UNIVERSITY OF TARTU
FALCUTY OF ARTS AND HUMANITIES
INSTITUTE OF ESTONIAN AND GENERAL LINGUISTICS

Zhenyu Wang

Lexical selection between synonymous medical terms

Master’s thesis

Supervisor: Professor (PhD) Virve-Anneli Vihman

Tartu 2025

Affirmation of authorship

I confirm that I have written the present thesis myself and have correctly cited the contributions of other authors. The work was written based on the thesis requirements of the Institute of Estonian and General Linguistics of the University of Tartu and is in conformity with the academic practice.

Zhenyu Wang

Abstract

Previous research reveals that the frequency effect plays a role in the lexical selection mechanism; however, it can be assumed that the frequency effect is an important yet restrictive factor in this process. Based on the usage-based theory, this experimental master's thesis mainly focuses on two questions with a language corpus analysis and an experimental method: (i) how medical students make lexical selections between native and foreign terms; (ii) whether the lexical selection mechanism is competitive or not. Based on the corpus-based analysis, 30 pairs of terms were selected and 300 test samples were prepared. During the experiment, I asked the participants to make a choice and produce a speech as soon as an audio recording containing two synonymous terms and a description of the concept denoted by the two ends. The subjects' speech production following the lexical selection was analysed with descriptive statistics and a logistic model. Based on the result of analysing data from a small sample (N = 10), frequency ratio is the only significant factor in predicting lexical choices. The results are consistent with the assumption in the existing literature that other factors also influence the effect of frequency. The results also revealed that foreign terms do not necessarily pale in comparison to native terms in terms of accuracy. Considering both the sample size and the differences between students, it cannot be ruled out that there are both Estonian-word-preferring language users and foreign-word-preferring language users. Future research on the same topic should pay attention to improving the methods, with inherent differences among participants and variations in term frequencies in corpora under the influence of differences between texts taken into account.

Key words: usage-based theory, lexical selection, medical terms, corpus-based frequency, synonyms

Aknowledgement

I first appreciate the support from my supervisor Virve-Anneli Vihman. Virve-Anneli Vihman always provided constructive comments to my wirting. She is also supportive when I got stuck in addressing issues in conducting research. I feel grateful for her suggestions and recommended books and articles. Without her encouragement and help, this work would not have been finished. Secondly, I feel thankful for the help from junior researcher Jesse Holmes. He assisted me in dealing with technical issues. Great thanks to Kulla Mellov, Mare Kõiva and Kasper Raud — they assisted in reviewing all the texts used in my experiment! Besides, I will not forget the devotion to recording stimuhli from Äivi Soo, Helena Brit Rosenberg, and Margo Huik, and pleasant moments in recording with them. Without their help, experiment would not have progressed smoothly.

Table of Contents

Introduction

This study addresses the question of lexical selection in the L1 language use of Estonian medical students. However, the study was inspired by the idea of exploring the association between code-switching and foreign morphemes-based vocabulary in the use of specialised language (e.g., computer science, industrial design, or other domains with specific professional or academic purposes). Research into bilingualism (e.g., Belhoul & Benmoussat, 2023; Gotti, 2015) has seen the role of code-switching in the speech where specialised language is frequently used. The initial idea emerged of observing code-switching triggered by orthographic and/or phonological overlap among specialised language users based on the bilingual mode hypothesis (Grosjean, 2013: 18) and triggering hypothesis (Clyne, 2003). However, Estonian is a language characterised by a long history of language contact with other languages. Adapting loanwords into Estonian has been an important task for language planners since the beginning of Estonian language planning (Paet, 2023: 46-50). Borrowing is universal across languages (Tadmor, 2009). Also, according to Deuchar (2020), borrowing is assumed to follow a period of bilingual language use incl. code-switching. Thus, it should be under review that borrowing is comparable with code-switching in terms of bringing in another language (cf. Grosjean, 2013: 18).

In prior research on bilingualism, most investigated cognates or loanwords refer to concrete concepts other than abstract concepts referred to by specialised terms. In Estonian, the parallel of terms based on native morphemes to those based on international terms borrowed from classical languages exists (Erelt, 2007: 119); this is characteristic of many specialities such as medicine science. Though foreign words share orthographical and/or phonological features with terms in donor languages, few studies have investigated the actual use of the two types of terms. Estonian linguists also hold various attitudes towards which term should be encouraged to use (cf. Erelt, 2007; Rätsep 2002 [1964]; Saari, 1981). Arguments centre around the advantages of word formation, comprehensibility, and accuracy of meaning; studies on the language use of medical terms in Estonian need empirical investigation. Apart from the frequency which has been long assumed to have a main effect on the mechanism of making lexical selections (cf. Dylman & Barry, 2017; Gollan et al., 2008; Kroll et al., 2006), recent research has also revealed other factors that may interact with the frequency.

Thus, the initial idea changed into a more practical research plan to investigate lexical selection between synonyms in medicine science. The core of research questions must turn to the use of lexemes not only for the consideration of feasibility but also due to the urgent need to explore and unravel the long unresolved problem, yet fundamental to other further investigations — as to how specialised language users make selections between parallel lexemes based on different morphemes.

1. Background

1.1. Lexical selection between synonyms

According to the three-stage process of speech production (based on the speech production model by Levelt 2000, Levelt 2001, as cited in Grosjean, 2013), lexical selection arises from the process of organising lexical concepts for the oral expression. The speaker in the following formulation stage is assumed to select word forms from the mental lexicon (perhaps a language-specific (Myers-Scotton,2007: 296-299) or a register-specific (Adamska-Salasiak, 2013) associated with lexical concepts before further steps in the next level. I will first review several thoroughly investigated models regarding lexical selection and competition of bilinguals because prior research has produced fruitful results around this topic, incl. the cognate facilitation effect (Costa et al., 2000, as cited in Santesteban & Schwieter, 2020: 4-5), weaker-link hypothesis (Gollan, et al., 2008) and bi-directional translation connection model (Dylman & Barry, 2017).

The discussed lexical competition herein consists of (i) the selection between two or more words from different languages (e.g., *strawberry* from English vs *fresa* from Spanish; cf. Sarkis & Montag, 2021: 168) for bilingual speakers — the cross-language competition; (ii) the selection between synonyms (e.g., *sofa* vs *couch*) irrespective of how many languages a speaker uses. If the mechanism is competitive, lexical selection requires more time to deal with the competition between a lexical candidate and its alternative. By contrast, a non-competitive mechanism will cease the further activation of other alternatives as soon as one candidate is activated and selected (Oppenheim, 2023). Among studies of cross-language competition, Costa et al. (2000) argued for a non-competitive model — the language-specific selection model. When the selected language is determined between conceptualisation and formulation stages, the lexical representation from the non-target language will impose limited influence on the selection mechanism. The further proposed cognate facilitation effect corroborates this model. This hypothesis does not deny the possible occurrence of competition on the phonological level (Costa et al., 2000; Kroll et al., 2006). Competition may occur at the phonological level due to the phonological forms co-activated by the phonological overlap between cognates. However, if competition exists, what can account for the facilitation effect?

One explanation is the global inhibitory control system (Kroll et al., 2006). Once the target language is determined, a bilingual relies on this system to inhibit the further activation of the non-target language. This inhibitory control system is asymmetrical (Kroll et al., 2006;

Christoffels et al., 2007); it requires more cost to inhibit the dominant L1 when shifting from L1 to L2. This property implies that the lexical selection is *de facto* competitive. However, other studies (e.g., Baus et al., 2013) have questioned this proposal and argued for the role of frequency and cognates as the discovered inhibition works only when words are either less frequently used or non-cognates. Gollan et al. (2008) further proposed the weaker-link hypothesis. They assumed that bilinguals may use words from L1 and L2 for given concepts less frequently than monolinguals who have no concern about acquiring, storing, processing and using words from two languages. As a result, the link between semantic and phonological representations for vocabulary in two languages becomes weaker. After these words undergo long-term use or training, the cost of inhibiting the non-target language should decrease significantly.

Compared with research into cross-language competition, studies concerning selections between synonyms within a language are fewer. Dlman and Barry (2017) also found the frequency effect in the competition between a common name (the more frequently used word in a synonym pair) and an alternative name (the less frequently used word in the same pair). They argued that it is analogous to the competition between one word in L1 and another in L2. The difference in frequency leads to there being no interference from the alternative lexeme when a speaker names the picture with a common name.

Regardless of the various topics, the existing literature revolves around the frequency effect. According to the usage-based theory (Bybee, 2007; Langacker, 1987; as cited by Divjak, 2019: 17, 41), the frequency effect is universal and critical to the development of single words and other aspects of a language (Divjak, 2019). Yet it is noteworthy that the frequency investigated in prior research (Dylman & Barry, 2017; Gollan et al., 2008) is limited to single-word frequency or the concept embodied in a word frequency list (Divjak, 2019: 20-25). For instance, both Dylman and Barry or Gollan et al. gathered words before experiments and recorded frequency with a subjective approach (Divjak, 2019: 25). By contrast, frequency effect in the usage-based theory, though yet playing an important role, is never a universal cure that can be used to account for all cases (Baayen, 2010; Divjak, 2019: 17). Other factors interacting with the frequency of L2 have been listed such as proficiency of L2 (Blumenfeld et al., 2016), morphological patterns (Baayen et al., 2007), age of acquisition (Baayen et al., 2016), etc. That is, when the frequency is investigated in a specific context, its effect may be fostered or lessened by other factors (Divjak, 2019: 27-29).

If the methodology in the work by Dylman and Barry is viewed from a usage-based perspective, the formation of common-and-alternative name list is less satisfying. Firstly, Dylman and Barry did not pay sufficient attention to other factors that would influence single word frequency. For instance, if the topic is the lexical selection among young people, it is questionable to include *swine*, which is archaic and a term for farming, as an alternative name to *pig*. Based on Bybee's exemplar theory (Bybee, 2007; as cited in Divjak, 2019: 42-45), in this case, the lower frequency of *swine* can be ascribed to the need to create a category for the association between the distinct item *swine* and the certain construction where *swine* occurs. By contrast, a general category is enough to store a construction with which any commonly used words such as *pig* can fit. It is questionable to name *swine* as the alternative to *pig* without considering the context. Other examples that need reviewing in their word list include *dog* vs *hound*, *donkey* vs *mule*, *frog* vs *toad*, etc. Granted, completely interchangeable synonyms are *per se* rare (Adamska- Sałaciak, 2013: 333-334), but it is necessary to define clearly synonyms before investigation.

To resolve these issues, in the present research, synonym is defined based on *equivalents* (Adamska-Sałaciak, 2013: 336-337) and *justified parallelism* (Est. *õigustatud rööpsus*) as words whose parallelism is accepted in common practices or justified in studies in a specific linguistic area according to the largest proximity of meanings at the same level of a semantic category; the mutual interchangeability is susceptible to register differences. It is normal for a term to have multiple meanings in various domains, whereas the alternative name has only one meaning in one specific domain. The ideal synonymous terms this study is aimed to investigate do not include polysemes, however.

