational variation of main allophones of intervocalic geminate plosives in P3. Abbreviations: Q2 – the second quantity degree, Q3 – the third quantity degree, Vcl – voiceless, PrtVcd – partly voiced.

In P3, the duration of the voiced transition at the beginning of the closure phase of the partly voiced allophone of geminates and the proportion from the closure were measured. The durational variation of the voiced transition is given in Figure 16. The variation was large, but most transitions lasted 10–30 ms. The mean duration of the transition formed a pattern /k/ (18 ms) < /t/ (21 ms) < /p/ (25 ms). The duration of the voiced transition decreased with the longer duration of the token and when the token was in Q3.

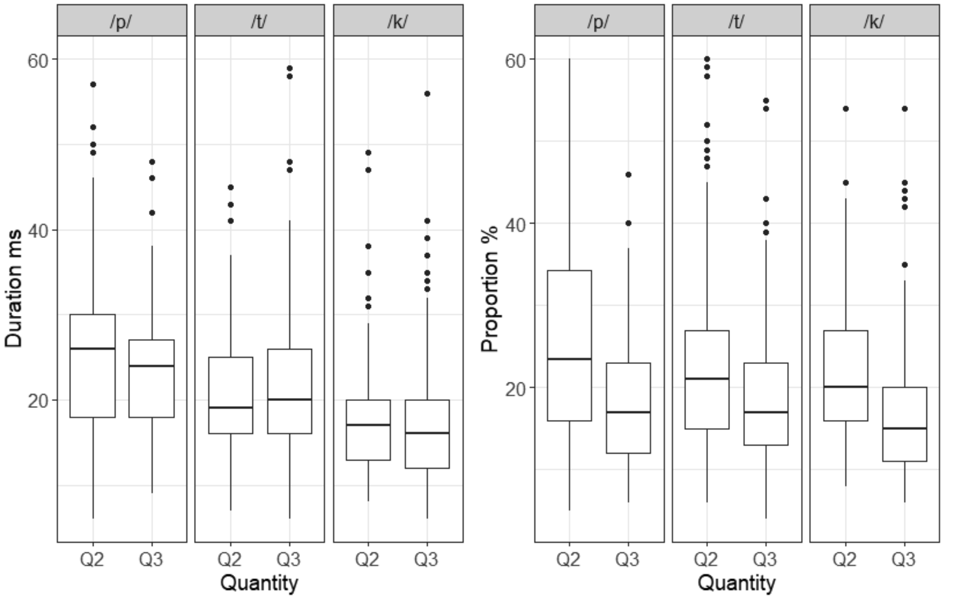

Figure 16. Variation of duration and proportion of the voiced transition of the closure phase in geminate plosives (P3).

The mean proportion of the voiced transition constituted about 22–27% in Q2 and 17–19% in Q3. Compared between POAs, the proportion of the voiced transition in /k/ was significantly smaller than in other POAs. The proportion of voiced transition decreased in Q3 compared to Q2 and in tokens with a longer duration.

4.2.3. Discussion

Concerning the durational variation of Estonian plosives, I hypothesised that the durational patterns according to POA and allophone are similar between singletons and geminates; there are significant differences between the durations between quantity degrees; the duration of geminates is shorter when the preceding vowel is long; the durational variation in singleton plosives is influenced by adjacent vowel context but not sentence stress.

The duration of plosives was first influenced by quantity, second by POA, and third by allophone. There were differences in durational patterns between singletons and geminates. The only similarity was the overall duration pattern according to POA: /p/ > /k/ > /t/, but the durational relations differed slightly. In the case of single plosives, /p/ and /k/ were close in duration and /t/ was remarkably shorter. In geminates, /p/ was the longest, while /k/ and /t/ were relatively close in duration. Comparing allophones, partly voiced allophone in singletons lasted longer than the voiceless allophone; in geminates, voiceless allophone showed longer durations. The durational patterns partly followed previous findings in Estonia. The same pattern between POAs in spontaneous speech was found in Türk (2018).

Eek (1974) measured durations in isolated words; the study only included /p/ and /t/, and the mean durations of /p/ were shorter than those of /t/.

As expected, the geminates were longer in Q3 than in Q2. The difference in mean durations of short and long geminates was at least 20 ms for all plosives. Q3 geminates were 1.17–1.27 times longer than Q2 geminates. The duration ratios between geminates in P3 and singletons in P1 were 1.72–2.02 for Q2/Q1 and 2.2–2.42 for Q3/Q1 (see ratios in Table 6). The duration of geminate tokens was correlated with the syllable structure: plosive tokens were shorter after long vowels and diphthongs. The duration ratios of geminates in different syllable structures (VC/VVC) were around 1.1–1.2. The results is close to the ratio in Eek (1974; 1975) (1.1), who found compensatory effects in syllable durations in order to retain the ratios necessary for quantity perception.

Table 6. The durational ratios between quantity degrees in Estonian plosives in different studies.

	Q2/Q1	Q3/Q2	Q3/Q1
/p/			
Eek (1974)	1.31	1.58	2.06
Türk (2018, 2019a)	1.7	1.2	2.1
P1, P3	1.72	1.27	2.2
/t/			
Eek (1974)	2.03	1.42	2.88
Türk (2018, 2019a)	2.1	1.3	2.6
P1, P3	2.02	1.2	2.42
/k/			
Eek (1974)	–	–	–
Türk (2018, 2019a)	1.9	1.4	2.6
P1, P3	2.02	1.17	2.36

Durational ratios between quantity degrees of Estonian plosives have been measured before by Eek (1974) in read speech and Türk (2018; 2019a) in spontaneous speech. The results in Türk (2018; 2019a) are close to P3: /p/ 1.2, /t/ 1.3, /k/ 1.4, the biggest difference in /k/. The results in Eek (1974) diverge: /p/ 1.58, /t/ 1.42 (/k/ was not measured). It must be noted that Eek (1974) used material only from one speaker so that idiolect may have played a role.

Markus et al. (2013) have calculated the ratios between the durations of Q2 and Q3 geminates in several Finno-Ugric languages with similar three-way quantity systems: Estonian, Livonian, Ingrian and Inari Saami. There, the Q3/Q2 ratio in Estonian (calculated on alveolar /t/ and /l/) was 1.4. The closest was Ingrian (1.3), then Inari Saami (1.52), and the durational ratio was the largest in Livonian (1.78). The durational ratios in Saami languages have also been measured by Türk et al. (2019); the ratios varied between 1.2 and 1.5. However, in most of the studied languages (except Ingrian), the ratios were calculated over more consonant types than just plosives.

When compared to languages with two-way quantity systems, the durational ratios were larger than the ratios of Q3/Q1, being around 3 in Japanese (Homma 1981) and Swiss German (Kraehenmann 2001), or near the ratios Q2/Q1, being around 2 in Italian (Payne 2005) and Maltese (Galea 2016). The exception was Cypriot Greek (Arvaniti & Tserdanelis 2000) where the geminate-singleton duration ratio was 1.6. However, in Cypriot Greek, the additional cue to closure duration was longer VOT in geminates.

The voiced transition at the beginning of the closure phase was the shortest in /k/ and the longest in /p/. Differences are explainable with aerodynamics during plosive articulation (Ohala 1983). The room behind the closure is small in the velar, so pressure build-up and glottal pulsing ceasing are fast. The room in the mouth is bigger during bilabial closure, and it takes more time for the pressure to increase to the point of voicelessness even after the implosion gesture has been finished. Therefore, the voicing also precedes longer.

The proportions of the voiced transition in P3 (22–27% in Q2 and 17–19% in Q3) are similar to Ariste (1933a) who measured closure voicing in word-medial geminates in isolated words. The voiced part constituted about 10–25% of the plosive duration. The proportions of the voiced transition were larger in P3 compared to Ariste (1933a) because, in P3, it was calculated from only the closure phase, not from the total duration of plosive. A similar tendency of /k/ to be less voiced than other POAs in geminates has been found, for example, in Tashlhiyt Berber voiced geminate plosives (Ridouane 2007).

The vowel context showed a minor influence on the plosive duration in singletons in P1. There were no strong patterns over allophones, and no clear patterns formed either in /p/ or /k/. Alveolar /t/ was the most influenced (as it was most influenced in allophonic variation), showing longest durations when V2 was /i/ or V1 was /ɑ/ or /o/. There was a tendency for longer mean durations near labial vowels in all POAs. This finding repeats the one in Raasik (2010).

The result for /t/ is partly in line with Türk (2019a), who studied the influence of adjacent vowels' height on the consonants' duration. She similarly found that consonants tended to be longer in duration when V2 was high vowel, which is in line with P1. She found the same effect also when V1 was high. In P1, the difference in durations according to V1 was based rather on frontness than height and front vowels /i/ and /æ/ were connected to shorter durations. The difference is probably because Türk (2019a) only used vowel height as a factor in modelling, not frontness.

The stressed position of the plosive carrying word had no significant effects on the durations of the same allophones of singleton plosives. The finding is in line with Suomi and Meister (2012) who found no differences between non-contrastively stressed and unstressed positions in singletons in Estonian and Kuzla and Ernestus (2011) who found the same in lenis plosives in German.

4.3. Characteristics of the burst phase

This chapter discusses research question 4: How do segmental and suprasegmental features influence the acoustics of the burst phase? The durational characteristics were studied in P1, P3 and P4 and non-durational characteristics were studied in P4.

4.3.1. Durational characteristics of the burst phase

All the studies (P1, P3, P4) confirmed that the main factor influencing the burst phase duration was the POA. The longest burst phase duration occurred in /k/ in all studies (see Figures 17 and 18). The influence of other controlled factors (vowel context, stress, quantity) was small or missing. So, all the hypotheses about durational characteristics were confirmed.

In singleton plosives in spontaneous speech (P1, Fig 17), the mean burst durations of singleton /p/ and /t/ were very close (22 and 23 ms, respectively), and the mean burst duration of /k/ was significantly longer (29 ms). There were slight durational differences between allophones of /t/, where the voiceless allophone showed longer bursts and in /k/, where the partly voiced allophone showed shorter bursts than other allophones. The vowel context influenced the bursts of /t/ and /k/, which showed significantly longer burst durations before /i/. Durational patterns were a little more consistent in voiceless and partly voiced allophones. The burst durations of /p/ were similar in all allophones and vowel contexts. The stressed position of the plosive carrying word showed no influence on the burst durations.

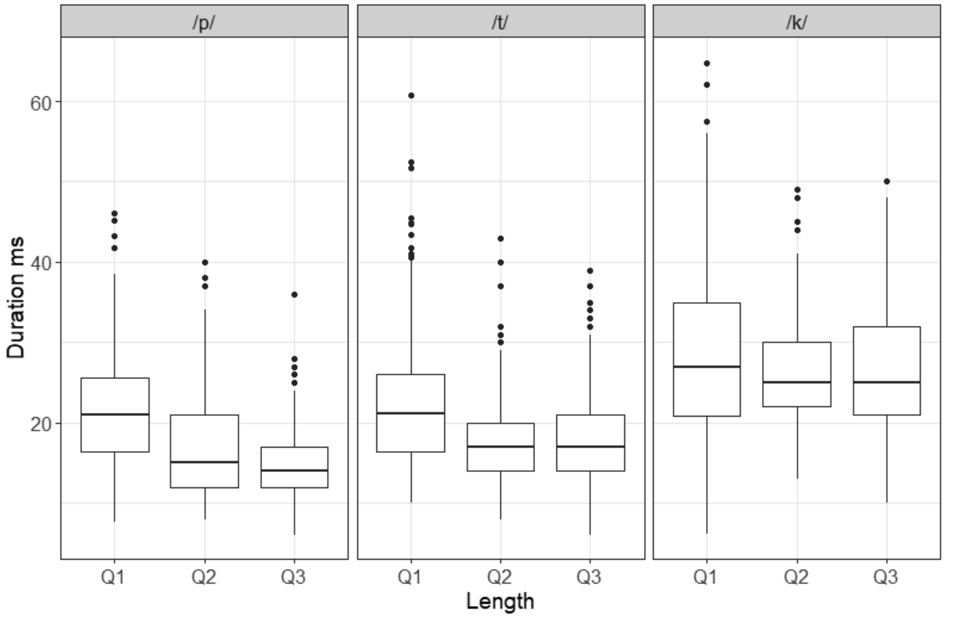

Figure 17. Variation of burst duration (ms) by POA and quantity in singletons – Q1 (P1) and geminates – Q2, Q3 (P3).

All other studies dealt with semi-spontaneous or read speech. VOT of geminates was compared in P3 (Fig 17) and VOT of all the length categories in P4 (Fig 18). There were very small differences in mean VOT between singletons and geminates, Q2 and Q3 geminates or between word-initial and medial singletons. In P3, the mean VOT of /p/ was 17 ms in Q2 and 15 ms in Q3, the mean VOT of /t/ 17 ms and 18 ms, respectively, and VOT of /k/ 26 ms in both length categories. In P4, the mean duration of VOT of /p/ was 17 ms in all length categories, /t/ 18 ms in all categories, /k/ between 23–26 ms, being the shortest in Q1 (differences were more considerable between median durations).