1.2. Lexical selections between medical terms

Medical terms in this study denote an open class of specialised vocabulary referring to medical concepts. Under the written language, specialised language has an indispensable role in reflecting scientific thinking processes and is used in various specific pragmatic environments as a register (Erelt, 2007). By contrast, general language herein coheres with Erelt's definition — it forms the core of both spoken and written language (Erelt, 2007: 18-19). On the one hand, specialised language places general language in the shade due to its advantage in the accuracy of expressions and accurate differentiation between concepts (Erelt: 2007: 35). On the other hand, the boundary between specialised language vocabulary and general language vocabulary is inconstant and unclear. The two-way communication of vocabulary between the two registers is an ordinary phenomenon. Terms form based on general language vocabulary; specialised vocabulary can also become popularised and widely used in general language (Erelt, 2007: 22-24).

According to the operationalisation herein, *specialised language/term* is preferred to *technical language/term* because the pragmatic contexts for terms are far wider than institutions, academic circles, industries, etc. Moreover, jargon, another branch of specialised language vocabulary, will not be covered herein as jargon refers to vocabulary designed for a particular group such as experts in medicine or other domains, whether based on definitions in dictionaries or from a linguistic perspective (Krieger & Gallois, 2016; Shulman et al., 2020). Jargon's audience, in other words, does not include laypeople.

Medical terms may feature foreignness in structure and origin. Medical science is a typical speciality where English dominates as a lingua franca in medical literature (its status as a lingual franca dates to the 1950s; Baethge, 2008; Van Weijen, 2012, as cited in Džuganová, 2019). English has become the main donor language in lexical borrowing in many languages, albeit with variations among languages. Before this change, to mark medical concepts, borrowing vocabulary or morphemes directly from Latin and Ancient Greek or via mediator languages such as French and German was the most common way to form medical terms in languages used in Europe. To form medical terms, coining medical terms with native lexemes and/or roots makes up another important approach besides borrowing. For instance, in modern English, terms based on native or Germanic roots: *blood clot*, *blood island*, etc., exist, whereas 95 per cent of medical vocabulary is assumed to originate from classical languages (e.g., Belialeva et al., 2017, as cited in Džuganová, 2019).

The contrast between borrowings and native words (NB considering words borrowed via calque, native word herein denotes words based on native morphemes other than words inherited from an earlier stage of the language, not simply the opposite to *borrowing*) is more significant in, for instance, German than in English. Jucks and Paus (2011) draw a line between classical language terms and German language terms (e.g., *Verbrauch* 'use' vs *Konsum* 'consumption', *Spritze* 'jab/shot' vs *Injektion* 'injection', etc.) and advance an experiment-based proposal that German language terms are more comprehensible, compared with terms borrowed from classical languages. Recall the two situations of vocabulary flow between general language and specialised language: (i) Terms form based on general language vocabulary; (ii) specialised vocabulary can also become popularised and widely used in general language (Erelt, 2007: 22-24). The proposed association between word formation and cognitive comprehension (Jucks & Paus, 2011) may contribute to understanding the two situations. Firstly, forming new medical terms based on native morphemes/words or adding new meanings to a native word that is commonly used and understandable in everyday speech. Secondly, popularising medical terms with the advantage of the comprehensible word formation of native morphemes.

However, it is questionable that native morphemes always have advantages over foreign morphemes in cognitive comprehension. Vocabulary borrowing is a common phenomenon and a major aspect of language contact (Tadmor, 2009: 55). We cannot ignore the likelihood that the comprehensibility of foreign morphemes gradually increases when those morphemes are adapted, developed and accepted through long-term use. Thus, foreign morphemes are not necessarily estranged from native speakers of a specific language. In case (ii) above, it is also possible for medical terms based on foreign morphemes to become popularised. In other words, both native terms and foreign terms can enter the spoken language.

The question addressed is whether native words are always preferred in medicine science. If the distinction between the two registers, the spoken language and written language, is constant and firm, based on the finding in the article by Jucks and Paus and dominant-L1 effect in the research into bilingual speech production (Grosjean, 2013), one reasonable assumption is that native terms are always preferred. However, this assumption becomes debatable given the influence of written language on spoken language vocabulary diversity (Berman & Nir, 2010; Chafe, 1994: 42-45; Shi & Lei, 2021; Strömqvist et al., 2002). Strömqvist et al. (2002) proposed in his work that the written language acquired via reading and writing would significantly shape the spoken language as an individual's literacy and cognitive abilities develop. Experimental findings also show that this influence can be instantaneous (Berman & Nir, 2010). Another factor can be the speaker's background. In discussing the code-switching between L1 and L2, Grosjean (2013) proposes a language continuum between L1 and L2. By adapting this continuum to the present study, I assume that the speaker's language background of language skills and the listener's are both likely to influence the selection between more native and less native lexemes. Accordingly, the lexical selection made by doctors, physicians, or medical students may vary significantly with interlocutors and environments.

1.3. Medical terms in Estonian context

Medical terms in Estonian nowadays also undergo the influence of English (Sinivalu & Soosaar, 2013). The recent practice of terminology in medicine has displayed a tendency to borrow terms with Latin and/or Old Greek origins from English according to English pronunciation (Erelt, 2007: 222). The Estonian language also has a long tradition of borrowing medical terms from classical languages via Romance languages or from German (Paet, 2023: 62, 64), thereby adopting a comparable system for medical terms to that in the research by Jucks and Paus — classical word vs Estonian word. Estonian linguists have established rules to distinguish *native words* (Est. *omasõna*) from *foreign words* (Est. *võõrsõna*) which have been borrowed recently. The present study adopts the *foreign word* instead of *the classical loanword* or *borrowing* and the *Estonian word* rather than the *native word/term*. The reason for the latter is clarified in the last section. Foreign words in Estonian denote loanwords not adapted completely (either not adapted or partially adapted) to the Estonian phonological system. As a main branch of loanwords, foreign words come from language families outside Uralic, characterised by the remote origin (cf. structural-based and origin-based foreignness in Soosaar, 2020: 350-256).

Many foreign medical terms share common features with other foreign words, such as word-initial syllables starting with *b*, *d*, *g*; *o* in non-initial syllables; long vowels in non-initial syllables, and others (cf. Paet, 2023; Viitso, 2003: 21-23). They are also characterised by the foreign suffixes borrowed together with the lexical item. Kasik (2015: 75) listed 34 foreign suffixes (e.g., *-iiv* '-ive' < ? < **-īvus*, *-teet* '-ty' < ? < **-tās*, etc.). Common suffixes for medical terms include *-iit* in *hepatiit* 'hepatitis', *-oos* in *skleroos* 'sclerosis' (this is the one not listed in that list of 34 foreign suffixes by Kasik), -ia in *aneemia* 'anaemia'. If the borrowed word is an adjective, it is adapted to the Estonian morphological system, attached to the suffix *-ne* or *-line* such as *bioloogiline* "biological". Foreign verbs also undergo the same adaptation because the dictionary form of the Estonian verb is *ma*-ending. Roots of foreign verbs are often modified to *eeri*- or *i*-ending (e.g., *kontrollima* "to check", *skaneerima* "to scan", etc.).

Though there have been few studies focusing on the use of the Estonian medical terms, terminological research findings provide food for thought. Based on the usage-based theory, by listing the present usage of 243 foreign words in the earliest *EKÕS 1918* (*Eesti keele õigekeelsussõnastik 1918*) 'orthological dictionary of Estonian language 1918', Paet (2023: 73) pointed out that nowadays it is foreign words rather than their native counterparts that have become fixed in usage. However, not all the foreign words investigated by Paet belong to specialised vocabulary. Saari (1981: 207, 282-285, as cited in Erelt, 2007: 140) proposed the argument that native words have the advantage in definition accuracy (Est. *täpsusvõime*) — especially in medicine science — and word formation productivity (Est. *moodustusvõime*). By contrast, Rätsep (1964/2002) agrees with Saari on the advantage of native derivatives and compounds in conveying meaning to ordinary Estonian language users, on the one hand. On the other hand, he disagrees as to what the disadvantages of foreign words are accordingly assumed to be. Rather, he affirms the merits of foreign words — their international comprehensibility and more accurate fixation of meaning. Erelt (2007: 144-148) reminded readers from a more practical viewpoint that it is advisable to consider the scope of the use of the term, while it is also unnecessary to struggle for native words.

1.4. Limitations of methodology previously used

In discussing the exploration of lexical selection, a topic that cannot be ignored is the way of measuring frequency. Most prior research investigates concrete words rather than abstract concepts (e.g., Dylman & Barry, 2017; Gollan et al., 2008; Sarkis & Montag, 2021). Among methodologies previously used, to measure single-word frequency and elicit lexical selection, one traditional and frequently used method is picturing naming (e.g., Dylman & Barry, 2017; Sarkis & Montag, 2021). Also, recall that limitations of measuring single-word frequency (e.g., Raymond & Brown, 2012; Baayen et al., 2016: 1207), picture naming is never valid to measure the frequency of medical terms. Considering the mutual influence of vocabulary between the written language and spoken language, and between the general language and the specialised language, measuring the frequency of medical terms in written language should be considered.

Among methods in previous studies, the most suitable for investigating abstract concepts is the 'telephone-chain study' used by Grosjean (2008). In his research into the influence of the interlocutor's background on the speech by bilinguals, participants were required to retold or describe what they had watched from cartoons of various topics to interlocutors (who did not necessarily show up) of different backgrounds. Thus, it is acceptable to play the audio stimulus including medical synonyms and a description of the concept referred to by the terms to elicit lexical selections. Accordingly, since the audio stimuli are aimed at concretising abstract concepts by playing explanatory texts, the present study is also aimed to explore whether a certain homogeneous structure will be primed if all the audio stimuli maintain homogeneity to a certain degree.

The priming effect denotes the process of repeating recently used syntactic structures or lexical choices when speaking. This effect has been investigated in many studies on speech production (Kootstra, 2020; Pickering 2004; Pickering, 2008). Apart from the effect of syntactic structure, another component that may be sensitive to priming would be the use of other foreign words, if the only foreign word in each of the audio stimuli is that foreign term. Moreover, previous studies have also noted the fuzzy distinction between foreign words or borrowing and single-word code-switching. Erelt (2007: 136-137) has pointed out a phenomenon of misusing fake foreign words (Est. *pseudo-* or *libavõõrsõna*) — due to limited English and other language skills, one uses English words as if they are authentic foreign words. More interestingly, in Estonian, quoted words (Est. *tsitaatsõna*) form another branch of the loanword family. The quoted word refers to a borrowed word, phrase and expression used in Estonian with the original form in the donor language intactly preserved. Its pronunciation is prescribed to be close to that in the donor language, whereas it is allowed for speakers to pronounce it based on the Estonian phonology (Paet, 2023: 14).

However, from the viewpoint of Myers-Scotton (2006: 253) and Haspelmath (2009: 41), some quoted words are rather comparable with single-word code-switching. Myers-Scotton proposed a 4M model to account for why and how this type of word is morphologically integrated into the matrix language. Haspelmath (2009: 41) proposed an intermediate stage where some words from other languages lie between established borrowings and code-switching because adaptation and integration of loanwords is a long and dynamic process. Though Hapspelmath's proposal appears to settle the nature of quoted words, this is not what the present study is intended to investigate. This research queries whether another language is brought in because of phonological and/or orthographical overlap (Clyne, 2003: 164; Kootstra et al., 2020) shared by the foreign word and its equivalent in a donor/mediator language, regardless of how that language is brought in — in the form of a quoted word or single-word code-switching.