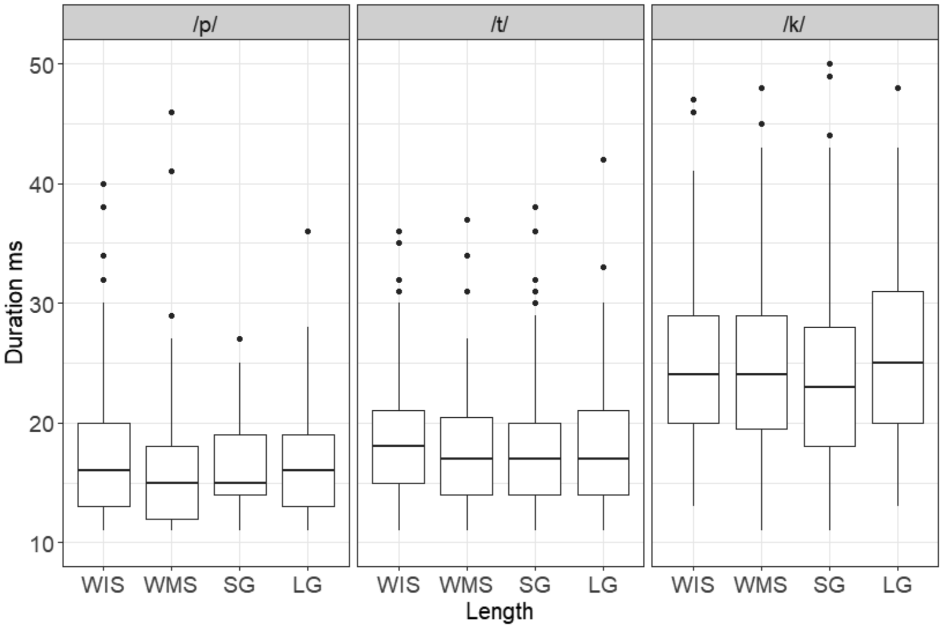

Figure 18. Variation of burst duration by POA and plosive category (from P4: 98). Abbreviations: WIS – word-initial singleton, WMS – word-medial singleton (Q1), SG – short geminate (Q2), LG – long geminate (Q3).

In P3, the most important factor influencing the VOT apart from the POA was the duration of the token. Longer tokens were more likely to have longer VOT. Q3, on the other hand, decreased the VOT. The length of the vowel preceding the plosive token and the allophone did not show any significant effects on the VOT. In P4, there was a slight tendency for longer VOT in geminates according to the mean values, but statistical modelling found no effect. Also, no significant differences emerged between the VOT of word-initial and word-medial plosives, i.e. stressed and unstressed positions.

4.3.2. Non-durational characteristics of plosive bursts

The intensity and spectral measures of plosive bursts were studied in P4. All the categories – word-initial singleton (WIS), word-medial singleton (WMS, Q1), short geminate (SG, Q2) and long geminate (LG, Q3) were included. The variation of non-durational measures of plosive bursts is given in Figure 18.

The burst intensity was measured as absolute and relative (see explanation in Chapter 3.2). Between POAs, the **absolute intensities** formed a pattern /p/ > /t/ > /k/. Mean absolute intensities were close between /p/ and /t/ (around 70 dB in singletons, around 65 dB in geminates); the intensity of /k/ was lower than other POAs in singletons (65–66 dB). In geminates, the differences between POAs were minor (64–67 dB).

The pattern in **relative intensity**, dependent on the intensity difference between the burst and the following vowel, was the opposite: /k/ > /t/ > /p/. Again, /p/ and /t/ showed similar values (3–5 dB), but /k/ showed higher relative intensity than other POAs (5.5–6.9 dB).

Length categories formed different patterns in different POAs, but WMS and SG tended to have lower relative intensities than WIS or LG. However, the differences were statistically significant only for SG. The difference in intensity between word-initial and word-medial singletons was evident in overall variation, with word-initial having greater relative intensity, but the statistical differences failed to confirm the differences.

The pattern emerging in COG values was /t/ > /k/ > /p/. The mean COG values of /t/ were much higher than in other plosives, and differences were especially remarkable in geminates. Two groups formed in /p/ and /t/ where singletons showed lower and geminates higher COG values (means /p/ singletons 1600–1800 Hz, geminates 1900–2200 Hz; /t/ singletons 2000–2100 Hz, geminates 2600–2800 Hz). For /k/, the difference was much smaller (means: singletons 1800–1900 Hz, geminates 2000 Hz).

The measures of standard deviation did not form any clear patterns. The primary influencer was POA, the pattern /t/ > /k/ > /p/. The mean SD values varied between 1220 and 1368 Hz for /p/, between 1662 and 1880 for /t/ and between 1279 and 1562 Hz for /k/. The SD values of singletons were lower than those of geminates, but no significant differences were revealed in statistical modelling.

The skewness values were positive for all POAs, which suggests more intensive lower frequencies in spectra. According to the POA, the pattern was /p/ > /k/ > /t/, whereas the skewness on /t/ was significantly lower than other POAs. There were differences in the skewness between singletons and geminates. Again, /k/ showed the closest mean values between all length categories (2.2–2.6). The differences were more considerable for /p/ and /t/: 3.19–3.34 in singletons and around 2.1 in geminates for /p/; around 1.9 in singletons and 1.25 in geminates for /t/. Differences emerged between word-initial and word-medial positions, whereas the skewness of word-initial position was lower than any word-medial categories.

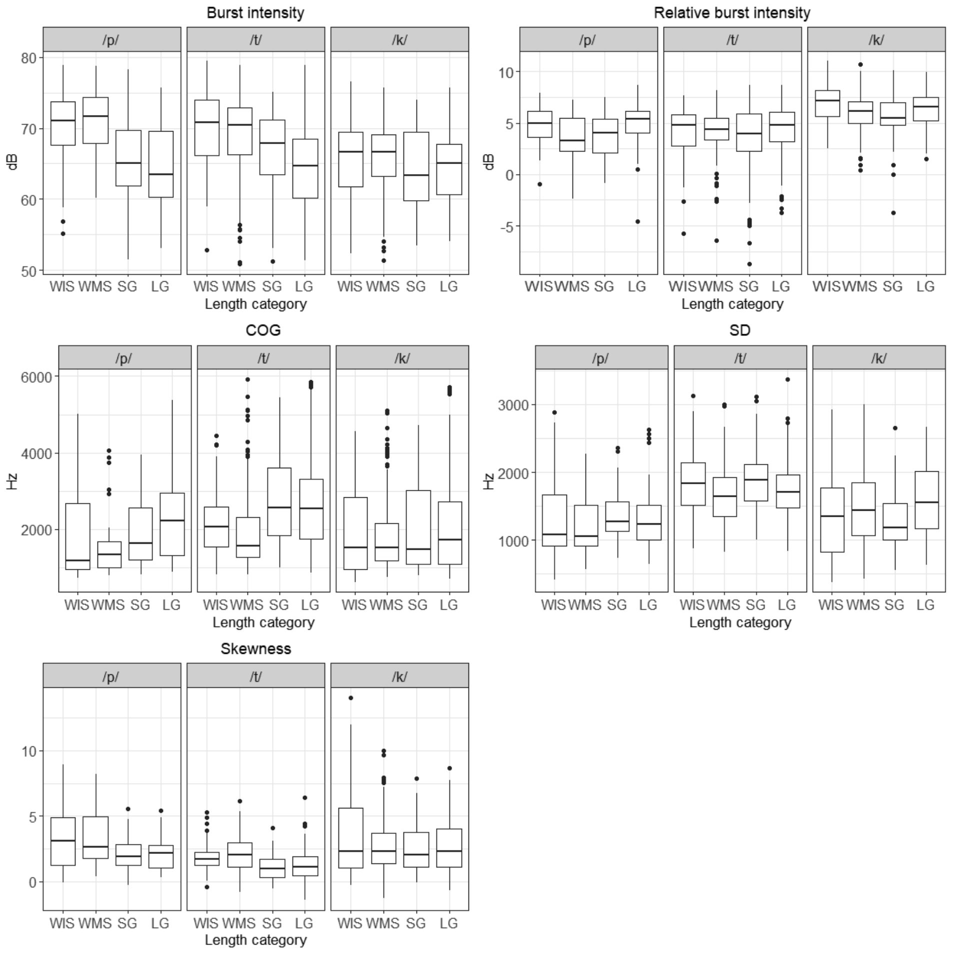

Figure 18. Variation of non-durational measures of plosive bursts by POA and plosive category (P4). From the upper left: absolute burst intensity, relative burst intensity, centre of gravity (COG), standard deviation (SD), skewness. Note the different scales on each subplot. Abbreviations: WIS – word-initial singleton, WMS – word-medial singleton (Q1), SG – short geminate (Q2), LG – long geminate (Q3).

4.3.3. Discussion

Concerning the variation in the burst phase, I controlled following hypotheses: there are no differences between burst durations either between singletons and geminates or word-initial and word-medial singletons; the burst intensity is higher in geminates than in singletons; there are no differences in intensity of primary-stressed word-initial and non-stressed word-medial singletons; geminates show higher COG, SD and lower skewness compared to singletons; there are no significant spectral differences between word-initial and word-medial bursts.

The Burst phase, as most of the other studied characteristics, was mostly influenced by POA; no differences emerged in length categories according to stress. The burst durations place Estonian in a group of languages with short-lag VOT, according to Lisker and Abramson (1964). Compared to the correlations in VOT between POAs given in Chodroff et al. (2019), labial – coronal 3 ms, coronal – dorsal 14 ms, dorsal – labial 17 ms, the differences were similar, although the differences between dorsal and velar categories were smaller in Estonian.

Compared to previous studies in Estonian that studied burst durations (Eek & Meister 1996b; Raasik 2010; Suomi & Meister 2012), most of the results align with those of P1, P3 and P4. As seen from Table 7, the mean burst durations were relatively close between the studies. Divergences are, first, longer burst durations in Raasik (2010) and P1, which represent spontaneous speech, and longer burst phase of word-initial /t/ in Eek and Meister (1996b) and shorter burst phases of /p/ and /k/ in Suomi and Meister (2012). Compared to studies in other languages with length distinction, the results are in line with Bengali (Kotzor, Wetterlin & Lahiri 2017), Italian (Esposito & Di Benedetto 1999), Japanese (Homma 1981) and Tashlhiyt Berber (Ridouane 2007), where no differences occurred in VOT between singletons and geminates.

Table 7. Comparison of mean burst durations of plosive categories in studies conducted on Estonian. Abbreviations: WIS – word-initial singleton, Q1 – singleton, Q2 – short geminate, Q3 – long geminate.

Study	Register	/p/	/t/	/k/
WIS, Eek and Meister (1996b)	Read	18	31	37
WIS, Suomi and Meister (2012)	Read	–	18	27
WIS, P4	Semi-spontaneous	17	18	25
Q1, Raasik (2010)	Spontaneous	25	28	26
Q1, Suomi and Meister (2012)	Read	13	16	19
Q1, P1	Spontaneous	22	23	29
Q1, P4	Semi-spontaneous	17	18	23
Q2, P2	Read, semi-spontaneous	17	17	26
Q2, P4	Semi-spontaneous	17	18	25
Q3, P2	Read, semi-spontaneous	15	18	26
Q3, P4	Semi-spontaneous	17	18	26

For singletons, the main difference in burst durations emerged between speech registers. The burst duration of singletons in spontaneous speech (Raasik 2010, P1) showed longer mean durations compared to the burst durations in semi-spontaneous (P4) or read speech (Soumi & Meister 2012), and the differences between POAs were smaller. It is up to further investigation to find the reasons for this kind of difference. Long bursts of /t/ and /k/ occurred also in lab speech in Eek and Meister (1996b), but these were based on only one speaker.

The hypothesis about higher intensity in geminates was rejected. Overall, geminates showed lower absolute intensity values than singletons. This result was contrary to the hypothesis based on similar data from Finnish (Doty, Idemaru & Guion 2007), where mean absolute intensities were higher in geminates. In relative intensities, the Q3 geminate had the highest intensity, but the Q2 geminates had the lowest. The higher relative intensity in Q3 geminates compared to Q2 and Q1 in Estonian can be explained by a substantial reduction in unstressed vowels following the Q3 consonant but not Q1 and Q2 (Eek & Meister 1997). Comparison of burst intensity to other languages with length distinction gave various results: Tashlhiyt Berber (Ridouane 2007) showed a similar pattern with higher intensities in geminates, but in Itunyoso Trique (Dicanio 2012) and Cypriot Greek (Tserdanelis & Arvaniti 2001) lower relative burst intensities occurred in geminates in some POAs or some speakers. The higher intensity of the word-initial position, i.e. primary stress, was not expected based on existing studies in other languages like English (Cho & Keating 2009) and Dutch (Cho & McQueen 2005) where no significant differences were found between stressed and unstressed positions. It is, however, difficult to draw any firm conclusions about the relative burst intensity because all the studies used slightly different calculation methods.

The spectral measures – COG, SD and skewness – were primarily influenced by place of articulation. Alveolar /t/ stood out with much higher COG and much lower skewness. Skewness values were positive for all POAs, suggesting more intensity in lower frequencies in the spectra. As for gemination, there were systematic differences in COG and skewness in /p/ and /t/, with geminates having higher COG and lower skewness than singletons. The SD values of geminates were higher than those of singletons, but no significant differences were revealed in statistical modelling. For /k/, the differences in COG and skewness were much smaller between length categories. It may be that the SD of /k/ is more affected by the following vowel. The vowel was not included as an independent variable in modelling in P3, but the strongest influence of vowel context on the spectra of /k/ compared to other POAs was found by Eek and Meister (1996c) in Estonian and by Nicolaidis et al. (2019) in Greek, so the same reason is probable here.

Word-initial singletons showed higher COG and lower skewness than word-medial singletons, but the values were close together compared to the geminate categories. The result is partly in line with Tabain et al. (2016) who compared stressed word-initial and unstressed word-medial plosives in Indonesian language Makasar and Australian languages Arrernte, Pitjantjatjara and Warlpiri without voicing contrast and found mixed behaviour in COG or SD of the burst spectra, with most languages apart from Pitjantjantjara showing small or no significant differences in spectral moments or differences occurring only for some POAs.