1.5. Research questions

The research is centred on two research questions. The first research question queries whether the corpus-based frequency of medical terms or the foreignness in word formation is the strong predictor of lexical selection when medical students are required to make acoustic decisions based on audio stimuli. I hypothesise that the role of frequency would be strong yet restricted insofar as it pales by comparison with word formation. Participants would, thus, make selections with a preference for Estonian words due to the L1 dominant effect and fewer cognitive loads required for comprehension. A further sub-question focuses on the competitiveness of the mechanism of lexical selections. The competitiveness should emerge in the process of making decisions between synonyms in which neither Estonian word nor foreign word is the dominant one, based on the corpus data on principle. There should be variation in initial latency. Intial latency denotes the time used from the end of stimulus to the beginning of speech production, which can be either the voicing onset or the beginning of release, depending on the initial phoneme (cf. Sarkis & Montag, 2021: 169). However, I expect the mechanism to be uncompetitive if the structural foreignness in word formation plays a role.

The second research question is centred on speech production following the lexical selection. I hypothesise that lexical or syntactic representations may be primed if audio stimuli maintain homogeneity to a certain degree and differ in only stimulus term pairs. A further interesting question arises as to whether another language is brought in if many other foreign words are also used due to lexical representation priming.

2. Method

Methods consist of a corpus-based frequency analysis and an elicitation experiment. I plan to perform a corpus-based analysis to fulfil these goals: (i) to find the pairs of terms that are in actual usage; (ii) to sift out stimuli by calculating their relative frequencies against each other (e.g., Estonian word to foreign word, or foreign word to Estonian word).

Since the planned experiment includes tasks to elicit participant responses, we sought ethical approval from the Research Ethics Committee of the University of Tartu. The study received approval on the 4th of March 2025, protocol number 395/T-4. Based on an agreement between the investigator and the ethics committee, all the participants should be recruited without compensation. The investigator is obliged to inform volunteer candidates of the voluntariness of participation. Before an experiment (including practice trials before testing trials), the participant is obligated to read the consent form, while the investigator stands by the side to help to explain if any questions emerge. Without the consent form signed unforcedly by both sides, the experiment cannot start.

2.1. Corpus study for stimulus selection

The corpus-based frequency analysis covers these tasks: (i) forming the principles of determining "stimuli candidates"; (ii) forming an initial list of medical term pairs; (iii) extracting and comparing corpus frequencies; and (iv) determining pairs for stimuli by cleaning and analysing data.

The first and second tasks proceeded simultaneously. To cast a wide net, sources of medical terms were first decided: (i) *Principes of anatomy and physiology*; (ii) *Vikipeedia meditsiini mõisteid* (Eng. 'Wikipedia Medical Concepts'); (iii) *Väike inglise-eesti-vene meditsiinisõnastik* 'Little English-Estonian-Russian Medical Dictionary; (iv) *Sõnaveeb*. Among these materials, material (iii) is not only the relatively newest and comprehensive publication of the indexed medical terms in three languages but also the suggested reading in the course *ladina keel meditsiinierialadele* 'Latin language for medical speciality (FLEK.05.008)'. Material (ii) covers more terms and hyperlinks to the page of medical concepts. Material (iv) is not medicine-oriented; however, it covers more than 135 specialised language databases.

Based on the operationalisation of medical terms and synonyms described in the ***Background*** section, I used flexible criteria to sift terms. Firstly, nouns are preferred to adjectives and verbs. Secondly, an ideal synonym proper to the experiment design should consist of only one Estonian word and only one foreign word. No matter how much information about an item in a dictionary or on web pages, it will not be considered if its foreign or Estonian equivalent cannot be found. Thirdly, if more than two variants exist to refer to the same concept, the one whose frequency in the corpus-based analysis is the highest is undeniably the most proper. Fourthly, words will not be considered if they are not included in *Sõnaveeb* yet found in other materials. These kinds of words are either (i) ones whose independent webpages on Wikipedia are absent or (ii) ones whose status cannot be determined — it is unclear whether they are phrases or compounds, yet they appear to look like a phrase with the gap between components. In situation (ii), words include *kõre-hingetoru põletik* 'laryngotracheitis', *vikerkesta eemaldus* 'iridectomy', etc. It might be confusing as the final list includes the word *luuhõrenemine* 'osteoporosis' and *südamepekslemine* 'tachycardia', both of which can however be written as *luu hõrenemine* and *südame pekslemine*. However, how to distinguish compounds from phrases is not the focus of the study. Considering the most economical way, the words included in *Sõnaveeb* are considered as a standard for comparison. Fifthly, words were ruled out if they have other definitions in domains outside medicine science such as *eemaldamine*. Sixthly, words will be excluded if their definition and usage are archaic and/or prejudicial such as *nymphmaania*.

Establishing a sub-corpus took longer than expected. It appears clear that pure corpora of medical written texts in Estonian are nowhere to be found except for Korp. However, Korp includes insufficient medical written texts to do further statistics. The amount of data on Korp is also much lower than on Estonian National Corpus 2023 (ENC 2023). Though ENC 2023 covers limited data on medical science, the present study does not assume that medical students acquire or familiarise themselves with medical concepts only from academic reading and textbooks. The higher frequency of a medical term in written texts, without respect to topics or sources (e.g., books, news, TV programmes, etc.), at least can be an indicator of a higher probability that a speaker will have been exposed to those terms in their life. ENC 2023 is preferred in the present research, though the approximation of students' exposure to medical terms is imperfect.

The sampled corpora constitute four sub-corpora: Corpora of Academic Texts, Balanced Corpus 1990-2008, Reference Corpus 1990-2008, and Timestamped 2014-2023. Only Corpora of Academic Texts, which makes up for 0.291% of ENC 2023, contain scholarly writing (*Computer Linguistics*, n.d.-a). Both Balanced Corpus 1990-2008 and Reference Corpus 1990-2008 are miscellaneous collections of written texts, including only a small part of academic or scientific writing (5 million words in Balance Corpus 1990-2008 and 3.6 million words in

Reference Corpus 1990-2008) (*Computer Linguistics*, n.d.-b; *The Mixed Corpus of Estonian*, n.d.). By contrast, Timestamped 2014-2023 are sheer corpora of news articles (Lexical Computing CZ s.r.o., 2025). The concordance function of Sketch Engine provides data on the number of occurrences of medical terms on the list. Many terms were ruled out due to zero count in the corpus. The final data frame includes 139 pairs of Estonian-and-foreign words.

Then, the frequency distribution of two types of terms was plotted via Python Pandas and Matplotlib packages. Two new columns were added to the dataset: (i) the relative frequency of foreign words and (ii) the relative frequency of Estonian words, to store the proportion of each term in the sample of 579543713 tokens (automatically calculated on Sketch Engine,). Based on comparisons of relative frequency, 139 pairs of terms can be divided into Estonian-word dominant Estonian-word set (DE set in the following text) (*min.* = 0.005, *max.* = 16.741, *MD* = 0.2725) and dominant foreign-word set (DF set in the following text) (*min.* = 0.012, *max.* = 10.367, *MD* = 0.573). After ranking these two subsets, 31 pairs out of 62 Estonian-dominant pairs and 38 pairs out of 77 Foreign-dominant pairs remain for further sifting according to the median.

The further sifting began with the calculation of frequency ratio: (i) the frequency of occurrence of an Estonian word against that of its foreign equivalent in the DF set; (ii) the frequency of occurrence of a foreign word against its Estonian equivalent in the DE set; both should between 0 and 1. When the frequency of occurrence of an Estonian word is approximate to that of its foreign equivalent, both ratios (i) and (ii) should be around 0.5. As shown in the histogram (**Fig. 1**), the ratio between 0 and 0.5 on the x-axis indicates the use of Estonian words is dominant in the written text. The smaller the ratio, the lower is the frequency of occurrence of the foreign equivalents. On the contrary, the ratio larger than 0.5 means that the dominance in use turns gradually to foreign words.

This histogram displays three pairs characterised by even competitiveness. The distribution can be broken into the other four ranges besides the range between 0.4 and 0.6. The distribution of pairs between 0.9 and 1 is comparable with the distribution of pairs between 0 and 0.1. However, a blank area is marked in the range between 0.1 and 0.2, whereas seven to eight pairs are demonstrated in the range between 0.8 and 0.9.

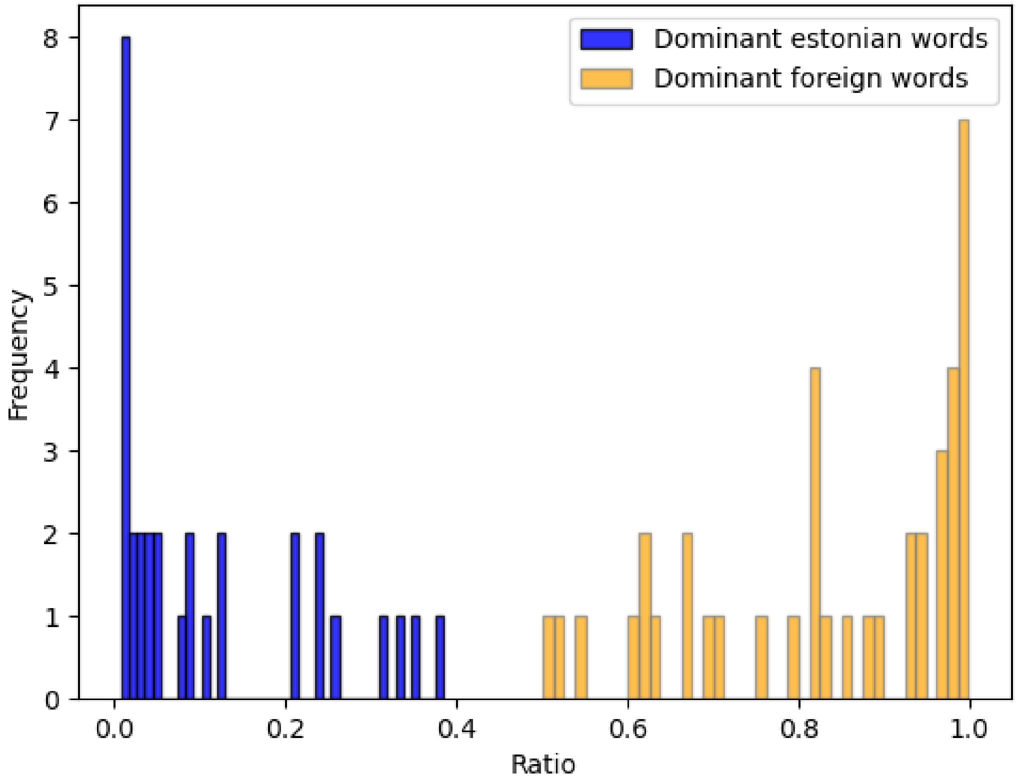


Figure 1. Histogram of the distribution of pairs of terms according to frequency ratio.

Thus, fifteen pairs (**Table. 1**) from the DE dataset for stimuli came out, based on the visualised distribution of frequency ratio. My supervisor also helped to remove the pairs consisting of the words that are seldom used in everyday speech. Six pairs characterised by Estonian words of the highest dominance are from the range between 0 and 0.1 (*min.* = 0.022, *max.* = 0.077); nine pairs are from the range between 0.2 and 0.4 (*min.* = 0.206, *max.* = 0.385). In the DF dataset, six pairs are characterised with the most dominant foreign *words* (*min.* = 0.911, *max.* = 0.979); the other nine pairs are from the range between 0.6 and 0.8 (*min.* = 0.608, *max.* = 0.799) (**Table. 1**).

Dominant Estonian set	Ratio	Dominant foreign set	Ratio
pankreas/kõhunääre ‘pancreas’	0.385	*dermatiit/nahapõletik* ‘dermatitis’	0.608
komedoon/mustpea ‘comedo’	0.351	*anoreksia/isutus* ‘anorexia’	0.617
erütrotsüüt/punalible ‘erythrocyte’	0.33	*artriit/liigesepõletik* ‘arthritis’	0.629
müoopia/lühinägevus ‘myopia’	0.318	*arter/tuiksoon* ‘artery’	0.665
tahhükardia/südamepekslemine ‘tachycardia’	0.255	*herpes/ohatis* ‘herpes’	0.666
poliomüeliit/lastehalvatus ‘poliomyelitis’	0.239	*osteoporoos/luuhõrenemine* ‘osteoporosis’	0.7
akupunktuur vs *nõelravi* ‘akupuncture’	0.235	*leukeemia/verevähk* ‘leukaemia’	0.708
skalp/peanahk ‘scalp’	0.234	*aneemia/kehvveresus* ‘anaemia’	0.762
amneesia/mälukaotus ‘amnesia’	0.206	*meningiit/ajukelmepõletik* ‘meningitis’	0.799
antidoot/vastumürk ‘antidote’	0.077	*astma/hingeldus* ‘astma’	0.911
metabolism/ainevahetus ‘metabolism’	0.049	*menstruatsioon/kuupuhastus* ‘menstruation’	0.93
hemorraagia/verejooks ‘haemorrhage’	0.048	*difteeria/kurgutõbi* ‘diphtheria’	0.962
prostata/eesnääre ‘prostate’	0.044	*eutanaasia/halastussurm* ‘euthanasia’	0.974
insomnia/unetus ‘insomnia’	0.041	*akne/vinnid* ‘acne’	0.975
pneumoonia/kopsupõletik ‘pneumonia’	0.022	*alkoholism/joomatõbi* ‘alcoholism’	0.979

Tabel 1. List of the final selection of 30 pairs of terms as stimuli in the experiment.

The final step is to devise stimuli. This step started with the creation of texts before recording. As discussed in the ***Background*** section, a proper stimulus should include both two terms and a description of the concept referred to by them. To make stimuli sound natural to Estonian native speakers, all the texts begin with a commonly used Estonian expression *A ehk B*, whose English equivalent is *A or B* — both are used when a speaker needs another way of saying a thing to connect one word with another that denotes the same concept (Eesti keele instituut, n.d.; Oxford University Press, n.d.). By collecting and sifting out materials from ENC 2023 or web texts on kliinik.ee, Tartu Ülikool kliinkum, and Wikipedia, I finished compiling texts for two conditions. In the first condition, all the texts begin with the *A* [*Estonian word/foreign word*] *ehk B* [*foreign word/Estonian word*] structure; in the second condition, descriptive sentences follow only the Estonian term. Examples are below:

Condition 1: *Erütrotsüüt ehk punalible on südame-veresoonkonna kaudu hapnikku vedav vererakk, millel puuduvad enamasti rakutuum ja paljunemisvõime. See ei tarbi ise hapnikku, vaid kannab seda teistele rakkudele. Ööpäevas teeb see hapnikku vedades tuhandeid selliseid ringe ning pärast nelja kuud see hävib.*

Condition 2: *Punalible on südame-veresoonkonna kaudu hapnikku vedav vererakk, millel puuduvad enamasti rakutuum ja paljunemisvõime. See ei tarbi ise hapnikku, vaid kannab seda teistele rakkudele. Ööpäevas teeb see hapnikku vedades tuhandeid selliseid ringe ning pärast nelja kuud see hävib.*

English translation: 'Erythrocyte or red blood cell is blood cell transporting oxygen via circulatory system; it mostly lacks a cell nucleus and the ability to reproduce. It does not consume oxygen itself but transfers it to other cells. Transporting oxygen, it makes thousands of these circles a day and dies after four months.'

To secure the general homogeneity of all stimuli texts for the third research question, loanwords and foreign words in the original text were replaced with their Estonian equivalents. Also, to make the stimuli hard to predict, I randomly placed either a foreign word or an Estonian word first in the *A ehk B* structure. In the DF set, seven pairs begin with foreign words; eight pairs begin with Estonian words. In the DE set, seven pairs begin with Estonian words; eight pairs begin with foreign words (The file storing my original 30 stimuli can be found here).

Before recording, text content must be edited by native speakers. My supervisor and another teacher assisted in revising grammatical errors and re-writing texts — when the content sounded unnatural. I found the native speakers who offered to record stimuli. Three volunteers (two women and a man) at the University of Tartu shared recording tasks; each finished one-third of this word. One volunteer helped to record four files for participants to practice. They also checked the content and improved the grammar. My supervisor provided equipment: Zoom handy recorder H2N, memory card reader, and extra batteries. Recording took place in a group study classroom booked in the university library. The total process lasted one and thirty minutes. The average length of each audio file is between 20 and 25 seconds.

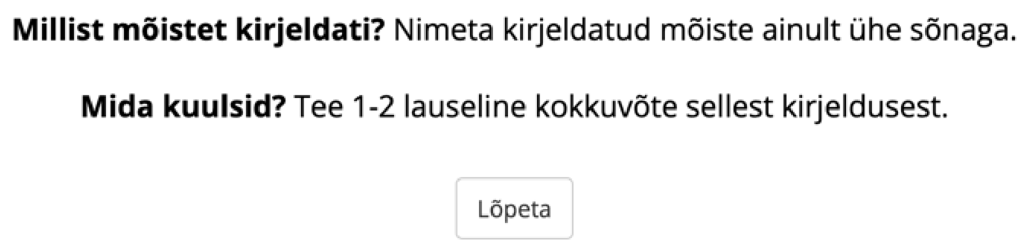


Figure 2. The screenshot of the displayed text when it is time for participant to respond to the stimulus.

The development of a platform to play all the audio files also started at the same time. The portal developed by Node.js and JsPsych plugins (de Leeuw et al., 2023) was deployed on Render (all coding files can be found on GitHub; the deployed website is here). It consists of various trials. Trials herein denote all the sections in this experiment: the logging-in trial, the introduction trial, 2 practice trials and 30 test trials. All the audio files will be preloaded before practice trials. As soon as a participant enters a practice trial or a test trial, the audio plays automatically. Between trials, there is a screen showing a button for participants to continue. Two questions (*Millist mõistet kirjeldati* 'What concept is described'? *Nimeta kirjeldatud mõistet ainult ühe sõnaga* 'Name the described concept in only one word'. *Mida kuulsite* 'What did you heard'? *Tee 1-2 lauseline kokkuvõte sellest kirjeldusest* 'Make a one-to-two sentence summary of the description'.), a button below to end the recording, will be displayed as the audio ends — signalling the beginning of the automatic recording (**Fig. 2**). In other words, between listening and recording, my platform provides no function to forward, back, pause or play the audio. Considering audio stimuli were recorded with the help of three volunteers, I divided 30 test trials into 3 blocks with two pause trials between them. All the texts displayed on the platform underwent the review by my supervisor and another teacher.

2.2. Participants

All participants reported themselves to be fluent in Estonian and English; they all now study in the six-year integrated master's study *arstiteadus* 'medicine science'. The ideal number of participants is up to 30; however, considering the limited time and resources, the goal became recruiting 15 to 20 participants. At least 10 participants in the experimental group (condition 1) should come from the University of Tartu's Medical Faculty, and 5 to 10 participants in the control group (condition 2). However, due to time restriction, I was forced to drop the control group with a focus on condition 1.

I accomplished the recruiting work through cooperation with an academic specialist working at the medical faculty. One thing that attracted my attention was the information provided by the specialist — all medical students appear to be required to attend and pass the Latin language course in their first year due to the fixed course schedule for different groups. Thus, I decided to recruit medical students who now study at least in the second year.

As planned, medical students received recruiting emails and found a link to MS booking where I created a registration form. Once notification of the booked timeslot arrived, I sent a follow-up email to the registrant. In the follow-up email, one can find their participant code and a link to the questionnaire aimed at gathering data on their background information incl. age, which year they study now, self-reported Estonian language proficiency, the time when they studied English first, English reading and writing skills, English listening skill, English speaking skill (self-reported in a 5-scale rating), the situation regarding whether they passed the course *ladina keel meditsiinierialadele* 'Latin language for medical speciality (FLEK.05.008)' (this information was required lest the one had not passed), and general description about what other languages they use every day. Self-reported data extracted from the questionnaire will be also used in data analysis.

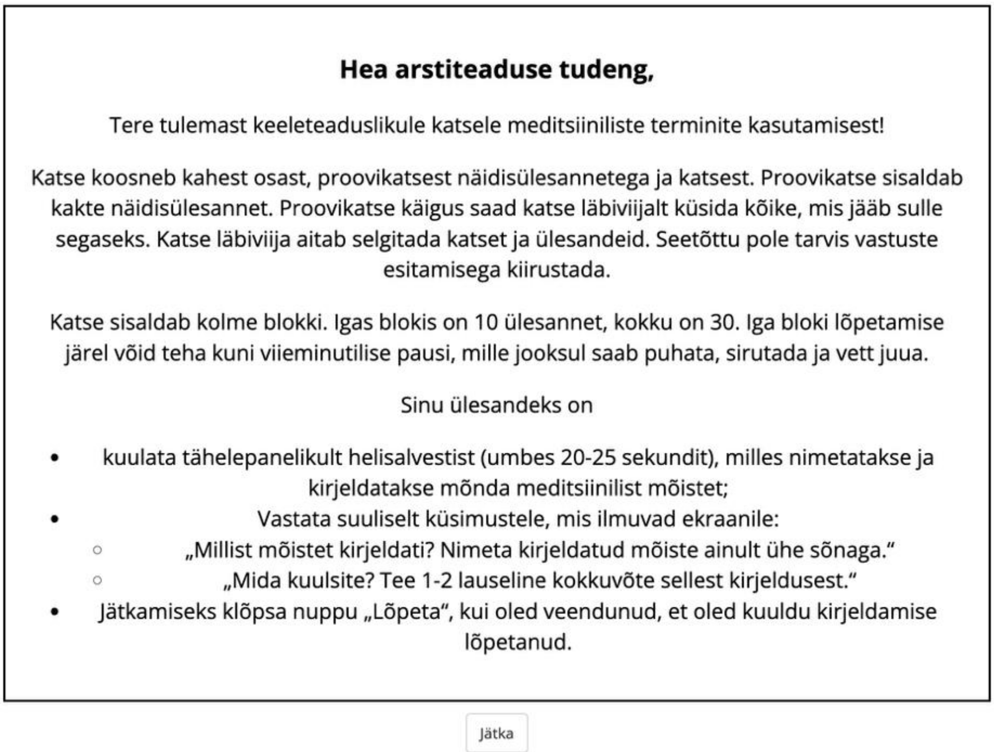
Both experiments and recruiting took place in the month of April. By the 24th, the last day for experiments, 13 medical students registered, yet two did not show up. All the 11 students now study in the third year or above at the University of Tartu's Medical Faculty. I decided to remove one participant's data because their self-reported language proficiency did not match their actual abilities.