Compared to any other studies (Al-Tamimi & Khattab 2011; Chodroff & Wilson 2014; Tabain et al. 2016; Nicolaidis et al. 2019), the plosive bursts in Estonian had lower COG and higher skewness, indicating more intensive spectra in the lower frequencies. There can be several reasons for that. First, as with calculating the relative intensity of the plosive bursts, all kinds of technical details, like filtering frequencies or the length of the time frames used for analysis differ from study to study and probably play some role. Second, the speaking situation may play a role. Additionally, the values of spectral moments may characterise the differences in strength of articulation between speech registers. Most of the studies used for comparison in P4 were conducted on lab speech, that is near the hyperspeech end of the H–H continuum (Lindblom 1990) and therefore more controlled. Chodroff and Wilson (2014) compared lab speech with read speech and found lower COG and higher skewness in plosives in read speech. Data in P4 is from semi-spontaneous speech that is probably more loosely pronounced than lab speech or read speech, so the lower COG and higher skewness can reflect that. There may be differences between languages and their phonological categories. Sundara (2005) compared spectra of the bursts of coronal plosives in Canadian French and Canadian English and found measures referring to the lower frequency range of the bursts in Canadian French. She explained it with differences in the coding of the voicing. In French, the voicing is coded in glottal movement in the closure, while in English, the difference mainly lies in the burst phase and, therefore, needs to be more perceivable there. As Estonian does not use the voicing distinction, and the main coding of the quantity distinction is the closure duration, the differences in the burst are only secondary and do not have to be as large.

4.4. Fortis and lenis distinction and Estonian

Regarding the question of fortis and lenis distinction and its suitability for the Estonian plosive system, corroboration can be sought from the results of the thesis and other studies. In Estonian tradition, geminates and word-initial singletons are fortis, and word-medial and word-final singletons are lenis (e.g., Ariste 1968: 39–41; Metslang et al. 2023: 57–58). The view is, at least partly, supported by the spelling tradition (see p. 22). The underlying mechanisms or acoustic correlates of fortis and lenis have not been tested.

As mentioned in the introduction, fortis contrast is phonological and is based on the abstract concept of muscular tension or respiratory effort during articulation. From the list of most frequent correlates of the phonological category of fortis (Burroni, Lau-Preechathammarach & Maspong 2021), duration, burst duration, intensity, spectral measures of the burst, and implicitly articulatory resistance were studied in this thesis. Articulatory resistance can be observed through reduction patterns. Next, I briefly look at the behaviour of the above correlates in Estonian.

The closure duration difference is one of the most frequently used correlates for the distinction. In Estonian, the closure durations are clearly different between singletons and geminates (see ratios in Table 6, p 34). Word-initial and word-medial singleton, on the other hand, are very close in duration, as reported by Suomi and Meister (2012) and Suomi et al. (2013), except in the case of contrastive stress, where the word-initial singleton could be 20 ms longer than the word-medial one (both positions go through lengthening). In P2, word-medial singletons were over 10 ms shorter than either word-initial or word-medial singletons (the difference was influenced by the allophonic distribution, stress was not controlled). So, based on duration, geminates can be considered fortis, while the nature of word-initial, and also word-final singleton is more obscure.

Next, differences in VOT or burst duration are often found in fortis languages (e.g., Korean Cho, Jun & Ladefoged 2002; Standard German Kuzla & Ernestus 2011; Swiss German Ladd & Schmid 2018), although not in all (e.g., Itunyoso Trique Dicanio 2012). In P3, where bursts of word-initial and word-medial plosives were compared in the same data, no differences in burst duration apart from the ones caused by POA could be found. Instead, longer burst durations were found in singletons in spontaneous speech (P1), which should represent the most lenition in articulation. Longer burst durations in utterance-initial position have been measured by Eek and Meister (1996b). However, the data set contains speech only from one speaker and did not contain other positions. So, burst durations seem not to take part in possible fortis and lenis contrast in any way, although the nature of the utterance-initial position remains ambiguous.

Spectral differences were found in P4 between singletons and geminates. Compared to singletons, geminates had higher COG and lower skewness, which signal burst in higher frequencies and, therefore, more tense or forced articulation (Al-Tamimi & Khattab 2011). Word-initial and word-medial singleton were very close to each other, although word-initial singleton showed somewhat higher mean values. Again, geminates can be interpreted as more forcefully articulated but not word-initial singleton.

In Estonian plosives in P4, absolute burst intensity was lower in geminates than in singletons, and relative intensity was higher only in long geminates (Q3), which is caused by the reduction of the following vowel and the short geminate (Q2) showed the lowest relative intensity of all length categories. Dicanio (2012) made similar discovery in Itunyoso Trique, where relative burst amplitude was higher in lenis series. So, differences in burst do not seem to play part in the distinction either.

A great amount of passive voicing occurs in Estonian singleton plosives, which suggests weak articulation, characteristic of lenis plosives. Partial voicing is also present in geminate closures (P3). However, glottal vibration may occur in fortis plosives as well, like in German (Kuzla & Ernestus 2011) or Jawon (Jaeger 1983). High reduction rates in singletons can be explained by short durations, as in Itunyoso Trique (Dicanio 2012). Word-initial singletons were very often voiceless in P2 compared to word-medial and word-final singletons, especially in utterance-initial position, but also utterance-medially. Word-final singletons were close to word-medial in reduction patterns, although the durations were closer to word-initial singletons.

Based on findings presented in previous paragraphs, it can be said that while geminates show somewhat more ‘fortis’ behaviour, the differences are not significant enough to motivate the use of fortis and lenis in Estonian, in terms of articulatory force. Word-initial fortis lenites in voiced environment according to Metslang et al. (2023: 58) or neutralises, according to Eek and Meister (1996a: 67).

Currently, there is no sufficient evidence to confirm the nature of word-initial position. The quantity system in Estonian is ternary, so it might need additional distinction for short and long geminate (weaker fortis?). There are languages with ternary systems, such as Korean, that has categories for lenis, fortis and aspirated (Cho, Jun & Ladefoged 2002) but these categories have more distinguishable features than just duration, like aspiration.

Further, Metslang et al. (2023: 58) say that fortis and lenis are allophones of the same phoneme in Estonian, and Ariste (1968) lists many cases that one of the contrastive pair realises as the other. That brings up the question which of the two is the main variant of the phoneme and if the choice is motivated by anything more than just the spelling. It is noted that word-initial plosives in Estonian are pronounced similarly despite the spelling (e.g., Ariste 1933a; Eek & Meister 1996), which usually means that *b*, *d*, and *g* at the beginning of loan words are not voiced. As we have seen in P2, passive voicing occurs in word-initial singletons in utterance-medial positions. But there were loan words starting with *b*, *d*, and *g* pronounced voiced in utterance-initial position that suggests intentional voicing instead.

To summarise, based on this data, the only relevant features for the fortis and lenis distinction that can be used in Estonian seems to be durational difference and greater reduction rate in singleton plosives. From the languages used for comparison, Itunyoso Trique (Dicanio 2012) is the closest in acoustic behaviour. Dicanio stated that if the lenition patterns, such as voicing, follow the durational patterns, there is no need to apply abstract feature of articulatory strength. That seems to be the case in Estonian as well. As duration is always included in the features of fortis and lenis distinction, then the distinction could be used in Estonian as well. But other features concerning articulatory strength did not prove to be significant or were ambiguous. So, one might ask if it is reasonable to use a category based on articulatory force without the feature of articulatory force.

5. CONCLUSIONS

The thesis studied the acoustic characteristics of plosives in Estonian. The work aims to describe the allophonic variation and acoustic characteristics of Estonian plosives and find out the main factors influencing the variation. The four studies concerned several aspects of the acoustics of plosives in Estonian. Allophonic variation was studied in singleton plosives in spontaneous and read speech in P1 and P2, respectively, and geminate plosives in semi-spontaneous and read speech in P3. Durational variation of singleton and geminate plosives were described in P1 and P3, respectively. The burst phase was studied in singletons in P1, in geminates in P3 and in all length degrees, including a word-initial position, in P4.

The work intended to answer the following research questions:

RQ1: What kind of allophonic variants emerge in plosives in Estonian connected speech?

RQ2: How do segmental (place of articulation, adjacent segments) and suprasegmental features (quantity, stress) influence the allophonic variation of plosives?

RQ3: How does the speech register influence the allophonic variation of plosives?

RQ4: How do segmental (place of articulation, adjacent segments) and suprasegmental (quantity, syllable structure, stress) features influence the durational variation of plosives?

RQ5: How do segmental (place of articulation) and suprasegmental (quantity, primary stress) features influence the acoustics of the burst phase?

The POA was the primary influencer in most aspects of plosive acoustics – the allophonic variation, burst duration, burst intensity and spectral moments. Quantity degree was the main basis for closure duration differences and was also significant in non-durational measures of the burst phase.

Two kinds of allophones emerged – full allophones with a distinguishable release burst and reduced allophones that had lost the burst phase. Full allophones differed in the voicing in the closure phase: voiceless, partly voiced and voiced. Reduced allophones included voiced closure without burst, fricative, approximant, and total loss. Reduced allophones occurred only in singleton plosives, especially in spontaneous speech, but partly voiced allophone also occurred in geminates.

The amount of reduced allophones differed between POAs, being the smallest in /p/ and the largest in /k/ among the singleton plosives. The influences on adjacent vowel context and the stressed position of the plosive carrying word were the strongest on /t/ on both allophonic distribution and durational variation but were small. The plosives in read speech showed less reduction than in spontaneous speech.

The durational behaviour depended on POA and quantity degree and, to a lesser extent, on allophone. The duration of geminates was also correlated to the phonological length of the preceding vowel, being shorter after long vowels. Durational variation by POA followed universals: /p/ had the longest duration and /t/ the shortest. The main factor distinguishing length categories was the closure duration. The durational ratios were 1.7–2 between Q2 and Q1, 2.2–2.4 between Q3 and Q1 and 1.2–1.27 between Q3 and Q2. The allophones were patterned in duration according to the presence of the burst and voicing in the closure phase.

The burst duration was solely influenced by the POA. There were almost no differences between length degrees or word-initial and word-medial positions. That follows the data of most other languages that use the length category. There occurred a difference between speech registers: burst duration values in singletons in spontaneous speech were longer than in read and semi-spontaneous speech.

P4 gave the first overview of burst intensity and three spectral moments of the bursts: centre of gravity (COG), standard deviation (SD), and skewness of the burst spectra in Estonian plosives. The primary influencer on burst acoustics was, again, POA. Gemination affected burst intensity and some spectral moments. Between length categories, mean absolute intensities were unexpectedly higher in singletons than in geminates. In relative intensities, Q2 geminate had the lowest intensity compared to other length categories. The big difference between the relative intensities of geminates may be caused by the extensive reduction in the unstressed syllable following the Q3 geminate but not singleton and Q2 geminate (Eek & Meister 1997).

The geminates had higher COG and lower skewness, suggesting intensity in higher burst frequencies both in measurable units and in spectral shape. The word-initial singleton had higher burst intensity than the word-medial categories, both in absolute and relative measures. Compared to the word-medial singleton, it also showed higher COG and lower skewness, but these were still closer to word-medial singleton than geminates. The values of spectral measures suggested bursts in lower frequencies in Estonian than in many other languages. That may be connected to the various filtering and amplification parameters, different speaking styles and probably also language-specific differences.

Additionally, a nature of possible correlates of fortis and lenis distinction besides short and long in Estonian was discussed. Conclusion after examining a set of acoustic measures available in published material about Estonian and other languages was that the distinction could be made primarily on the base of duration, because other features were either irrelevant (burst phase duration) or did not give sufficient information (burst spectra).

In summary, the thesis studies showed extensive coarticulatory reduction in Estonian plosives, which was especially prevalent in singleton plosives and occurred in all speech registers studied. While many results confirm universal tendencies in plosive acoustics related to the POA and gemination, some new features were found, especially those concerning the burst phase. The burst phase of plosives in connected speech has only been studied durationally. P4 was the first study of burst spectra in Estonian and one of the first in the world to cover the burst spectra in connection with gemination (only other to the author's knowledge is Al-Tamimi & Khattab 2011 in Lebanese Arabic).

The studies of this thesis represented a corpus-linguistic approach. The corpus data of connected speech used in the studies is a difficult material that may contain unbalanced data and capturing all the important factors influencing certain phenomena is difficult. Still, the use of language, as it is really spoken combined to quantitative methods allows to find factors that might not be considered in controlled samples. Tendencies found give the basis for further investigation.

Plosives are sophisticated sounds, and many aspects still need further study. Based on the findings of the current thesis, elaboration is welcome, for example, on behaviour of burst phases of geminates in spontaneous speech to find out if similar differences occur as in spontaneous speech. The spectral moments of singleton plosives in structures of different quantities (as singletons can occur in all quantity degrees) could be compared with geminates to corroborate the possible importance of spectral moments in quantity distinction. Features of word-initial position stayed rather unstudied, so further studies are welcome.

KOKKUVÕTE

Sulghäälikute häälduse varieerumine eesti keeles

Sissejuhatus

Käesolev uurimus käsitleb eesti keele klusiilide ehk sulghäälikute /p t k/ akustilist varieerumist. Töö peamine eesmärk on kirjeldada eesti keele sulghäälikute allofoonilist varieerumist ja akustilisi omadusi laiemas tüpoloogilises perspektiivis ning täita mõned lüngad eesti keele sulghäälikute kirjelduses.

Sulghäälikute hääldamisel on artikulatoorselt eristatavad kolm faasi: sulu moodustamine (implosioon), sulu hoidmine (oklusioon), mille jooksul rõhk kõnetraktis suureneb, ja sulu vallandumine (eksplosioon) (Ladefoged & Johnson 2011: 14). Akustiliselt on eristatavad sulufaas, mis võib olla helitu, poolheliline või heliline ja vallandumisfaas. Neid faase saab iseloomustada akustiliste omaduste alusel, nagu kestus, helilisus sulufaasi ajal või vallandumise spekter.