Data on participants' background information exhibits heterogeneity in some variables while maintaining homogeneity in others such as the time when they first studied English and Estonian language proficiency. Participants' average age is 24 (*max.* = 31, *min.* = 21, MD = 23). No participant studies in the second year. Four students are from the third year and 2 students are from each of the remaining years. All participants learned Russian apart from English; the second most learned is German; however, in their answers to the question regarding the languages they use in everyday life, Russian appears to be less dominant. Only one participant clearly states that they use Russian with their family member. I also found subtle variations in participants' English proficiency in reading and writing, speaking, and listening; I will cover this point in the **Result** section.

2.3. Procedure

The experiment took place at the 'Biomeedikum' building, where medical students spend most of their time studying and attending classes. Academic specialists helped to book rooms ahead of experiments. If a computer is installed in the booked room, participants can work on that or use the experimenter's laptop. Participants signed the paper consent form before the experiment. According to the consent form, unprocessed data will be sent and stored in an encrypted folder in *Sharepoint Cloud* of TÜ; the backup will be also stored in the experimenter's laptop for further data analysis. Data analysis will never involve reconstructing the identity of the subject; nor will personal information be made public in the result presentation. Participants have the right to request the deletion of data within two weeks from the end of the experiment. All the derived voice data will be deleted the latest in August 2025.

To complete experiment tasks, each participant sat in front of either my laptop or the computer installed in the booked room (all experiments were done on my laptop); experimenter sat one or two desks away from the participant. Equipment includes the same H2N recorder by which I backed up responses and headphones with an external microphone. The experimenter had launched the deployed platform of which the first scene displays the title *Tänan osalemast katses* 'thanks for joining the experiment' and a button to continue and to preload all the audio files. The next step is to input the participant code. In the initial plan, participants are assigned to two different conditions randomly. Participant codes will be processed, thus guiding participants to different experimental conditions. However, since the plan changed eventually, the basic structure of all the assigned codes became the same. Then, the participant read instructions that inform them of how many sections the whole experiment consists of, how many trials there are, under each section what kinds of tasks are waiting for them, and what exact actions they should take (**Fig. 2**). Exhibited in the screenshot below, wearing headphones, participants should first listen to audio being played in both practice trials and test trials with full attention. As soon as the audio ends, participants should proovide their lexical selection in whatever language they use and generalise what they heard with one or two sentences in a loud voice. Though I set two pauses for them between three text trial blocks, only a few participants took a rest. The average time they spent on my experiment was between 20 and 30 minutes; the longest experiment lasted 40 minutes.



Hea arstiteaduse tudeng,

Tere tulemast keeleteaduslikule katsele meditsiiniliste terminite kasutamisest!

Katse koosneb kahest osast, proovikatsest näidisülesannetega ja katsest. Proovikatse sisaldab kakte näidisülesannet. Proovikatse käigus saad katse läbiviijalt küsida kõike, mis jääb sulle segaseks. Katse läbiviija aitab selgitada katset ja ülesandeid. Seetõttu pole tarvis vastuste esitamisega kiirustada.

Katse sisaldab kolme blokki. Igas blokis on 10 ülesannet, kokku on 30. Iga bloki lõpetamise järel võid teha kuni viieminutilise pausi, mille jooksul saab puhata, sirutada ja vett juua.

Sinu ülesandeks on

- kuulata tähelepanelikult helisalvestist (umbes 20-25 sekundit), milles nimetatakse ja kirjeldatakse mõnda meditsiinilist mõistet;
- Vastata suuliselt küsimustele, mis ilmuvad ekraanile:
  - „Millist mõistet kirjeldati? Nimeta kirjeldatud mõiste ainult ühe sõnaga."
  - „Mida kuulsite? Tee 1-2 lauseline kokkuvõte sellest kirjeldusest."
- Jätkamiseks klõpsa nuppu „Lõpeta", kui oled veendunud, et oled kuuldu kirjeldamise lõpetanud.

Jätka

Figure 3. The screenshot of the displayed instructions on the platform for the experiment.

As mentioned, participants should first familiarise themselves with tasks in two practice trials (the first section of the experiment). From the instruction trial to practice trials, the participants were expected to ask questions to the experimenter until they understood well what they should do to fulfil tasks in the experiment; they were also informed to do so in the instruction trial. In case the participant asked nothing, while performing wrongly in practice trials without realising errors, the experimenter must stop the participant from continuing and help them to understand tasks with detailed explanations. If necessary, the participant should reload the page so that the participant can restart; test trials start only when nothing is unclear to the participants. The time by the test trials is the only time when the experimenter is allowed to provide help. The communication language is only Estonian.

As soon as the experiment ended, the dataset was downloaded automatically to the local host, with a screen showing appreciation to participants. The recorded acoustic data is stored in base 64 format in one column of the dataset, waiting for further processing — they must be transformed into waveform format on other platforms. Then transcribing work started. I did transcription on Praat. Multifaceted properties of sound waves on Praat are useful for marking the voicing onset of the first utterance; also, reading sound waves is a bonus in transcribing work. Transcription work lasted until the end of the week after the final experiment. The supervisor provided help in reviewing part of the transcription of which data is hard to understand for non-native speakers. Besides transcription, I also marked initial latency for the second research question based on variations in formants, spectrogram, intensity, and pitch (**Fig. 4**). Moreover, information regarding (i) whether the structure *A ehk B* is primed; (ii) whether the foreign/Estonian equivalent to the selection is used in the rest of the speech; (iii) the count of other foreign words used in the rest of the speech; (iv) used foreign words; (v) used foreign morphemes; (vi) the presence or absence of codeswitching. All the transcripts that consist of lexical selection and the rest of the speech are stored first temporarily in a folder in Google drive.

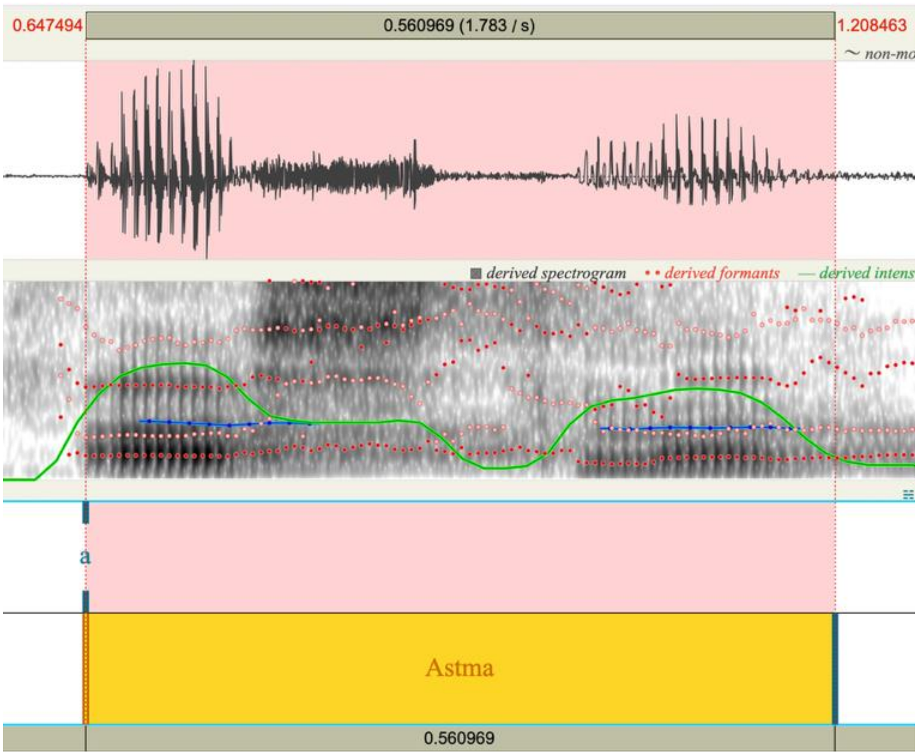


Figure 4. A screenshot of the transcribed sound with voicing onset märked.

In transcribing, I found other lexical selections — other Estonian words or noun phrases parallel to those presented in stimuli: *punaverelible* or *verepunalible* vs *punalible*; *lühinägelikkus* vs *lühinägevus*; *anoreksia nervosa* vs *anoreksia*; *herpes simpleks* or *herpes infektsioon* vs *herpes*. In the longer spoken response, the phrase *luukoe hõrenemine* which was not selected other than *luuhõrenemine* was used. After the discussion with my supervisor, I decided to view them as acceptable word selections. Though alternatives parallel they differ from the expected stimuli in the morphemes in word formation: (i) new morphemes added, (ii) suffix changed, or (iii) the original noun replaced by a noun phrase. Yet the base bears no change in general. The only unacceptable lexical selection is *poilonakkuhaigus*, which appears to be a hybrid of the native word *nakkushaigus* 'infectious disease' and a foreign word *polio* (NB. Whether *polio*, as the shortened *poliomyelitis*, is acceptable in Estonian language planning is unknown. If not, it would be more of a quoted word '*tsitaatsõna*'. However, it is outside of the focus of my study, and I call it a foreign word as any related details may not contribute substantively to the progress of this research.).

2.4. Analysis

The first research question involves these variables: (i) a dependent variable, the foreignness of the lexical selection; (ii) independent variables incl. the order of the lexical selection in the stimulus and frequency ratio. I coded the order and the dependent variable as binary data; coding work was done on R Studio. The dependent variable in the second research question is continuous data. According to the distinction in data types of dependent variables, a mixed effect logistic model was used to analyse data for the first research question and a linear mixed effect model for the second research question. The mixed-effect logistic model is a statistic model used to predict the binary variable, with both fixed-effect independent variables (e.g., frequency ratio in this study) and random-effect variables (e.g., variation between participants and items) considered. The mixed-effect lineaar model is, by contrast, used to predict the continual variable.