Sulghäälikute kolm peamist hääldusкohta on bilabiaalne [p], alveolaarne või alveolaardentaalne [t] ja velaarne [k]. Lisaks moodustuskohal põhinevatele eristustele võib keeles kasutada ka muid kategooriaid. Kolm sagedasemat sulghäälikute fonoloogilist kategoriseerimissüsteemi maailma keeltes on helilisus, mille puhul põhineb kategooriate eristamine sulufaasi helilisusel või vallandumisfaasi kestusel, fortis, mille puhul kategooriaid eristab hääldusе tugevus, ja välde, mille puhul põhineb vastandus hääliku kestusel.

Fonoloogilised eristused esinevad sageli kas sõna-algulises või sõna-siseses positsioonis ja allofooniline varieerumine keeles sõltub sellest, kui palju ja milliseid fonoloogilisi kategooriaid keeles kasutatakse; arvatakse, et vähem eristusi võimaldab rohkem ühe eristava kategooria siseselt (Keating, Linker & Huffman 1983).

Sulghäälikute akustilisel kirjeldamisel kasutatakse peale üldiste kestusandmete veel mitmeid näitajad, mis üldiselt iseloomustavad vallandumisfaasi. Helilisuse algamise aeg (*voice onset time*, VOT) on ajavahemik sulu vallandumisest kuni järgneva hääliku helilisuse tekkeni (Lisker & Abramson 1964). VOT on positiivne, kui helilisus algab pärast sulu vallandumist ja negatiivne, kui helilisus algab enne sulu vallandumist. VOT on tugevas korrelatsioonis sulu vallandumise üldise kestusega, kuid ei pruugi alati kattuda. Vallandumise kestust mõjutab esmajärjekorras hääldusкoht: mida tagapool on hääldusкoht, seda pikem on vallandumine ja VOT, mida kiiremad on hääliku moodustamisel osalevad artikulaatorid, seda lühem (Cho & Ladefoged 1999: 209–2013). Kõige pikem on üldiselt tagapoolse hääldusкohaga velaari [k], ja kõige lühem bilabiaalse [p] vallandumine. Mittekestuslikest tunnustest kasutatakse vallandumisfaasi iseloomustamisel enim intensiivsust ja spektraalmomente. Vallandumise intensiivsust kasutatakse häälduse tugevuse kirjeldamisel (Hayward 2000: 43), seda mõõdetakse detsibellides. Suhtelise intensiivsuse arvutamisel vallandumisfaasi inensiivsus jagatakse järgneva vokaali intensiivsusega või lahutatakse sellest. Vallandumisfaasi spektri iseloomustamiseks kasutatakse komplekti näitajaid, mida nimetatakse spektraalmomentideks (Forrest et al. 1988). Esimesed kaks spektraalmomenti iseloomustavad spektri sagedusvahemikku, neid mõõdetakse hertsides (Hz), need on spektri raskuskese (*centre of gravity*, COG), mis viitab kõige intensiivsemale sagedusvahemikule spektris ja spektri standardhälve (*standard deviation*), mis iseloomustab spektri hajuvust ehk kogu kaetud sagedusvahemikku (Forrest et al. 1988; Jongman, Wayland & Wong 2000). Kaks viimast spektraalmomenti iseloomustavad spektri kuju võrreldes normaaljaotusega, need on mõõtühikuta näitajad. Spektri asümmeetriakordaja (*skewness*) näitab, kas vallandumise spekter on intensiivsem kõrgematel või madalamatel sagedustel (normaaljaotuse kalde väärtus on null, kõrgemad sagedused negatiivse väärtusega) (Forrest et al. 1988; Jongman, Wayland & Wong 2000; Tabain & Butcher 2015). Ekstsess ehk järsakuskordaja (*kurtosis*) näitab, kas spekter on pigem ühtlane või esineb seal inensiivsustippe (*spectral peak*) (normaaljaotuse väärtus on 1, ühtlase spektri väärtus madalam) (Forrest et al. 1988; Tabain & Butcher 2015; Guthrie 2020). Häälduskoha järgi on sulghäälikute spektrid järgmised: bilabiaali spekter on hajus ja intensiivsem madalatel sagedustel (kõrge SD, kõrge asümmetriaväärtus), alveolaari spekter on hajus ja intensiivsem kõrgematel sagedustel (kõrge COG, madal asümmeetriaväärtus), velaari spekter on kompaktne ja intensiivsem keskmistel sagedustel (madal COG, madal SD) (Forrest et al. 1988; Cole et al. 2007; Chodroff & Wilson 2014; Nicolaidis et al. 2019).

Seotud kõnes leiavad aset koartikulatsiooniprotsessid, mis on tingitud artikulaatorite omavahelistest vastasmõjudest. Koartikulatsioon väljendub akustiliselt assimilatsioonis ja reduktsioonis ehk häälduse nõrgenemises. Häälduse nõrgenemise korral muutuvad segmendid lühemaks, nende hääldus muutub ebamäärasemaks ja segment võib kaduda (Zellers, Schuppler & Clayards 2018: 1). Assimilatsioon on häälikusegmendi tunnuse levimine teistele lähedalasuvatele segmentidele, seda on kasutatud koartikulatsiooniga paralleelselt (Farnetani & Recasens 2013). Assimilatsioon võib väljenduda kolme moodi: häälduskohaassimilatsioon, kus ühe segmendi häälduskoht mõjutab kõrvalasuva segmendi oma (alveolaarne /n/ muutub velaarse /k/ ees samuti velaariks [ŋ], *kang* [kɑŋk]), hääldusviisiassimilatsioon, mille puhul ühe segmendi hääldusviis kandub kõrvalasuvale segmendile (näiteks järjend [hts] sõnas *lihtsalt* assimileerub sibilandiks [sːs] lisːsɑlt) ja helilisusassimilatsioon, mille puhul levib ühe segmendi helilisus (näiteks muutuvad helilised konsonandid helitu obstruendi järel samuti helituks *lehm* [leh:m̥].

Sagedasemad sulghäälikutega toimuvad koartikulatsiooninähtused on helilisusassimilatsioon ja häälduse redutseerumine, mis väljendub vallandumisfaasi nõrgenemises või kaos. Lisa Davidson (2016; 2018) on ameerika inglise keele põhjal eristanud nelja tüüpi sulufaasi helilisust: helilisuse edasikandumine (*bleed*), mille puhul muutub sulu algusosa eelmise hääliku mõjul heliliseks, negatiivne VOT, mille puhul ilmub helilisus sulu lõpuosas enne vallandumist, läbiv helilisus (*through*), mille puhul on heliline kogu sulufaas ja helilisuse ajutine ilmumine sulu jooksul (*hump*). Sagedasemad on *bleed*- ja *through*-tüüp.

Reduktsiooninähtusi on palju uuritud nt romaani keeltes (Torreira ja Ernestus 2011; Recasens ja Mira 2012). On uuritud lause fraasistruktuuri, silbistruktuuri ja prosoodia mõju allofoonide kestusele ja kvaliteedile (Cole jt 2007; Kuzla ja Ernestus 2011).

Eesti keele konsonandisüsteemis on neli sulghäälikufoneemi: bilabiaalne /p/, alveolaarne /t/, palataliseeritud alveolaarne /tʲ/ ja velaarne /k/. Kõik sulghäälikud on helitud ja aspireerimata.

Eesti keel kuulub koos liivi, isuri ja mõnede saami keeltega keelte hulka, millel on kolmene vältesüsteem (vt Türk 2019b). Konsonantide puhul vastab esimene välde (Q1) lühikesele ehk üksikhäälikule, teine välde (Q2) lühikesele geminaadile ja kolmas välde (Q3) pikale geminaadile. Välde on suprasegmentaalne nähtus, mis hõlmab lisaks segmendi kestusele kõnetakti rõhuliste ja rõhutute silpide kestuse suhet ning põhitoonikontuuri. Sõnaalguline häälik vältevastanduses ei osale, kuid kannab esisilbirõhku (välja arvatud mõnedes onomatopoeetilistes ja laensõnades) (Asu jt 2016: 127). Helilised kaashäälikud võivad esineda üksikhäälikutena pärast lühikesi ja pikki vokaale ning geminaatidena ainult pärast lühikesi vokaale. Helitud obstruendid võivad seevastu esineda nii üksikhäälikutena kui ka geminaatidena nii lühikeste kui ka pikkade vokaalide ja diftongide järel. Sõnaalguse üksikhäälikuid on kestuse aspektist vaadeldud peamiselt välteuurimustes. Sõnaalguliste häälikute kestused kattuvad tavaliselt sõnasiseste üksikhäälikutega (Eek & Meister 2003; Suomi & Meister 2012; Suomi jt 2013).

Lisaks kvantiteedile on eesti keeles kasutatud ka fortise ja leenise eristust (Ariste 1933a; 1939; 1968; Metslang jt 2023: 57–58). On öeldud, et geminaadid on fortised ja sõna keskosas ja lõpus asuvad üksikhäälikud helikontekstis on leenised. Sõnaalguline sulghäälik on kas fortis (Ariste 1933a; Metslang jt 2023: 57–58) või neutraliseerunud (Eek & Meister 1996a: 67). Seda vaadet toetab eesti ortograafia, mis kasutab samu grafeeme – *p, t, k* – sõnaalguliste sulghäälikute ja geminaatide tähistamiseks ning *b, d, g* sõnasiseste lühikeste sulghäälikute tähistamiseks. Seni ei ole tehtud eksperimentaalseid uuringuid, et leida erinevusi artikulatoorse pingutuse korrelaatides. Käesolev väitekiri pakub lühikest arutelu eesti sulghäälikute akustika kohta fortise ja leenise tunnuste vaatest.

Eesti sulghäälikute vallandumisfaasi on seni vähe uuritud. Üksikhäälikute vallandumisfaasi kestusi on uuritud loetud (Suomi & Meister 2012) ja spontaanses kõnes (Raasik 2010). Sõnaalgulise sulghääliku vallandumise spektreid on kirjeldanud Eek ja Meister (1996c). Geminaatide vallandumisfaase ei ole üldse uuritud, samuti ei ole olnud ühtegi uuringut, mis võrdleks vallandumisi sõnaalgulises ja sõnasiseses positsioonis. Publikatsioonid P1 ja P2 vaatlevad vastavalt lühikeste sulghäälikute ja geminaatide kestusaspekte, P4 lisab uusi teadmisi eesti sulghäälikute vallandumisfaaside kestuse, spektraalse koostise ja intensiivsuse kohta.

Eesti sulghäälikute koartikulatsiooninähtustest on märgatud sulufaasi heliliseks muutumist lühikestes sulghäälikutes nii loetud (Ariste 1933b; Suomi & Meister 2012) kui ka spontaanses kõnes (Raasik 2010). Suomi ja Meister (2012) ja Raasik (2010) leidsid ka vallandumisfaasi kadu. Velaarne sulghäälik /k/ redutseerus kõige rohkem. Geminaatide puhul on sulufaasi osalist helilisust leidnud Ariste (1933a; 1933b). Publikatsioonid P1, P2 ja P3 uurivad edasi eesti üksikhäälikute ja geminaatide allofoonilist varieerumist ja vähendamist, lisades võrdluse allofoonilisele varieerumisele kahes kõneregistris.

Töö peamised uurimisküsimused olid:

1) Millised allofoonilised variandid tekivad eesti keele klusiilides seotud kõnes?
2) Kuidas mõjutavad segmentaalsed (häälduskoht, häälikuline kontekst) ja suprasegmentaalsed (välde, rõhk) tunnused sulghäälikute allofoonilist varieerumist?
3) Kuidas mõjutab allofoonilist varieerumist kõneregister?
4) Kuidas mõjutavad segmentaalsed (häälduskoht, häälikuline kontekst) ja suprasegmentaalsed (välde, silbistruktuur, rõhk) tunnused sulghäälikute kestuste varieerumist?
5) Kuidas mõjutavad segmentaalsed ja suprasegmentaalsed tunnused sulghäälikute vallandumisfaasi akustikat?

Doktoritöö koosneb sissejuhatavast osast ja neljast artiklist. Sissejuhatavas osas on kuus peatükki. Esimene peatükk tutvustab doktoritöö eesmärke ja uurimisküsimusi ning annab ülevaate publikatsioonidest. Teine peatükk selgitab põhimõisteid ja teoreetilisi lähtekohti: sulghäälikute artikulatoorset ja akustilist iseloomustamist, sulghäälikute fonoloogilise kategoriseerimise süsteeme eri keeltes, koartikulatsiooni mõjusid sulghäälikute allofoonilisele varieerumisele ning tutvustab eesti keele sulghäälikusüsteemi ja selle kohta tehtud uuringuid. Kolmas peatükk tutvustab töös kasutatud materjale ja metoodikat. Neljandas peatükis esitatakse tulemused ja arutelu. Viies peatükk võtab kokku olulisemad tulemused ja järeldused. Kuues peatükk sisaldab töö eestikeelse kokkuvõtet. Järgnevad töös kasutatud viited.

Publikatsioonid käsitlevad eri aspekte eesti keele sulghäälikute akustikast. Artiklis P1 uurisin lühikeste sulghäälikute allofoonilist varieerumist ja seda mõjutavaid tegureid spontaanses kõnes. Artikkel P2 võttis kokku lühikeste sulghäälikute allofoonide ja tekstitunnuste vahelisi seoseid uurinud pilootuuringu tulemused. Artikkel P3 uuris geminaatsulghäälikute akustikat poolspontaanses ja loetud kõnes ning häälduskoha, välte ning eelneva vokaali pikkuse mõju. Artikkel P4 keskendus vallandumisfaasi akustikale. Uurisin häälduskoha, välte ja pearõhulise positsiooni mõju vallandumisfaasi kestusele, intensiivsusele ja spektri kujule kõigis kolmes vältes ja sõnaalgulises positsioonis.

Materjal ja meetod

Materjal pärineb kahest eesti keele foneetilisest korpusest. Spontaanse ja poolspontaanse kõne materjalid (P1, P3, P4) on võetud Tartu Ülikooli eesti keele spontaanse kõne foneetilisest korpusest (EKSKFK) (Lippus 2018; Lippus et al. 2021; Lippus et al. 2023), loetud kõne näited (P2, P3) Eesti Keele Instituudi kõnesünteesikorpusest (Piits 2016).

EKSKFK koosneb mitmest alamkorpusest: helistuudios salvestatud spontaansed dialoogid, peamikrofoniga salvestatud poolspontaansed loengud ja ettekanded, välitöödel salvestatud dialoogid. Salvestised on 16 bit / 44.1 kHz PCM wave formaadis. Kõnelejad esindavad kõiki eagruppe ja piirkondi. Helifailid märgendatakse programmiga Praat mitmel tasandil: sõnad, häälikud SAMPA (Wells 2005) sümbolitega, silbid, taktid, morfoloogia, häälelaad, lausungid, intonatsioonifraasid. Korpuse käesolevas versioonis (v1.3) on 135 tundi kõnet 207 kõnelejalt. EKI kõnesünteesikorpus sisaldab suuremalt jaolt stuudios ette loetud tekste. Iga kõneleja korpuses on alamkorpus lausetest, mis katavad kõik eesti keeles leiduvad häälikukombinatsioonid ja lisaks tekste uudistest, ilukirjandusest vms. Helifailid märgendatakse automaatselt sõna ja hääliku tasandil, kasutades programmi WebMAUS (Kisler, Reichel & Schiel 2017).

Kõigi uurimuste jaoks märgendati programmis Praat (Boersma & Weenink 2018) lisaks olemasolevale häälikutasandi märgendusele käsitsi sulghääliku sulufaasi helilised ja helitud osad ja vallandumisfaas. Märgendus ja kestused eraldati Praati skripti abil. P4 jaoks eraldati lisaks vallandumisfaaside kestustele ka vallandumisfaasi absoluutse ja suhtelise intensiivsuse, spektri raskuskeskme, spektri standardhälbe ja spektri asümmeetria väärtused (signaali võimendus al 500 Hz, Hanni filter 500–10000 Hz, Hammingi aken 10 ms sulu vallandumise algusest). Keskmine vallandumise intensiivsus (dB SPL) mõõdeti 10 ms jooksul alates vallandumise algusest (põhitooni miinimum 500 Hz). Vallandumise suhtelise intensiivsuse arvutamiseks lahutati vallandumise esimese 10 ms keskmine intensiivsus järgneva vokaali esimese 50% kestuse keskmisest intensiivsusest.

P1 uurimismaterjaliks olid sagedasemad täistähenduslike sõnade (nimisõnad, verbid, omadussõnad, arvsõnad, järgarvsõnad) lemmad, mis esinesid vähemalt 50 korda EKSKFK sagedussõnastikus (Lippus 2019) ja mille esimese ja teise silbi piiril oli intervokaalne lühike sulghäälik. Materjal P1 jaoks koguti EKSKFK avaliku otsimootoriga. Otsimootor tagastab päringule otsingusümboli kahesekundilise kontekstiga helifaili ja märgendusfaili. Andmestik sisaldas 1491 häälikujuhtu 19 lemmast. Andmeanalüüsiks kasutati MS Exceli tarkvara.

P2 andmestiku moodustasid heliliste häälikute või helilise hääliku ja pausi vahel asuvad sõnaalgulised, sõnasisesed ja sõnalõpulised lühikesed sulghäälikud ühe nais- ja ühe meesdiktori loetud uudistekstidest. Andmestik pärines EKI kõnesünteesikorpusest ja sisaldas 250 häälikujuhtu nais- ja 262 meeskõnelejalt. Andmeanalüüsiks kasutatai SYSTATi ja Ri (R Core team 2023) tarkvara. Loodi binaarsed logistilised mudelid, et ennustada sulu heliliseks muutumist ja vallandumisfaasi kadu. Mudelite prognoosivõimet hinnati diskriminantanalüüsiga.

P3 admestiku moodustasid esimese ja teise silbi piiril asuvad intervokaalsed geminaatsulghäälikud teise- ja kolmandavälteslistes sõnades. Materjal P3 jaoks pärines mõlemast korpusest: 12 kõnelejat EKSKFK monoloogide allkorpusest ja seitse kõnelejat EKI kõnesünteesikorpusest. Andmestik sisaldas 1795 häälikujuhtu.

P4 andmestik pärines EKSKFK monoloogide alamkorpusest. Kolme nais- ja kolme meeskõneleja kõnest eraldati 2346 sulghääliku vallandumisfaasi, mis olid kas sõnaalgulised või sõnasisesed ja esinesid intervokaalses helitu sulufaasiga sulghäälikus.

P3 ja P4 andmeanalüüs tehti programmis R (R Core team 2023) kasutades paketti lme4 (Bates et al. 2015), *post-hoc* testina kasutati Tukey Honest Significant Difference (TukeyHSD) funktsiooni ja P4 emmeans paketti (Lenth 2022). Kasutati lineaarseid segamudeleid (*lmer*, *glmer*).

P3 koostati lineaarsed segamudelid (*lmer*), et hinnata häälduskoha, välte, allofooni, silbistruktuuri ja kõnestiili mõju hääliku kogukestusele ning vallandumisfaasi ja sulufaasi helilise siirde kestusele ning generaliseeritud lineaarne segamudel (*glmer*), et hinnata häälduskoha, välte, silbistruktuuri, kõnestiili ja hääliku kestuse mõju helilise siirde esinemise tõenäosusele sulufaasi alguses. Juhuslikud muutujad olid kõneleja ja sõna.

P4 koostati lineaarsed segamudelid (*lmer*), et hinnata häälduskoha ja pikkuskategooria mõju vallandumise kestusele, absoluutsele ja suhtelisele intensiivsusele, spektri raskuskeskmele, spektri standardhälbele ja spektri asümmeetriakordajale. Juhuslikud muutujad olid mudelites kõneleja ja sulghäälikule järgnev vokaal.

Tulemused

Uurimustes P1 ja P2 vaatlesin lühikeste sulghäälikute ja uurimuses P3 geminaatide allofoonilist jagunemist, allofoonide kestusi ja variatsiooni mõjutavaid tegureid. Eristusid kaht tüüpi allofoonid: realiseerunud – vallandumisfaas on säilinud, erinevused sulufaasi helilisuses (helitu, poolheliline, heliline); redutseerunud – vallandumisfaas on kadunud, sulufaasi hääldus nõrgenenud (heliline sulufaas, frikatiiv, poolvokaal, täielik kadu). Redutseerunud allofoonid esinesid vaid lühikestes sulghäälikutes.

Lühikestes sulghäälikutes esines häälduse nõrgenemist nii loetud (P2) kui spontaanses kõnes (P1). Mõlemas registris olid realiseerunud allofoonid enamuses, kuid spontaanses kõnes esines redutseerunud allofoone rohkem kui loetud kõnes. Kõige rohkem redutseerunud allofoone esines /k/ puhul, spontaanses kõnes esines neid kokku rohkem kui realiseerunud allofoone. Eri häälduskohad käitusid isemoodi: /p/ puhul esines kõige rohkem poolhelilist allofooni, /t/ puhul helilist vallandumisfaasiga allofooni, /k/ puhul redutseerunud allofooni. Geminaatides (P3) esines samuti sulufaasi heliliseks muutumist, osaliselt helilist sulufaasi esines rohkem kui täiesti helitut. Kõige suurem osaliselt helilise sulufaasi hulk oli sarnaselt lühikestele sulghäälikutele /p/, kuid kõige väiksem /k/ esinemuste hulgas.

Artiklis P1 vaatlesin häälikulise ümbruse ja rõhu mõju lühikeste sulghäälikute allofoonilisele jaotusele ja kestusele. Spontaanses kõnes ei ilmnenud tugevaid mõjusid. Suurimad olid ümberkaudsete vokaalide mõjud kestusele ja allofoonilisele jaotusele /t/ puhul. Rohkem redutseerunud allofoone, eriti frikatiive, esines eesvokaalide (/i, e, æ/) ümbruses kõigi häälikute puhul, /t/ puhul esines /i/ naabruses võrreldes teiste kontekstidega ka rohkem täielikku kadu. Labiaalvokaalide /o, u/ ümbruses esines /p/ ja /k/ puhul rohkem poolvokaalset allofooni; /i/ puhul oli sarnases ümbruses vähem redutseerunud allofoone. Sulghäälikut sisaldava sõna sattumine rõhulisse positsiooni mõjutas allofoonilist jagunemist, kuid mitte sama allofooni esinemuste kestust. Rõhulises positsioonis esines rohkem realiseeritud ja rõhutus positsioonis rohkem redutseerunud ja juhuslikke allofoone. Häälduskohtade võrdluses oli suurim erinevus kahe konteksti vahel /t/ puhul. Helitu allofooni osakaal oli /t/ ja /k/ puhul mõlemas positsioonis umbes sama, kuid /p/ puhul rõhutus asendis suurem. Osaliselt heliline allofoon esines kõigi häälduskohtade puhul rohkem rõhulises positsioonis. Kestuste osas olid vokaal-kontekstist rohkem mõjutatud helitu ja poolhelilise allofoon, heliliste allofoonide puhul regulaarseid mõjusid ei ilmnenud. Rõhulise ja rõhuta positsiooni võrdluses olid samade allofoonide kestused mõlemas positsioonis sarnased.

Sulghäälikute kestus sõltus häälduskohast, allofoonist ja vältest. Kõige pikema kestusega häälduskohtade võrdluses oli kõigis väldetes /p/, järgnesid /k/ ja /t/. Lühikeste häälikute puhul oli kõige pikema kestusega poolheliline allofoon, järgnesid helitu, heliline ja vallandumisfaasita heliline, frikatiivne allofoon oli kestuselt lähedane vallandumisfaasita helilise allofooniga ning kõige lühem oli poolvokaalne allofoon. Geminaatidel oli pikema kestusega helitu ja lühem poolheliline allofoon. Välteid eristas ootuspäraselt sulufaasi kestus, kolmandas vältes geminaadid kestsid 1.17–1.27 korda kauem kui teises vältes geminaadid.

Vallandumisfaasi kestus sõltus täielikult häälduskohast. Pikkuskategooriate võrdlusel ei leitud statistiliselt olulisi erinevusi ei lühikeste ja geminaatsulghäälikute vahel ega ka sõnaalguliste ja sõnasiseste lühikeste sulghäälikute vahel. Võrreldes uuringutega teistes pikkusvastandusega keeltes, on tulemus kooskõlas näiteks bengali (Kotzor, Wetterlin & Lahiri 2017), itaalia (Esposito & Di Benedetto 1999), jaapani (Homma 1981) ja tashlhiyt berberi keelega (Ridouane 2007), kus samuti ei esinenud erinevusi sulufaasi kestuses lühikeste ja geminaatsulghäälikute vahel. Ainus erinevus kõigi uurimuste lõikes ilmnes lühikestes sulghäälikutes spontaanse kõne ning loetud kõne vahel, kusjuures spontaanse kõne vallandumisfaasid olid pikema kestusega kui loetud kõnes. Ka varasemates uurimustes, kus mõõdeti eesti keele sulghäälikute vallandumisfaasi kestusi (Raasik 2010; Suomi & Meister 2012), on jõutud sarnaste tulemusteni. Lühikeste sulghäälikute vallandumisfaasid spontaanses kõnes (Raasik 2010, P1) olid pikemad võrreldes poolspontaanse (P4) või loetud kõnega (Soumi & Meister 2012) ning erinevused häälduskohtade vahel olid spontaanses kõnes väiksemad.

Mittekestuslikest tunnustest otsisin artiklis P4 seaduspärasid vallandumise absoluutse ja suhtelise intensiivsuse, spektri raskuskeskme (COG), spektri standardhälbe (SD) ja spektri kalde variatsioonis. Siin oli variatsiooni peamine mõjutaja samuti häälduskoht: COG /t/ > /k/ > /p/, SD /t/ > /k/ > /p/, kalle /p/ > /k/ > /t/. Vallandumise intensiivsus oli absoluutväärtustes kõrgeim /p/-l ja madalaim /k/-l, kuid suhtelises (suhe järgmise vokaali intensiivsusega) vastupidine. Ootamatu oli tulemus lühikeste ja geminaathäälikute vahelises võrdluses: vallandumise intensiivsuse absoluutväärtused olid kõrgemad lühikestes sulghäälikutes kui geminaatides. Suhteliste väärtuste võrdluses oli kõige kõrgem intensiivsus kolmanda välte geminaatidel, kuid teise välte geminaatide oma oli kõige madalam. Suur erinevus kolmanda välte geminaatide ja lühikeste sulghäälikute suhtelises intensiivsuses on tõenäoliselt seotud sulghäälikule järgneva rõhutu vokaali tugeva redutseerumise ja seetõttu väiksema intensiivsusega kolmandas vältes (Eek & Meister 1997). Spektraalmomentide puhul ilmnesid regulaarsed erinevused lühikeste ja geminaatsulghäälikute vahel. Geminaatide vallandumistel olid võrreldes lühikeste häälikutega kõrgemad raskuskeskmed ja madalam kalle, mis näitab nende vallandumist kõrgematel sagedustel, see omakorda võib viidata tugevamale hääldusele (Al-Tamimi & Khattab 2011). Sõnaalguliste lühikeste sulghäälikute spektraalmomentide väärtused viitasid vallandumisele kõrgematel sagedustel kui sõnasisestel lühikestel sulghäälikutel, kuid siiski olid need sõnasisestele lühikestele häälikutele lähemal kui geminaatidele. Võrreldes teistes keeltes tehtud uurimustega, olid eesti keele sulghäälikute vallandumise spektri raskuskeskmed märkimisväärselt madalamad ja asümmeetriakordajad kõrgemad kui nt inglise (Chodroff & Wilson 2014) ja kreeka keele (Nicolaidis et al. 2019) kohta käivates uurimustes. See võis olla tingitud heli eeltöötluse parameetritest, kõneregistri erinevustest või ka keeltevahelistest erinevustest. Käesolevas uurimuses kasutati poolspontaanset kõnet, aga suurem osa spektraalmomente käsitlenud uurimustest on kasutanud laborikõnet. Kuna seotud kõne on tõenäolisemalt pisut nõrgema hääldusega, on tõenäoline, et sulud vallanduvad madalamatel sagedustel. Erinevusi kõneregistrite vahel leidsid ka Chodroff ja Wilson (2014), kes mõõtsid spektraalmomente laborikõnes salvestatud silpides ja loetud kõnes ning leidsid, et spektri raskuskeskmed olid loetud kõnes madalamad ja spektri asümmeetriakordaja väärtused kõrgemad loetud kõnes.