3. Results

3.1. Descriptive analysis

Overall, the foreign terms were preferred to *Estonian words,* as shown in Figure 3. In 299 trials, selections of foreign terms were found in 233 trials, 78 per cent of all the trials. By contrast, 66 cases where selections of Estonian terms were demonstrated (**Fig. 5**). Among 10 participants, two participants selected all foreign terms. With their data removed, the trend reflected in **Figure 7** is yet the strong preference towards foreign lexemes. The plot of counts of lexical selections according to the range of frequency ratio shows, first, that the number of selections of Estonian words shrinks as the frequency ratio increases. In the 0.2-to-0.3 range and the 0.3-to-0.4 range, selections of Estonian words make up around 30 per cent of all the lexical selections. However, when the frequency ratio increases and skews towards foreign words, the number of selections of Estonian words drops significantly. The highest proportion of the selections of Estonian words appear in the range between 0.6 and 0.7, around 13 per cent. Secondly, it is also noticeable that selections of foreign words still make up the larger proportion (around 63 per cent), compared with selections of Estonian words, when the frequency ratio drops to the bottom (**Fig. 6**).

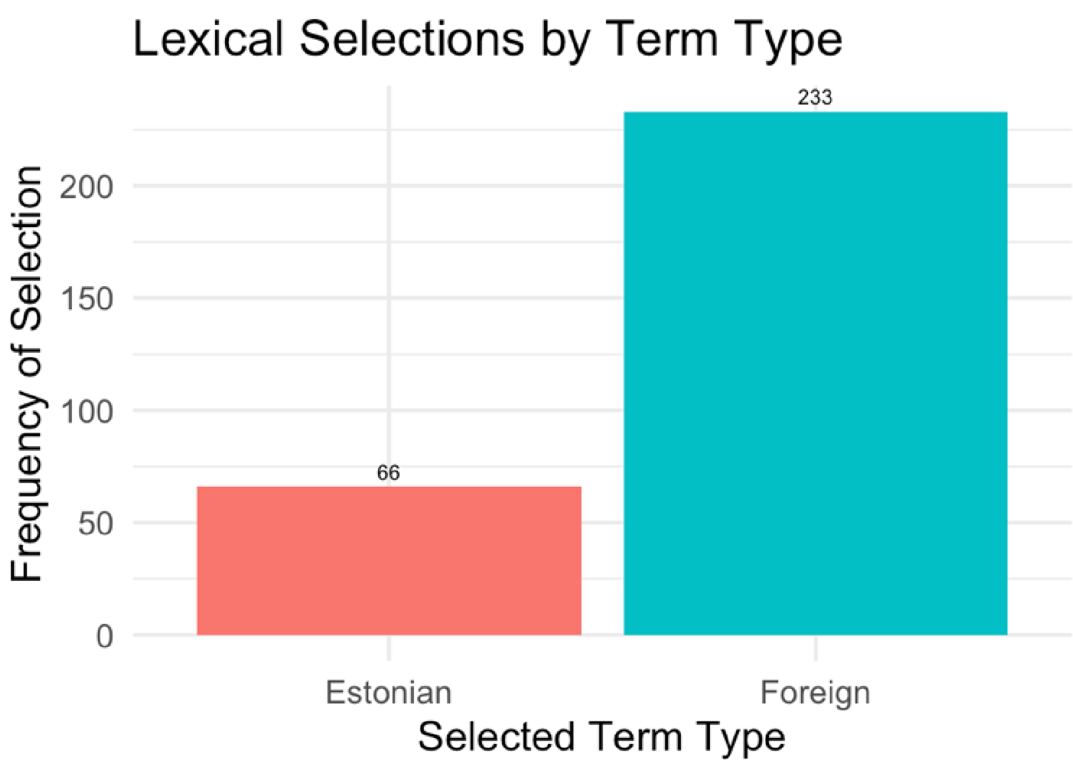


Figure 5. Overall lexical selections in 299 trials.

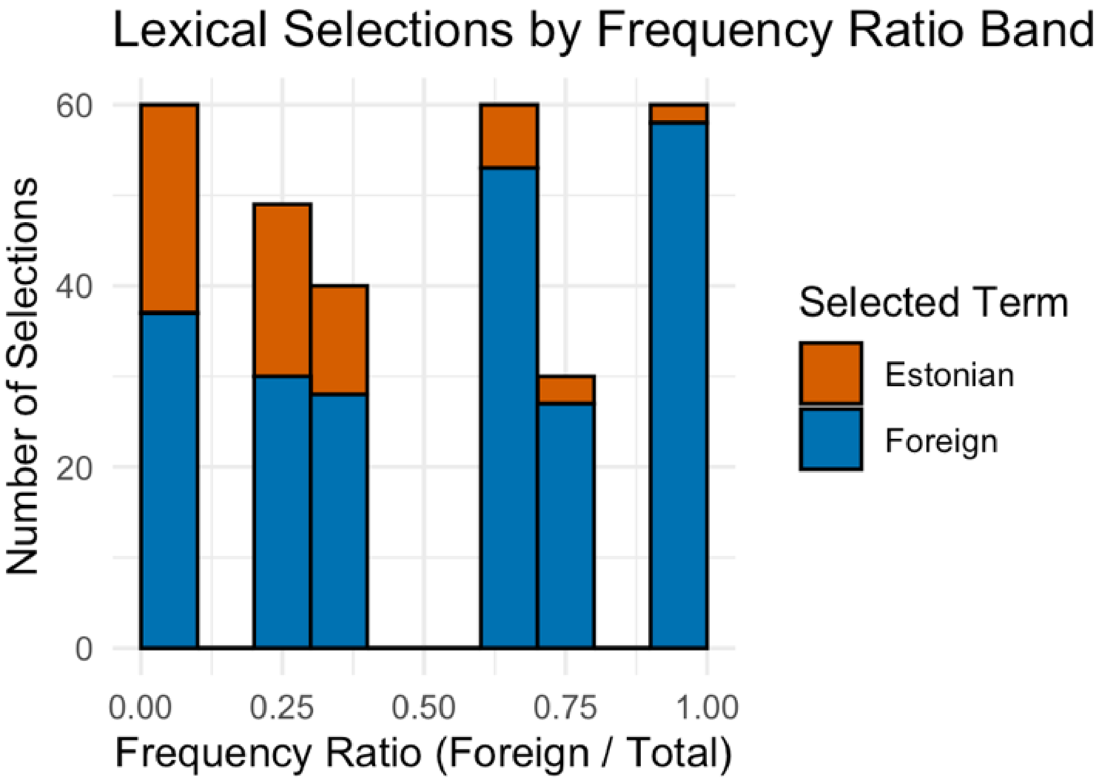


Figure 6. A bar chart to visualise number of selections of the two types of terms according to frequency ratio.

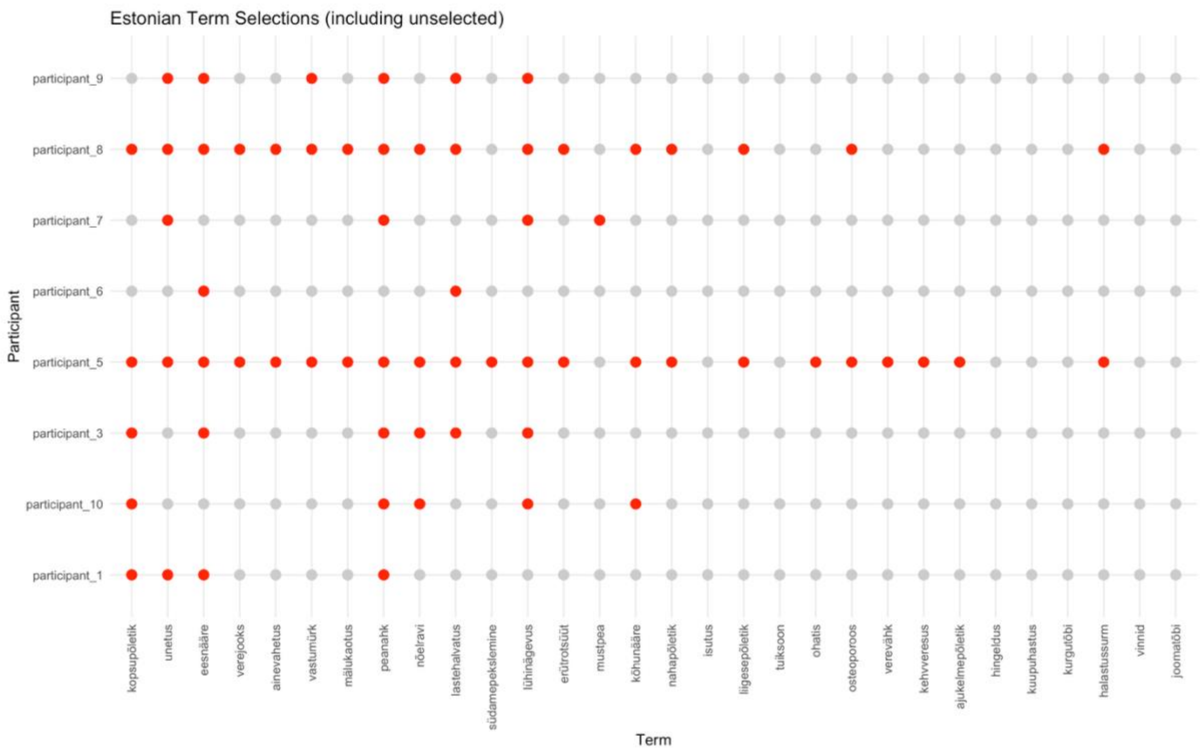


Figure 7. A visualised summary of information on who selected what types of terms, with red dots to indicate the selected Estonian words and grey dots to indicate the selected foreign words (two participants removed due to all selections of foreign words).

The close view of the selections of each participant (**Fig. 7**) demonstrantes that nobody selected Estonian terms in these seven trials: (i) *astma* vs *hingeldus*, (ii) *menstruatsioon* vs *kuupuhastus*, (iii) *difteeria* vs *kurgutõbi*, (iv) *akne* vs *vinnid*, (v) *alcoholism* vs *joomatõbi*, (vi) *anoreksia* vs *isutus*, and (vii) *arter* vs *tuiksoon*. In the trials (i), (ii), (iii), (iv), (v), the frequency ratios of foreign terms are above 90 per cent, whereas the trial (vi) and the trial (vii) are characterised by rather lower frequency ratios of foreign lexemes. In the opposite set where the frequency ratios of foreign terms are below 0.5, the selections of foreign terms still dominate in these trials: (i) *antidoot* vs *vastumürk*, (ii) *metabolism* vs *ainevahetus*, (iii) *amneesia* vs *mälukaotus*, and (iv) *hemorraagia* vs *verejooks*; the frequency ratios of Estonian lexemes are above 0.9.

3.2. Inferential analysis

3.2.1. Lexical selection and frequency ratio

The descriptive analysis results indicate that a strong preference for foreign words dominates in lexical selection, throughout 30 trials. To clarify whether frequency ratios predict lexical selection, several mixed-effect logistic models were established and compared. With the likelihood ratio test, I decided to include both variability within participants and variability between stimuli as random effect predictors (χ = 15.009, $p < 0.001$), yet excluded the interaction between the frequency ratio and the order of presentation in the stimulus in the model because the result proved it to be an insignificant predictor of lexical selection (Estimate = -1.3760, $p > 0.1$). In other words, the effect of frequency ratio and the changes of the effect are stable and not modulated by the effect of order. Moreover, considering that the likelihood ratio test indicates that the model involving the interaction effect does not strengthen the explanatory power of the model, the interaction was not included as a predictor in the model (χ = 1.051, $p > 0.05$).

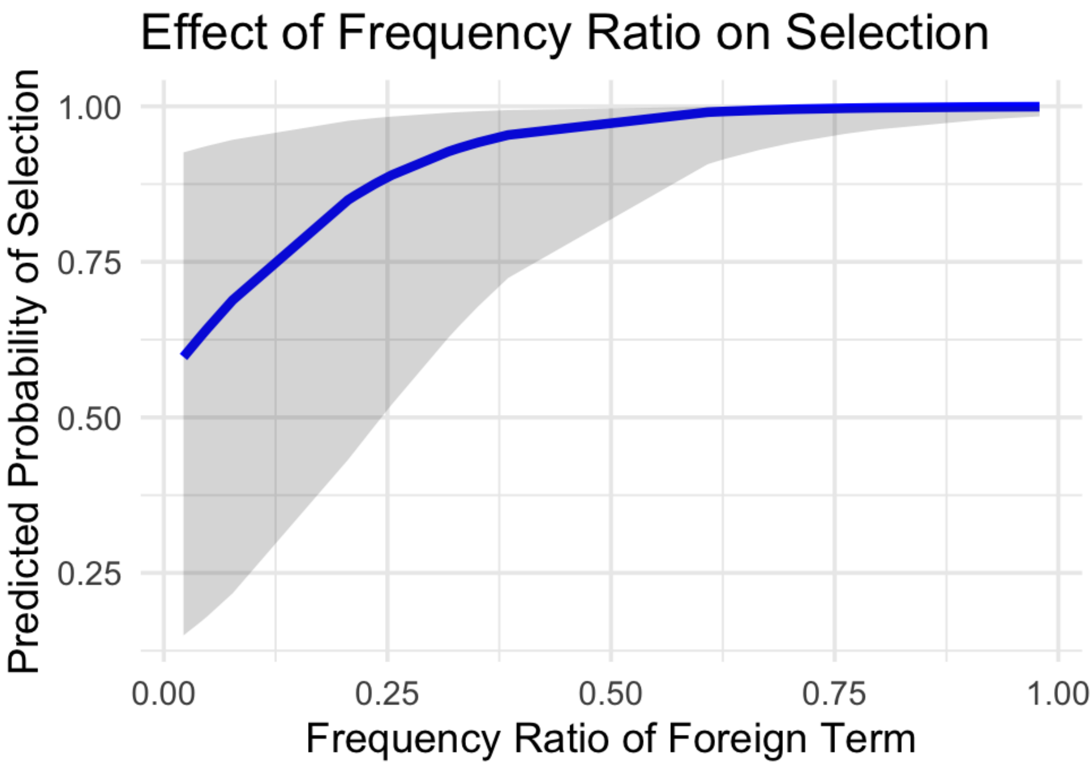


Figure 8. A visualisation of the logistic modelled predicted probability of selection foreign lexemes based on frequency ratios.

Group	Name	Variance	Std. Dev
items	(Intercept)	2.764	1.662
anon_participant	(Intercept)	6.573	2.564

Table. 2 A table of random effect of participant and stimuli pairs to indicate the variations between participants and stiimuli

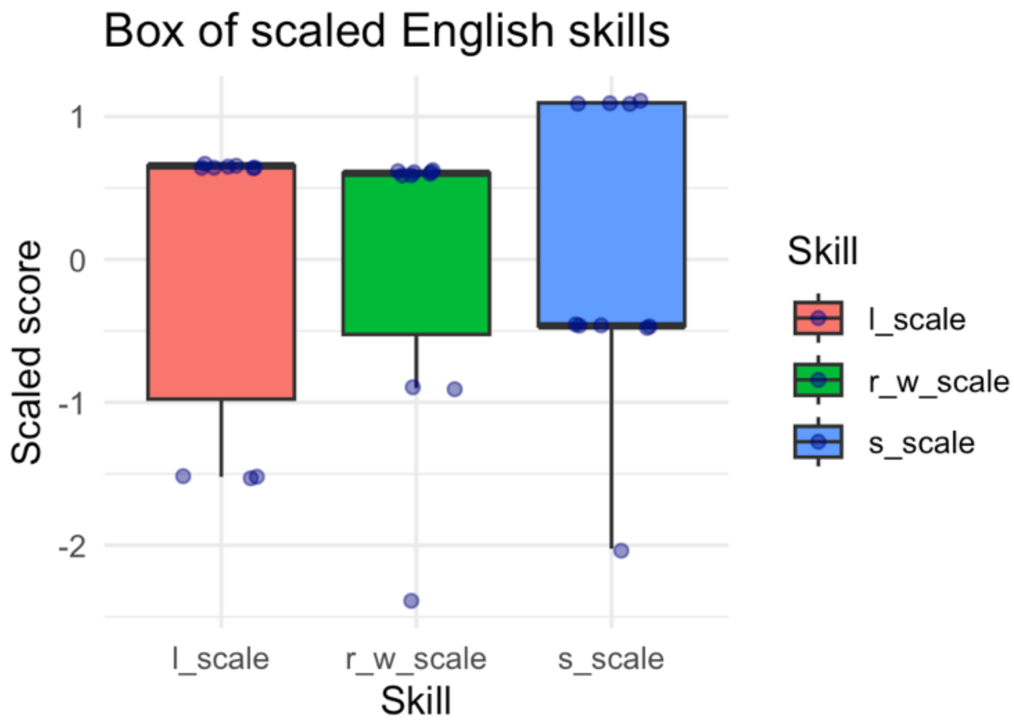


Figure 9. A boxplot of scaled (transformed into z-score) self-reported English proficiency.

As noted in the **Method** section, subtle variations exist in participants' background information incl. age (*max.* = 31, *min.* = 21, MD = 23), year of study (four from the third year; others are from the remaining higher years of study), and self-reported English proficiency. As shown in the boxplot (**Fig. 9**), though below the median, some scores are not outliers. The only outlier emerges in the reading-and-writing skill. However, all the scores of self-reported proficiencies are above 3. After finding no significant correlation between variables regarding participant background information — three sets: (i) scaled age to scaled year of study ($| r | < 0.7$); (ii) scaled age to scaled average English proficiency ($| r | < 0.7$); (iii) scaled year of study to scaled average English proficiency ($| r | < 0.7$) — I included all fixed-effect predictors (stimulus-level: frequency ratio, order; participant-level: age, year of study, average English skills) and random-effect variables (participants and items) into the model. Another problem to solve before the analysis was whether to include the scaled average English proficiency or the three scaled English skills. By comparing the results of two models, the correlations between these fixed-effect predictors were found to be low enough that the concern about collinearity is unnecessary. Secondly, both models demonstrated the evidence that frequency ratios across these models significantly influence the foreignness of lexical selection ($p < 0.0001$). However, with the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), a model including the scaled average English skills is better for both the accuracy of prediction and the model simplicity (model parsimony) (AIC = 190.9716, BIC = 213.1742). The model specification is as follows:

model_RQ1_bg_all_1 <- glmer(selected_foreign ~ freq_ratio + order_first + age_scale + year_scale + eng_ost_scale + (1 | anon_participant) + (1 | items), family = binomial, data = merged_frame)

	Estimate	**Std. Error**	**z-value**	**Pr (> | z |)**
(Intercept)	0.09087	1.12952	0.080	0.9359
freq_ratio	7.26572	1.74965	4.153	3.29e-05 ***
order_first	0.29424	0.50748	0.580	0.5620
age_scale	1.84558	1.31849	1.400	0.1616
year_scale	-1.60900	1.12097	-1.435	0.1512
eng_ost_scale	-1.82832	1.05443	-1.734	0.0829

Signif. codes: 0 '***', 0.001 '**', 0.01 '*', 0.05 '.', 0.1 ' ', 1

Table. 3 A table of summary of the logistic mixed-effect model assessing the effect of mixed-effect predictors (*freq_ratio*: frequency ratio; *order_first*: order of presentation in stimuli; *age_scale*: scaled age; *year_scale*: scaled year of study, *eng_ost_scale*: scaled English proficiency).

	(Intr)	*frq_rt*	*ordr_f*	*ag_scl*	*yr_scl*
freq_ratio	-0.362				
order_first	-0.223	-0.011			
age_scale	0.164	0.129	-0.014		
year_scale	-0.114	-0.156	-0.004	-0.609	
eng_ost_scl	-0.092	-0.164	0.007	-0.302	0.057

Table. 4 A table of correlation of mixed-effects (*frq_rt*/*freq_ratio*: frequency ratio; *ordr_f*/*order_first*: order of presentation in stimuli; *ag_scl*/*age_scale*: scaled age; *yr_scl*/*year_scale*: scaled year of study, *eng_ost_scl*: scaled English proficiency).

The final model demonstrated a comparable result to the tested models. Correlations between fixed-effect predictors are all checked for collinearity. The intercept or the log-odds of selecting the foreign word when all fixed-effect predictors are null is 0.09087; the corresponding probability, thus, is 0.523. The main effect of frequency ratio is still the only significant predictor of lexical selection (Estimate = 7.26573, $p < 0.0001$) (**Table. 3**). It is noteworthy that the estimate of frequency ratio or the slope of the logistic model is quite large. The basic equation to simulate the relationship between the probability of selecting a foreign word and the frequency ratio is:

$$\textit{log-odds} = b_0 + ratio * 7.26573$$

Given that the intercept b_0 approximates 0, with every 1-unit increasing in frequency ratio, the log-odds (the logarithm of odds) of selecting a foreign word increases by around 7.27. In other words, the odds of selecting a foreign word increase by a factor of 1430.429 (after exponentiating 7.26573). On the one hand, this model exhibits a strong preference for foreign terms insofar as the probability of selecting foreign terms is much higher than predicted based on the corpus-based frequency ratio despite a lower frequency ratio. On the other hand, the confidence interval that increases as the frequency ratio drops from 0.6 indicates that the model is becoming uncertain with the prediction (**Fig. 8**).