Võrdlesin eesti keele sulghäälikute akustilisi tunnuseid fortise ja leenise vastanduses oluliste tunnustega. Leidsin, et olulisematest tunnustest (kestus, vallandumise kestus, vallandumise intensiivsus, redutseerumismustrid) rakenduvad eesti keeles ainult kestuse ja redutseerumismustrite tunnused. Viimased võivad olla aga mõjutatud omakorda kestustest (Dicanio 2012). Saab öelda, et geminaatide ja lühikeste häälikute kategoriseerimiseks võib fortise ja leenise vastandust kasutada, kuid sõnaalgulist positsiooni ei ole võimalik olemasolevate andmete põhjal fortiseks või leeniseks määrata.

Kokkuvõtteks

Doktoritöös uuriti eesti keele sulghäälikute akustilisi tunnuseid. Töö põhieesmärk oli kirjeldada sulghäälikute allofoonilist ja kestuslikku varieerumist ja leida seda mõjutavaid tegureid. Töö tegeles viie uurimisküsimusega, mis puudutasid üldist allofoonilist varieerumist eesti keele sulghäälikutes ning mitmete segmentaalsete ja suprasegmentaalsete faktorite mõju allofoonilisele varieerumisele ning sulufaasi ja vallandumisfaasi akustikale.

Tulemused näitasid, et suures osas käituvad eesti keele sulghäälikud nii allofoonilises jaotuses kui kestuste varieerumises foneetilistele universaalidele vastavalt, kuid esineb ka omapäraseid mustreid, näiteks lühikese velaari /k/ suurem redutseerumine võrreldes mitmete teiste keeltega. Redutseerumine esines kõigis uuritud kõneregistrites. Esmakordselt uuriti töös vallandumisfaasi intensiivsuse ja spektri ehituse seoseid vältevaheldusega. Kui spektri ehitus käitus vastavalt üldistele seaduspärasustele, siis intensiivsusmustrid olid ootamatud, näidates madalamaid väärtusi geminaatides.

Töö panustas sulghäälikute akustika uurimisse uute teadmistega kõneregistrite erinevuste, geminaatsulghäälikute akustika ja vallandumisfaasi akustika kohta. Uuringutes kasutatud seotud kõne korpuseandmed ei võimalda kontrollida kõiki olulisi tegureid, mis teatud nähtusi mõjutavad, kuid siin leitud tendentsid annavad aluse edasiseks uurimiseks keerukamate meetoditega.

REFERENCES

Abramson, Arthur S. 1986. The perception of word-initial consonant length: Pattani Malay. *Journal of the International Phonetic Association* 16(1). 8–16. https://doi.org/10.1017/S0025100300003054.

Al-Tamimi, Jalal & Ghada Khattab. 2011. Multiple Cues for the Singleton-Geminate Contrast in Lebanese Arabic: Acoustic Investigation of Stops and Fricatives. In Wai Sum Lee & Eric Zee (eds.), *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII): August 17–21, 2011*, 212–215. Hong Kong: City University of Hong Kong.

Al-Tamimi, Jalal & Ghada Khattab. 2018. Acoustic correlates of the voicing contrast in Lebanese Arabic singleton and geminate stops. *Journal of Phonetics* 71. 306–325. https://doi.org/10.1016/j.wocn.2018.09.010.

Ariste, Paul. 1933a. Eesti sulghäälikud k, p, t ja b, d, g. *Eesti Keel* Nr. 3. 73–82.

Ariste, Paul. 1933b. Eesti sulghäälikud k, p, t ja b, d, g. *Eesti Keel* Nr. 6. 170–180.

Ariste, Paul. 1939. *Eesti keele hääldamine*. Tartu: Eesti Kirjanduse Selts. http://hdl.handle.net/10062/87222. (7 June, 2024).

Ariste, Paul. 1968. *Eesti keele foneetika: õpik ajaloo-keeleteaduskonna keeleosakondade üliõpilastele*. 3. tr. Tartu: Tartu Riiklik Ülikool.

Arvaniti, Amalia. 1999. Cypriot Greek. *Journal of the International Phonetic Association*. Cambridge University Press 29(2). 173–178. https://doi.org/10.1017/S002510030000654X.

Arvaniti, Amalia & Georgios Tserdanelis. 2000. On the phonetics of geminates: evidence from Cypriot Greek. In *Proceedings of the 6th International Conference on Spoken Language Processing*, 559–562.

Asu, Eva Liina, Pärtel Lippus, Karl Pajusalu & Pire Teras. 2016. *Eesti keele hääldus* (Eesti keele varamu 2). Tartu: Tartu Ülikooli kirjastus. http://hdl.handle.net/10062/57960.

Asu, Eva Liina & Pire Teras. 2009. Estonian. *Journal of the International Phonetic Association* 39(3). 367–372. https://doi.org/10.1017/S002510030999017X.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1). 1–48. https://doi.org/10.18637/jss.v067.i01.

Boersma, Paul & David Weenink. 2018. Praat: doing phonetics by computer [Computer program]. www.praat.org.

Boersma, Paul & David Weenink. 2023. Praat: doing phonetics by computer [Computer program]. www.praat.org.

Brato, Thorsten. 2015. A pilot study of acoustic features of word-final affricated /t/ and /ts/ in educated Ghanaian English. In Rita Calabrese, J. K. Chambers & Gerhard Leitner (eds.), *Variation and Change in Postcolonial Contexts*, 61–78. Newcastle: Cambridge Scholars Publishing.

Browman, Catherine P. & Louis Goldstein. 1992. Articulatory phonology: an overview. *Phonetica* 49(3–4). 155–180.

Burroni, Francesco, Raksit Lau-Preechathammarach & Sireemas Maspong. 2021. Unifying Initial Geminates and Fortis Consonants via Laryngeal Specification: Three case studies from Dunan, Pattani Malay, and Salentino. *Proceedings of the Annual Meetings on Phonology*. https://doi.org/10.3765/amp.v9i0.5200.

Cho, Taehong, Sun-Ah Jun & Peter Ladefoged. 2002. Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics* 30(2). 193–228. https://doi.org/10.1006/jpho.2001.0153.

Cho, Taehong & Patricia Keating. 2009. Effects of initial position versus prominence in English. *Journal of Phonetics* 37(4). 466–485. https://doi.org/10.1016/j.wocn.2009.08.001.

Cho, Taehong & Peter Ladefoged. 1999. Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics* 27(2). 207–229. http://dx.doi.org/10.1006/jpho.1999.0094.

Cho, Taehong & James M. McQueen. 2005. Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics* 33(2). 121–157. https://doi.org/10.1016/j.wocn.2005.01.001.

Chodroff, Eleanor, Alessandra Golden & Colin Wilson. 2019. Covariation of stop voice onset time across languages: Evidence for a universal constraint on phonetic realization. *The Journal of the Acoustical Society of America* 145(1). EL109–EL115. https://doi.org/10.1121/1.5088035.

Chodroff, Eleanor & Colin Wilson. 2014. Burst spectrum as a cue for the stop voicing contrast in American English. *The Journal of the Acoustical Society of America* 136(5). 2762–2772. https://doi.org/10.1121/1.4896470.

Clark, John, Collin Yallop & Janet Fletcher. 2007. *An Introduction to Phonetics and Phonology*. Third edition. Wiley.

Cole, Jennifer, Heejin Kim, Hansook Choi & Mark Hasegawa-Johnson. 2007. Prosodic effects on acoustic cues to stop voicing and place of articulation: Evidence from Radio News speech. *Journal of Phonetics* 35(2). 180–209. https://doi.org/10.1016/j.wocn.2006.03.004.

Cox, Felicity & Sallyanne Palethorpe. 2007. Australian English. *Journal of the International Phonetic Association* 37(3). 341–350. https://doi.org/10.1017/S0025100307003192.

Davidson, Lisa. 2016. Variability in the implementation of voicing in American English obstruents. *Journal of Phonetics* 54. 35–50. https://doi.org/10.1016/j.wocn.2015.09.003.

Davidson, Lisa. 2018. Phonation and laryngeal specification in American English voiceless obstruents. *Journal of the International Phonetic Association* 48(3). 331–356. https://doi.org/10.1017/S0025100317000330.

Davis, Stuart. 2011. Geminates. In Marc Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell Companion to Phonology*, 1–25. 1st ed. Wiley. https://doi.org/10.1002/9781444335262.wbctp0037.

Dicanio, Christian T. 2012. The Phonetics of Fortis and Lenis Consonants in Itunyoso Trique. *International Journal of American Linguistics* 78(2). 239–272. https://doi.org/10.1086/664481.

Doty, Christopher S, Kaori Idemaru & Susan G Guion. 2007. Singleton and geminate stops in Finnish – acoustic correlates. *INTERSPEECH-2007* 4.

Duez, Danielle. 1995. On Spontaneous French Speech – Aspects of the Reduction and Contextual Assimilation of Voiced Stops. *Journal of Phonetics* 23(4). 407–427. https://doi.org/10.1006/jpho.1995.0031.

Eek, Arvo. 1974. Observations on the Duration of Some Word Structures I. *Estonian Papers in Phonetics* 18–31.

Eek, Arvo. 1975. Observations on the Duration of Some Word Structures II. *Estonian Papers in Phonetics* 7–54.

Eek, Arvo & Einar Meister. 1996a. Eesti sõnaalguliste sulghäälikute akustika ja tajumine. *Keel ja Kirjandus* (3). 164–170.

Eek, Arvo & Einar Meister. 1996b. Eesti sõnaalguliste sulghäälikute akustika ja tajumine. *Keel ja Kirjandus* (4). 241–253.

Eek, Arvo & Einar Meister. 1996c. Eesti sõnaalguliste sulghäälikute akustika ja tajumine. *Keel ja Kirjandus* (5). 314–321.

Eek, Arvo & Einar Meister. 1997. Simple perception experiments on Estonian word prosody: foot structure vs. segmental quality. In Ilse Lehiste & Kristiina Ross (eds.), *Estonian Prosody: papers from a symposium. Proceedings of the International Symposium on Estonian Prosody*, 149–164. Tallinn, Estonia: Institute of Estonian Language.

Eek, Arvo & Einar Meister. 2003. Foneetilisi katseid ja arutlusi kvantiteedi alalt (I). Häälikukestusi muutvad kontekstid ja välde. *Keel Ja Kirjandus* 46(11). 815–837.

Engstrand, Olle. 1990. Swedish. *Journal of the International Phonetic Association*. Cambridge University Press 20(1). 42–44. https://doi.org/10.1017/S0025100300004060.

Ernestus, Mirjam, Iris Hanique & Erik Verboom. 2015. The effect of speech situation on the occurrence of reduced word pronunciation variants. *Journal of Phonetics* 48. 60–75. https://doi.org/10.1016/j.wocn.2014.08.001.

Esposito, Anna & Maria Gabriella Di Benedetto. 1999. Acoustical and perceptual study of gemination in Italian stops. *The Journal of the Acoustical Society of America* 106(4). 2051–2062. https://doi.org/10.1121/1.428056.

Fant, Gunnar. 1960. *Acoustic theory of speech production*. The Hague: Mouton.

Farnetani, Edda & Daniel Recasens. 2013. Coarticulation and Connected Speech Processes. In William J. Hardcastle, John Laver & Fiona E. Gibbon (eds.), *The Handbook of Phonetic Sciences*, 316–352. 2nd edn. John Wiley & Sons, Ltd. https://doi.org/10.1002/9781444317251.ch9.

Forrest, Karen, Gary Weismer, Paul Milenkovic & Ronald N. Dougall. 1988. Statistical analysis of word-initial voiceless obstruents: Preliminary data. *The Journal of the Acoustical Society of America* 84(1). 115–123. https://doi.org/10.1121/1.396977.

Fougeron, Cécile & Caroline L. Smith. 1993. French. *Journal of the International Phonetic Association*. Cambridge University Press 23(2). 73–76. https://doi.org/10.1017/S0025100300004874.

Fowler, Carol A. & Elliot Saltzman. 1993. Coordination and coarticulation in speech production. *Language and speech* 36 (Pt 2–3). 171–195. https://doi.org/10.1177/002383099303600304.

Galea, Luke. 2016. *Syllable structure and gemination in Maltese*. Universität zu Köln text.thesis.doctoral. (22 January, 2021).

Genee, Inge & Fangfang Li. 2023. Voice Onset Time in a language without voicing contrast: An acoustic analysis of Blackfoot oral stops. *Journal of the International Phonetic Association* 1–19. https://doi.org/10.1017/S0025100323000051.

Guthrie, William F. 2020. NIST/SEMATECH e-Handbook of Statistical Methods (NIST Handbook 151). National Institute of Standards and Technology. https://doi.org/10.18434/M32189.

Hamzah, Mohd Hilmi, Janet Fletcher & John Hajek. 2016. Closure duration as an acoustic correlate of the word-initial singleton/geminate consonant contrast in Kelantan Malay. *Journal of Phonetics* 58. 135–151. https://doi.org/10.1016/j.wocn.2016.08.002.

Harrington, Jonathan. 2010. *Phonetic analysis of speech corpora*. Chichester, U.K. ; Malden, MA: Wiley-Blackwell.

Harrington, Jonathan. 2013. Acoustic Phonetics. In William J. Hardcastle, John Laver & Fiona E. Gibbon (eds.), *The Handbook of Phonetic Sciences*, 81–129. 2nd edn. John Wiley & Sons, Ltd. https://doi.org/10.1002/9781444317251.ch3.

Hayward, Katrina. 2000. *Experimental Phonetics*. Harlow: Longman.

Homma, Yayoi. 1981. Durational relationship between Japanese stops and vowels. *Journal of Phonetics* 9(3). 273–281. https://doi.org/10.1016/S0095-4470(19)30971-4.

Idemaru, Kaori & Susan G. Guion. 2008. Acoustic covariants of length contrast in Japanese stops. *Journal of the International Phonetic Association* 38(2). 167–186. https://doi.org/10.1017/S0025100308003459.

Jaeger, Jeri J. 1983. The fortis/lenis question: evidence from Zapotec and Jawoñ. *Journal of Phonetics* 11(2). 177–189. https://doi.org/10.1016/S0095-4470(19)30814-9.

Jessen, Michael. 1998. *Phonetics and Phonology of Tense and Lax Obstruents in German*. Amsterdam, NETHERLANDS: John Benjamins Publishing Company.

Johnson, Keith. 2003. *Acoustic and auditory phonetics*. 2. ed., [repr.]. Malden, Mass.: Blackwell.

Jongman, Allard, Ratree Wayland & Serena Wong. 2000. Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America* 108(3). 1252–1263. https://doi.org/10.1121/1.1288413.

Kawahara, Shigeto. 2005. Voicing and geminacy in Japanese: An acoustic and perceptual study. *University of Massachusetts occasional papers in linguistics* 31. 87–120.

Keating, Patricia, Taehong Cho, Cécile Fougeron & Chai-shune Hsu. 2004. Domain-initial articulatory strengthening in four languages. In John Local, Richard Ogden & Rosalind Temple (eds.), *Phonetic Interpretation: Papers in Laboratory Phonology VI* (Papers in Laboratory Phonology), 145–163. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511486425.009.

Keating, Patricia, Wendy Linker & Marie Huffman. 1983. Patterns in allophone distribution for voiced and voiceless stops. *Journal of Phonetics* 11(3). 277–290. https://doi.org/10.1016/S0095-4470(19)30827-7.

Kisler, Thomas, Uwe Reichel & Florian Schiel. 2017. Multilingual processing of speech via web services. *Computer Speech & Language* 45. 326–347. https://doi.org/10.1016/j.csl.2017.01.005.

Kohler, Klaus J. 1984. Phonetic Explanation in Phonology: The Feature Fortis/Lenis. *Phonetica*. De Gruyter Mouton 41(3). 150–174. https://doi.org/10.1159/000261721.

Kotzor, Sandra, Allison Wetterlin & Aditi Lahiri. 2017. Bengali geminates: Processing and representation. In Haruo Kubozono (ed.), *The Phonetics and Phonology of Geminate Consonants*. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198754930.003.0009.

Kraehenmann, Astrid. 2001. Swiss German stops: geminates all over the word. *Phonology* 18(1). 109–145.

Kubozono, Haruo. 2017. Introduction to the phonetics and phonology of geminate consonants. In Haruo Kubozono (ed.), *The Phonetics and Phonology of Geminate Consonants*, 1–10. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198754930.003.0001.

Kühnert, Barbara & Francis Nolan. 2006. The origin of coarticulation. In William J. Hardcastle & Nigel Hewelett (eds.), *Coarticulation: theory, data and techniques* (Cambridge Studies in Speech Sciene and Communication). digit. print., 1. paperback ed. Cambridge: Cambridge University Press.

Kuzla, Claudia & Mirjam Ernestus. 2011. Prosodic conditioning of phonetic detail in German plosives. *Journal of Phonetics* 39(2). 143–155. https://doi.org/10.1016/j.wocn.2011.01.001.

Ladd, D. Robert & Stephan Schmid. 2018. Obstruent voicing effects on F0, but without voicing: Phonetic correlates of Swiss German lenis, fortis, and aspirated stops. *Journal of Phonetics* 71. 229–248. https://doi.org/10.1016/j.wocn.2018.09.003.

Ladefoged, Peter & Keith Johnson. 2011. *A Course in Phonetics*. sixth edition. Cengage Learning.

Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages* (Phonological Theory). Oxford, OX, UK ; Cambridge, Mass., USA: Blackwell Publishers.

Lahiri, Aditi & Jorge Hankamer. 1988. The timing of geminate consonants. *Journal of Phonetics* 16(3). 327–338. https://doi.org/10.1016/S0095-4470(19)30506-6.

Lehiste, Ilse. 1960. Segmental and syllabic quantity in Estonian. *American studies in Uralic linguistics*. Bloomington: Indiana University Publications 1. 21–82.

Lehiste, Ilse. 1966. *Consonant Quantity and Phonological Units in Estonian* (Uralic and Altaic Series 65). Bloomington: Indiana University.

Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge (Mass.) : London: MIT Press.

Lehiste, Ilse, Pire Teras, Valts Ernštreits, Pärtel Lippus, Karl Pajusalu, Tuuli Tuisk & Tiit-Rein Viitso. 2008. *Livonian prosody* (Suomalais-Ugrilaisen Seuran Toimituksia = Mémoires de La Société Finno-Ougrienne 255). Helsinki: Suomalais-ugrilainen Seura.

Lenth, Russell V. 2022. emmeans: Estimated Marginal Means, aka Least-Squares Means. https://CRAN.R-project.org/package=emmeans.

Lindblom, B. 1990. Explaining Phonetic Variation: A Sketch of the H&H Theory. In William J. Hardcastle & Alain Marchal (eds.), *Speech Production and Speech Modelling*, 403–439. Dordrecht: Springer Netherlands. http://dx.doi.org/10.1007/978-94-009-2037-8_16.

Lippus, Pärtel. 2018. Eesti keele spontaanse kõne foneetiline korpus v.1.0.4. Center of Estonian Language Resources. https://doi.org/10.15155/1-00-0000-0000-0000-00154L.

Lippus, Pärtel. 2019. Foneetikakorpuse sagedussõnastik. DataDOI. https://doi.org/10.15155/RE-62.

Lippus, Pärtel, Kätlin Aare, Anton Malmi, Tuuli Tuisk & Pire Teras. 2021. Phonetic Corpus of Estonian Spontaneous Speech v1.2. Institute of Estonian and General Linguistics, University of Tartu. https://doi.org/10.23673/RE-293.

Lippus, Pärtel, Kätlin Aare, Anton Malmi, Tuuli Tuisk & Pire Teras. 2023. Phonetic Corpus of Estonian Spontaneous Speech v1.3.

Lippus, Pärtel, Eva Liina Asu, Pire Teras & Tuuli Tuisk. 2013. Quantity-related variation of duration, pitch and vowel quality in spontaneous Estonian. *Journal of Phonetics* 41(1). 17–28. https://doi.org/10.1016/j.wocn.2012.09.005.

Lisker, Leigh & Arthur S. Abramson. 1964. A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements. *WORD* 20(3). 384–422. https://doi.org/10.1080/00437956.1964.11659830.

Maddieson, Ian. 2013. Voicing and Gaps in Plosive Systems. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. https://wals.info/chapter/5. (27 June, 2019).

Markus, Elena, Pärtel Lippus, Karl Pajusalu & Pire Teras. 2013. Three-way opposition of consonant quantity in Finnic and Saamic languages. In *Nordic Prosody. Proceedings of the XIth conference*, 225–234. Tartu: Peter Lang Verlag.

Meister, Lya. 2011. *Eesti vokaali- ja kestuskategooriad vene emakeelega keelejuhtide tajus ja häälduses. Eksperimentaalfoneetiline uurimus* (Dissertationes Philologiae Estonicae Universitatis Tartuensis 30). Tartu: Tartu Ülikooli kirjastus. https://dspace.ut.ee/handle/10062/18924.

Menzerath, P. & A. de Lacerda. 1933. *Koartikulation, Steuerung und Lautabgrenzung: eine experimentelle Untersuchung* (Phonetische Studien). F. Dümmler.

Metslang, Helle, Mati Erelt, Külli Habicht, Tiit Hennoste, Reet Kasik, Pire Teras, Annika Viht, et al. 2023. *Eesti grammatika*. Tartu Ülikooli Kirjastus. https://hdl.handle.net/10062/94933. (20 March, 2024).

Moon, Seung-Jae & Björn Lindblom. 1994. Interaction between duration, context, and speaking style in English stressed vowels. *The Journal of the Acoustical Society of America* 96(1). 40–55. https://doi.org/10.1121/1.410492.

Moran, Steven & Daniel McCloy (eds.). 2019. *PHOIBLE 2.0*. Jena: Max Planck Institute for the Science of Human History. https://phoible.org/. (3 May, 2023).

Nicolaidis, Katerina, Anna Sfakianaki, George Vlahavas & George Kafentzis. 2019. An acoustic study of Greek voiceless stops. In Sasha Calhoun, Paola Escudero, Marija Tabain & Paul Warren (eds.), *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019*, 1184–1188. Canberra, Australasia: Australasian Speech Science and Technology Association Inc.

Ohala, John J. 1983. The Origin of Sound Patterns in Vocal Tract Constraints. In Peter F. MacNeilage (ed.), *The Production of Speech*, 189–216. New York, NY: Springer. https://doi.org/10.1007/978-1-4613-8202-7_9.

Ohala, John J. 2013. The Relation between Phonetics and Phonology. In William J. Hardcastle, John Laver & Fiona E. Gibbon (eds.), *The Handbook of Phonetic Sciences*, 653–677. 2nd ed. John Wiley & Sons, Ltd. https://doi.org/10.1002/9781444317251.ch17.

Okada, Hideo. 1991. Japanese. *Journal of the International Phonetic Association*. Cambridge University Press 21(2). 94–96. https://doi.org/10.1017/S002510030000445X.

Payne, Elinor M. 2005. Phonetic variation in Italian consonant gemination. *Journal of the International Phonetic Association* 35(2). 153–181. https://doi.org/10.1017/S0025100305002240.

Piits, Liisi. 2016. Eesti Keele Instituudi kõnesünteesikorpus. Center of Estonian Language Resources. https://doi.org/10.15155/3-00-0000-0000-0000-05BDAL.

Prillop, Külli. 2015. Hääliku-, silbi- ning jalavälde: ühe nähtuse mitu tahku. *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics* 6(3). 169–195. https://doi.org/10.12697/jeful.2015.6.3.07.

R Core team. 2023. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Raasik, Liis. 2010. Intervokaalsete lühikeste klusiilide laad eesti keele spontaankõnes. Master thesis. University of Tartu, ms.

Recasens, Daniel. 2016. The Effect of Contextual Consonants on Voiced Stop Lenition: Evidence from Catalan. *Language and Speech* 59(1). 139–161. https://doi.org/10.1177/0023830915581720.

Recasens, Daniel & Aina Espinosa. 2009. An articulatory investigation of lingual coarticulatory resistance and aggressiveness for consonants and vowels in Catalan. *The Journal of the Acoustical Society of America*. Acoustical Society of America 125(4). 2288–2298. https://doi.org/10.1121/1.3089222.

Recasens, Daniel, Maria Dolors Pallarès & Jordi Fontdevila. 1997. A model of lingual coarticulation based on articulatory constraints. *The Journal of the Acoustical Society of America* 102(1). 544–561. http://dx.doi.org/10.1121/1.419727.

Reetz, Henning, Simone Mikuteit & Aditi Lahiri. 2019. Multiple functions of f0 in specifying the voicing, quantity and aspiration of East Bengali stops and affricates. In Sasha Calhoun, Paola Escudero, Marija Tabain & Paul Warren (eds.), *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019*, 2397–2400. Canberra, Australasia: Australasian Speech Science and Technology Association Inc.

Ridouane, Rachid. 2007. Gemination in Tashlhiyt Berber: an acoustic and articulatory study. *Journal of the International Phonetic Association*. Cambridge University Press 37(2). 119–142. https://doi.org/10.1017/S0025100307002903.

Roach, Peter. 2004. British English: Received Pronunciation. *Journal of the International Phonetic Association*. Cambridge University Press 34(2). 239–245. https://doi.org/10.1017/S0025100304001768.

Schuppler, Barbara, Mirjam Ernestus, Odette Scharenborg & Lou Boves. 2011. Acoustic reduction in conversational Dutch: A quantitative analysis based on automatically generated segmental transcriptions. *Journal of Phonetics* 39(1). 96–109. https://doi.org/10.1016/j.wocn.2010.11.006.

Singh, Kushagra & Nachiketa Tiwari. 2016. The structure of Hindi stop consonants. *The Journal of the Acoustical Society of America* 140(5). 3633–3642. https://doi.org/10.1121/1.4964398.

Sundara, Megha. 2005. Acoustic-phonetics of coronal stops: A cross-language study of Canadian English and Canadian French. *The Journal of the Acoustical Society of America*. Acoustical Society of America 118(2). 1026–1037. https://doi.org/10.1121/1.1953270.

Suomi, Kari & Einar Meister. 2012. A preliminary comparison of Estonian and Finnish plosives. *Linguistica Uralica* (3). 187–193 : ill. https://doi.org/10.3176/lu.2012.3.04.

Suomi, Kari, Einar Meister, Riikka Ylitalo & Lya Meister. 2013. Durational patterns in Northern Estonian and Northern Finnish. *Journal of Phonetics* 41(1). 1–16. https://doi.org/10.1016/j.wocn.2012.09.001.

Suomi, Kari, Juhani Toivanen & Riikka Ylitalo. 2008. 4 On the phonological interpretation of the quantity opposition. In *Finnish sound structure: phonetics, phonology, phonotactics and prosody* (Studia Humaniora Ouluensia 9), 39–42. Oulu: University of Oulu.

Tabain, Marija, Gavan Breen, Andrew Butcher, Anthony Jukes & Richard Beare. 2016. Stress Effects on Stop Bursts in Five Languages. *Laboratory Phonology* 7(1). UNSP 16. https://doi.org/10.5334/labphon.38.

Tabain, Marija & Andrew Butcher. 2015. Lexical stress and stop bursts in Pitjantjatjara: Feature enhancement of neutralized apicals and the coronal/velar contrast. *Journal of Phonetics* 50. 67–80. https://doi.org/10.1016/j.wocn.2015.02.004.

Torreira, Francisco & Mirjam Ernestus. 2011. Realization of voiceless stops and vowels in conversational French and Spanish. *Laboratory Phonology* 2(2). 331–353. https://doi.org/10.1515/labphon.2011.012.

Tserdanelis, Georgios & A. Arvaniti. 2001. The Acoustic Characteristics of Geminate Consonants in Cypriot Greek. In *Proceedings of the Fourth International Conference on Greek Linguistics*, 29–36. Thessaloniki: University Studio Press.

Türk, Helen. 2018. The three-way distinction of consonant duration in Estonian. In Katarzyna Klessa, Jolanta Bachan, Agnieszka Wagner, Maciej Karpiński & Daniel Śledziński (eds.), *Proceedings of the 9th International Conference on Speech Prosody 2018*, 656–660. ISCA. https://doi.org/10.21437/SpeechProsody.2018-133.
Türk, Helen. 2019a. Interactions between Segmental Context and Quantity: Temporal Patterns of Geminates in Estonian. *Linguistica Uralica* 55(4). 241–260. https://doi.org/10.3176/lu.2019.4.01.
Türk, Helen. 2019b. *Consonantal quantity systems in Estonian and Inari Saami*. Thesis. https://dspace.ut.ee/handle/10062/64625. (11 January, 2023).
Türk, Helen, Pärtel Lippus, Karl Pajusalu & Pire Teras. 2019. The acoustic correlates of quantity in Inari Saami. *Journal of Phonetics* 72. 35–51. https://doi.org/10.1016/j.wocn.2018.11.001.
Türk, Helen, Pärtel Lippus & Juraj Simko. 2017. Context-dependent articulation of consonant gemination in Estonian. *Laboratory Phonology* 8(1). 26 . https://doi.org/10.5334/labphon.117.
Vicenik, Chad. 2010. An acoustic study of Georgian stop consonants. *Journal of the International Phonetic Association* 40(1). 59–92. https://doi.org/10.1017/S0025100309990302.
Volenec, Veno. 2015. Coarticulation. In Jasmine Davis (ed.), *Phonetics: Fundamentals, Potential Applications and Role in Communicative Disorders*, 47–86. Nova Science Publishers, Inc.
Wells, John. 2005. SAMPA computer readable phonetic alphabet. *SAMPA computer readable phonetic alphabet*. http://www.phon.ucl.ac.uk/home/sampa/index.html. (12 November, 2016).
Zellers, Margaret, Barbara Schuppler & Meghan Clayards. 2018. 1. Introduction, or: why rethink reduction? In Francesco Cangemi, Meghan Clayards, Oliver Niebuhr, Barbara Schuppler & Margaret Zellers (eds.), *Rethinking Reduction*, 1–24. De Gruyter. https://doi.org/10.1515/9783110524178-001.

PUBLICATIONS

CURRICULUM VITAE

Name:	Liis Ermus
Citizenship:	Estonian
Date of birth:	January 3, 1984
Address:	Institute of the Estonian Language, Roosikrantsi 6, 10119 Tallinn, Estonia
Telephone:	+372 5805 1781
E-mail:	liis.ermus@eki.ee

Education

2018–2024	University of Tartu, doctoral studies in Estonian and Finno-Ugric Linguistics
2007–2010	University of Tartu, MA in Estonian and Finno-Ugric Linguistics
2004–2007	University of Tartu, BA in Estonian and Finno-Ugric Linguistics

Professional experience

2019–2021	Tallinn University, lecturer of "Structure of Estonian I" (phonetics)
2018–	Institute of the Estonian Language, archive specialist, junior researcher,
2016–2018	Institute of the Estonian Language, archive administrator, junior Researcher
2014–2016	Institute of the Estonian Language, archive administrator, sound technician
2011–2014	Institute of the Estonian Language, assistant, sound technician

Membership in professional organisations

2022–...	Member of the Estonian Association for Applied Linguistics
2022–...	Member of the Board of the Estonian Association for Applied Linguistics
2020–...	Member of International Speech Communication Association (ISCA)
2019–...	Member of International Phonetic Association
2011–...	Member of the Mother Tongue Society
2012–2024	Member of the expertise board of scientific collections (Estonian Ministry of Education and Research), expert of national collection of Estonian
2018–2020	Member of the Societas Linguistica Europaea

ELULOOKIRJELDUS

Nimi: Liis Ermus
Kodakondsus: eesti
Sünniaeg: 3. jaanuar 1984
Aadress: Eesti Keele Instituut, Roosikrantsi 6, 10119 Tallinn, eesti
Telefon: +372 5805 1781
E-post: liis.ermus@eki.ee

Haridus

2018–2024	Tartu Ülikool, doktoriõpe eesti ja soome-ugri keeleteaduse erialal
2007–2010	Tartu Ülikool, MA eesti ja soome ugri keeleteadus
2004–2007	Tartu Ülikool, BA eesti ja soome ugri keeleteadus

Teenistuskäik

2019–2021	Tallinna Ülikool, külalislektor aines “Eesti keele struktuur I” (foneetika)
2018–…	Eesti Keele Instituut, arhiivispetsialist-nooremteadur
2016–2018	Eesti Keele Instituut, arhiivihaldur-nooremteadur
2014–2016	Eesti Keele Instituut, arhiivihaldur-helitehnik
2011–2014	Eesti Keele Instituut, assistent-helitehnik

Teadusorganisatsiooniline tegevus

2022–...	Eesti Rakenduslingivstika Ühingu liige
2022–...	Eesti Rakenduslingivstika Ühingu juhatuse liige
2020–...	International Speech Communication Association (ISCA) liige
2019–...	Rahvusvahelise foneetikaühingu (IPA) liige
2011–...	Emakeele Seltsi liige
2012–2024	Haridus- ja Teadusministeeriumi teaduskollektsioonide ekspertnõukogu liige, eesti keele rahvuskollektsiooni ekspert
2018–2020	Societas Linguistica Europaea liige

DISSERTATIONES PHILOLOGIAE ESTONICAE UNIVERSITATIS TARTUENSIS

1. **Ülle Viks.** Eesti keele klassifikatoorne morfoloogia. Tartu, 1994.
2. **Helmi Neetar.** Deverbaalne nominaaltuletus eesti murretes. Tartu, 1994.
3. **Ülo Valk.** Eesti rahvausu kuradi-kujutelm. Tartu, 1994.
4. **Arvo Eek.** Studies on quantity and stress in Estonian. Tartu, 1994.
5. **Reet Kasik.** Verbid ja verbaalsubstantiivid tänapäeva eesti keeles. Tartu, 1994.
6. **Silvi Vare.** Nimi- ja omadussõnatuletus tänapäeva eesti kirjakeeles. Tartu, 1994.
7. **Heiki-Jaan Kaalep.** Eesti keele ressursside loomine ja kasutamine keeletehnoloogilises arendustöös. Tartu, 1998.
8. **Renate Pajusalu.** Deiktikud eesti keeles. Tartu, 1999.
9. **Vilja Oja.** Linguistic studies of Estonian colour terminology. Tartu, 2001.
10. **Külli Habicht.** Eesti vanema kirjakeele leksikaalsest ja morfosüntaktilisest arengust ning Heinrich Stahli keele eripärast selle taustal. Tartu, 2001.
11. **Pire Teras.** Lõunaeesti vokaalisüsteem: Võru pikkade vokaalide kvaliteedi muutumine. Tartu, 2003.
12. **Merike Parve.** Välted lõunaeesti murretes. Tartu, 2003.
13. **Toomas Help.** Sõnakeskne keelemudel: Eesti regulaarne ja irregulaarne verb. Tartu, 2004.
14. **Heli Laanekask.** Eesti kirjakeele kujunemine ja kujundamine 16.–19. sajandil. Tartu, 2004.
15. **Peeter Päll.** Võõrnimed eestikeelses tekstis. Tartu, 2005.
16. **Liina Lindström.** Finiitverbi asend lauses. Sõnajärg ja seda mõjutavad tegurid suulises eesti keeles. Tartu, 2005.
17. **Kadri Muischnek.** Verbi ja noomeni püsiühendid eesti keeles. Tartu, 2006.
18. **Kanni Labi.** Eesti regilaulude verbisemantika. Tartu, 2006.
19. **Raili Pool.** Eesti keele teise keelena omandamise seaduspärasusi täis- ja osasihitise näitel. Tartu, 2007.
20. **Sulev Iva.** Võru kirjakeele sõnamuutmissüsteem. Tartu, 2007.
21. **Arvi Tavast.** The translator is human too: a case for instrumentalism in multilingual specialised communication. Tartu, 2008.
22. **Evar Saar.** Võrumaa kohanimede analüüs enamlevinud nimeosade põhjal ja traditsioonilise kogukonna nimesüsteem. Tartu, 2008.
23. **Pille Penjam.** Eesti kirjakeele *da*- ja *ma*-infinitiiviga konstruktsioonid. Tartu, 2008.
24. **Kristiina Praakli.** Esimese põlvkonna Soome eestlaste kakskeelne keelekasutus ja koodikopeerimine. Tartu, 2009.
25. **Mari Mets.** Suhtlusvõrgustikud reaalajas: võru kõnekeele varieerumine kahes Võrumaa külas. Tartu, 2010.
26. **Karen Kuldnokk.** Militaarne retoorika. Argumentatsioon ja keeleline mõjutamine Eesti kaitsepoliitilises diskursuses. Tartu, 2011.

27. **Kai Tafenau.** Uue Testamendi tõlkimisest Rootsi ajal: käsikirjad, tõlkijad ja eesti kirjakeel. Tartu, 2011.
28. **Külli Prillop.** Optimaalsusteoreetiline käsitlus eesti keele fonoloogilisest kujunemisest. Tartu, 2011, 261 lk.
29. **Pärtel Lippus.** The acoustic features and perception of the Estonian quantity system. Tartu, 2011, 146 p.
30. **Lya Meister.** Eesti vokaali- ja kestuskategooriad vene emakeelega keelejuhtide tajus ja häälduses. Eksperimentaalfoneetiline uurimus. Tartu, 2011, 145 lk.
31. **Kersti Lepajõe.** Kirjand kui tekstiliik. Riigieksamikirjandite tekstuaalsed, retoorilised ja diskursiivsed omadused. Tartu, 2011, 141 lk.
32. **Tiit Hennoste.** Grammatiliste vormide seoseid suhtlustegevustega eestikeelses suulises vestluses. Tartu, 2013, 236 lk.
33. **Helena Metslang.** Grammatical relations in Estonian: subject, object and beyond. Tartu, 2013, 443 p.
34. **Kristel Uiboaed.** Verbiühendid eesti murretes. Tartu, 2013, 227 lk.
35. **Helen Plado.** Kausaalsuhete adverbiaallaused eesti keeles. Tartu, 2013, 244 lk.
36. **Annika Küngas.** Pragmaatiliste markerite kujunemine ja funktsioonid eesti keeles *lt*-sõnade näitel. Tartu, 2014, 200 lk.
37. **Maarika Teral.** Arvutipõhine eesti keele õpe: vahendid ja hinnangud nende efektiivsusele Tartu ülikooli keelekursuste näitel. Tartu, 2015, 175 lk.
38. **Anni Jürine.** The development of complex postpositions in Estonian: a case of grammaticalization via lexicalization. Tartu, 2016, 315 p.
39. **Kairi Tamuri.** Basic emotions in read Estonian speech: acoustic analysis and modelling. Tartu, 2017, 238 p.
40. **Sven-Erik Soosaar.** Eesti keele kultuursõnavara arengujooni – eesti keel kui piirikeel. Tartu, 2017, 157 lk.
41. **David Ogren.** Object case variation in Estonian *da*-infinitive constructions. Tartu, 2018, 212 p.
42. **Reet Hendrikson.** Kas sõjasõna sünnib sõtta? Erialakeele tõhusus sõjandusterminoloogia näitel. Tartu, 2018, 252 lk.
43. **Eva Velsker.** Ajaväljendid eesti murretes. Tartu, 2018, 216 lk.
44. **Riina Reinsalu.** Juhendavad haldustekstid žanriteoreetilises raamistikus. Tartu, 2019, 168 lk.
45. **Ann Siiman.** Vormikasutuse varieerumine ning selle põhjused osastava ja sisseütleva käände näitel. Tartu, 2019, 209 lk.
46. **Tiina Laansalu.** Kesk-Harjumaa kohanimede kujunemine ja nimesiire. Tartu, 2020, 216 lk.
47. **Mariko Faster.** Hargla kihelkonna vetenimed Eesti-Läti piirialal. Tartu, 2020, 218 lk.
48. **Maarja-Liisa Pilvik.** Action nouns in a constructional network: A corpus-based investigation of the productivity and functions of the deverbal suffix *-mine* in five different registers of Estonian. Tartu, 2021, 252 p.

49. **Hanna Pook.** Pronoun use and variation in Estonian dialects: *kes* ‘who’, *mis* ‘what’ and *keegi* ‘someone’. Tartu, 2023, 225 p.
50. **Andra Annuka-Loik.** Naer ja naerev hääl eesti suulises suhtluses. Tartu, 2023, 202 lk.
51. **Tiina Paet.** Võõrainese kinnistumine eesti keeles: keelekorralduslik ja leksikograafiline vaade. Tartu, 2023, 260 lk.
52. **Lydia Risberg.** Sõnatähendused ja sõnaraamat. Kasutuspõhine sisend eesti keelekorraldusele. Tartu, 2024, 286 lk.