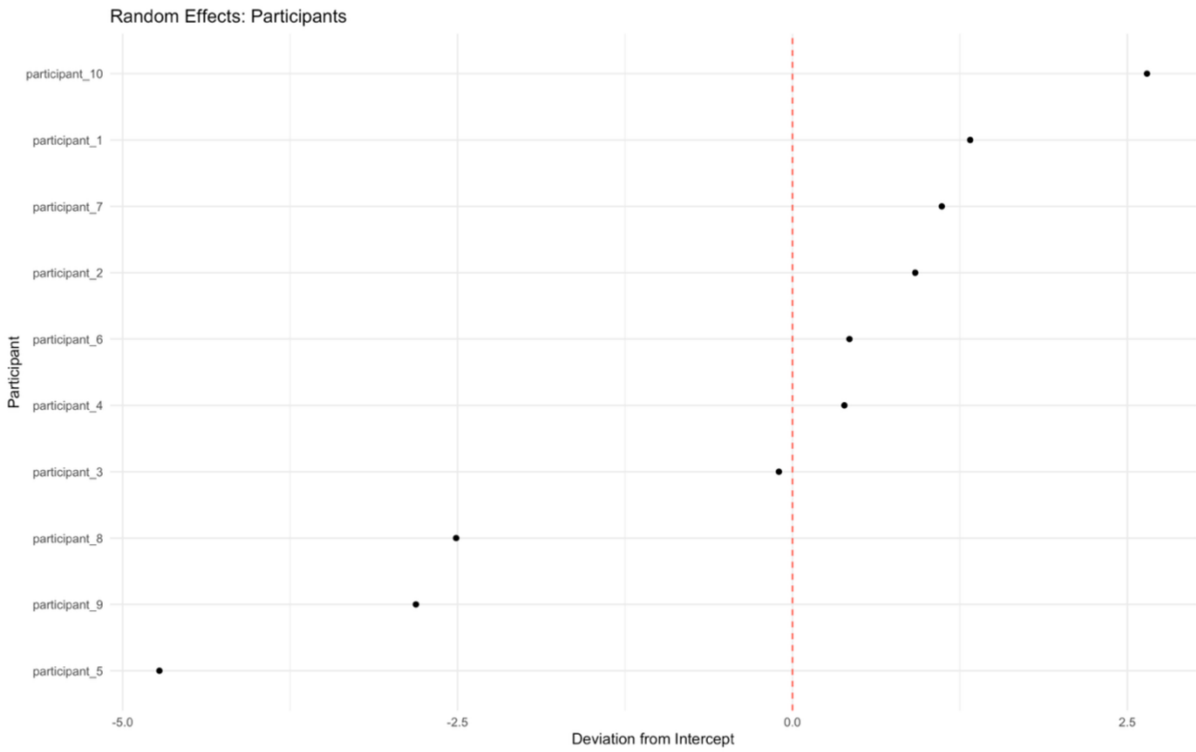


Figure 10. A visualisation of random effect of participants to display participant-level tendencies to making lexical selections if all the mixed-effect predictors are null.

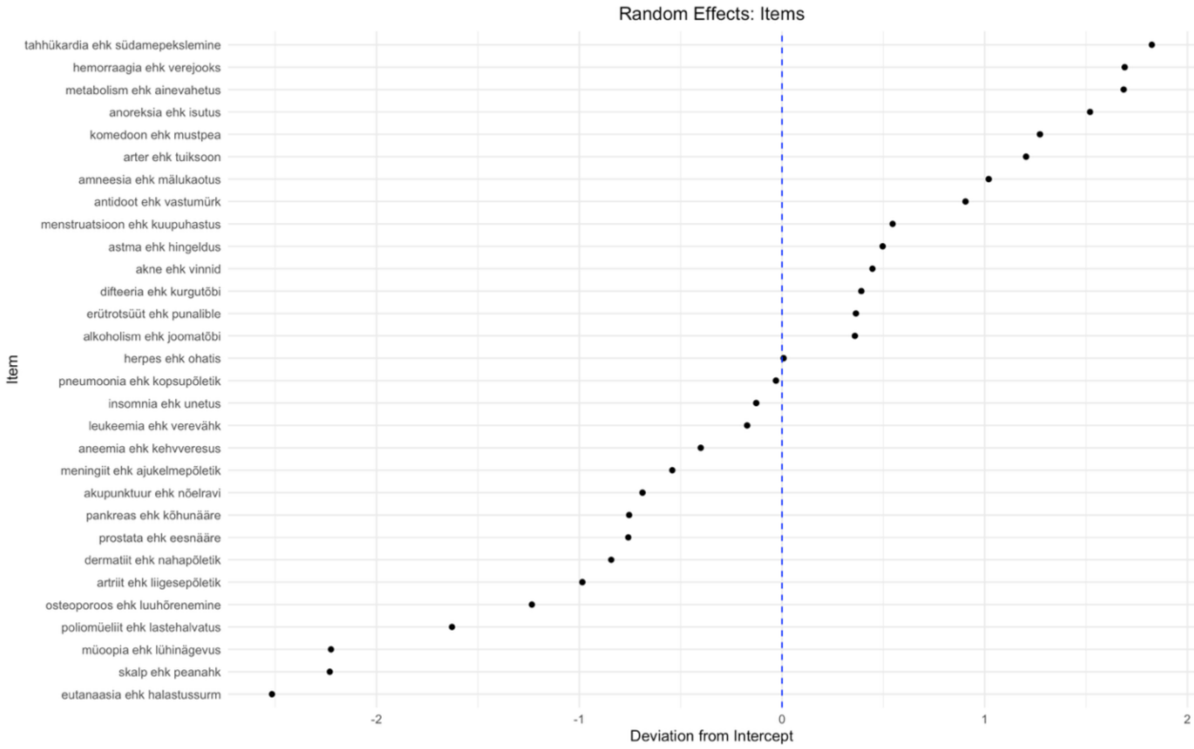


Figure 11. A visualisation of random effect of stimuli pairs to display item-level anomalies if all the mixed-effect predictors are null.

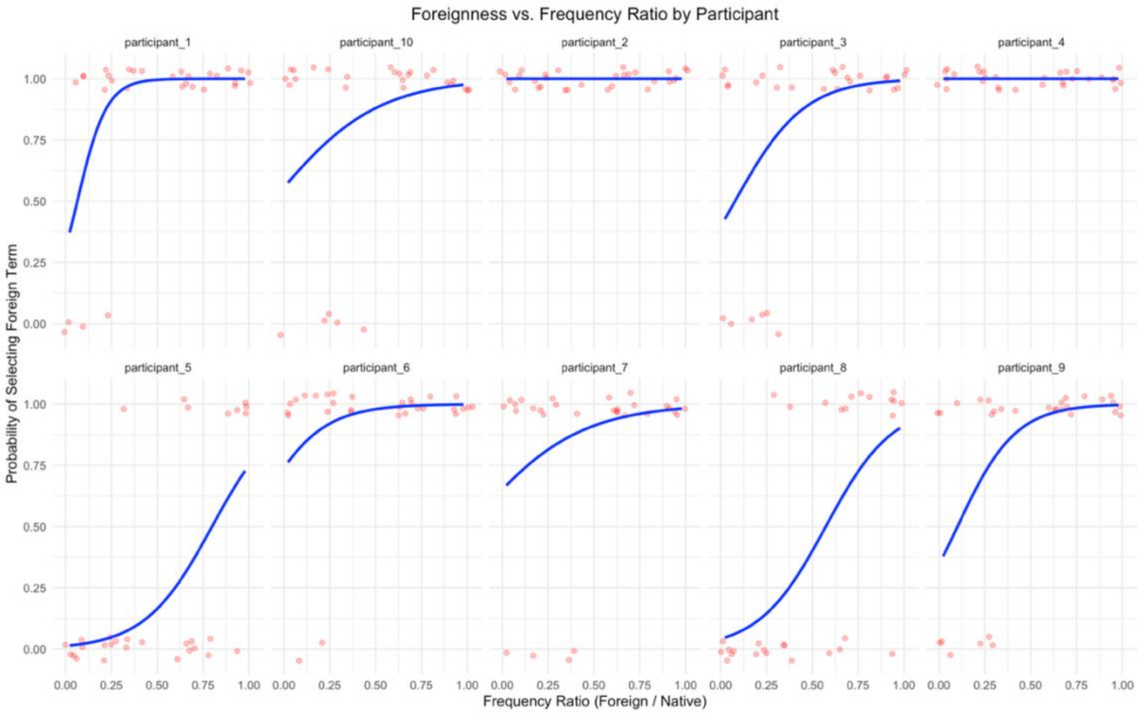


Figure 12. A set of 2 * 5 panels to display the predicted relationship between lexical selections and frequency ratios per participant.

In random effects (**Table. 2**), each participant differs significantly (Variance = 6.573, SD = 2.564) from each other in their general tendency to make lexical selection after the model assumes the data regarding background information (age, year of study, scaled average English proficiency) as predictors to influence participants in the same way. By contrast, the variation between stimulus pairs is insiginificant (Variance = 2.764, SD = 1.662). This finding proves that the behaviours deviating from the prediction exist and that the established model cannot account for them. **Figure 7** shows that both Participant 5 and Participant 8 preferred Estonian terms (22 Estonian terms by Participant 5 and 17 Estonian terms by Participant 8), compared with others. Their selection patterns are biased towards Estonian words (**Fig. 12**). In **Figure 10**, both deviate from the average intercept the farthest. Interestingly, in **Figure 11**, if stimuli deviate from the intercept in the direction against the predicted — for instance, these pairs: (i) *antidoot* vs *vastumürk*, (ii) *metabolism* vs *ainevahetus*, (iii) *amneesia* vs *mälukaotus*, (iv) *hemorraagia* vs *verejooks*, etc. — the selections against the predicted were made by Participant 5 and Participant 8.

	age	***study of year***	***scaled average English proficiency***
participant_10	25	The sixth year	0.863
participant_9	22	The fourth year	-0.980
participant_3	21	The third year	0.249

Table. 5 A table of data on distinctions in ages, years of study, and scaled average English proficiency between participant 10, participant 9, and participant 3.

However, though the selection patterns of the rest of the participants in **Figure 12** are comparable, the corresponding random effects are not the same in **Figure 10**. For instance, both Participants 2 and 4 selected foreign terms in all 30 trials (**Fig. 12**). On principle, they should deviate the same from the intercept. However, they differ from each other in age (Participant 2: *31*, Participant 4: *22*), year of study (Participant 2: the fourth year, Participant 4: the third year), and scaled average English skills (Participant 2: *0.863*, Participant 4: *-2.208*). Thus, **Figure 10** demonstrates that compared with Participant 4, Participant 2 should express a higher preference for foreign words due to random effect. Also, the selection patterns of Participant 9, Participant 10, and Participant 3 (**Fig. 7, Fig. 12**) are also comparable. Their deviation from the intercept, nevertheless, appears to be confusing (**Fig. 10**) — one almost goes in congruence with the intercept, whereas the other two are plotted in two contrasting places, above + *2.5* and below *-2.5*. It is difficult to interpret this phenomenon without considering the distinctions in their background information (**Table 5**).

3.2.2. Initial latency and lexical selections

I established a linear mixed-effect model to examine whether fixed-effect predictors (frequency ratio, the order of lexical selection, foreignness of lexical selection, and interaction between any two predictors) and random-effect predictors (participant and stimuli) have any influences on initial latency. Before the inferential analysis, it is necessary to filter the outliers in initial latency. For instance, Participant 2 spent quite a long time (11s in trial 19 and 12s in trial 20) before lexical selection in two trials, whereas the longest time for other participants was at most 4s. After removing these outliers, exploratory analysis displayed none of the statistically significant effects. I decided to add a new variable, the use of the term that was not selected, coded as a binary value in the original dataset. In coding this information, I did not distinguish the places of the terms that were not selected. However, the descriptive analysis displayed that the priming action was quite rare throughout the experiment. Only one participant primed *A ehk B* structure in 70 per cent (21 trials) of all the trials; another primed only twice (**Fig. 13**, **Fig. 14**). After removing 30 trials of the participant who primed the most frequently, I found no predictors of significant effect on initial latency.

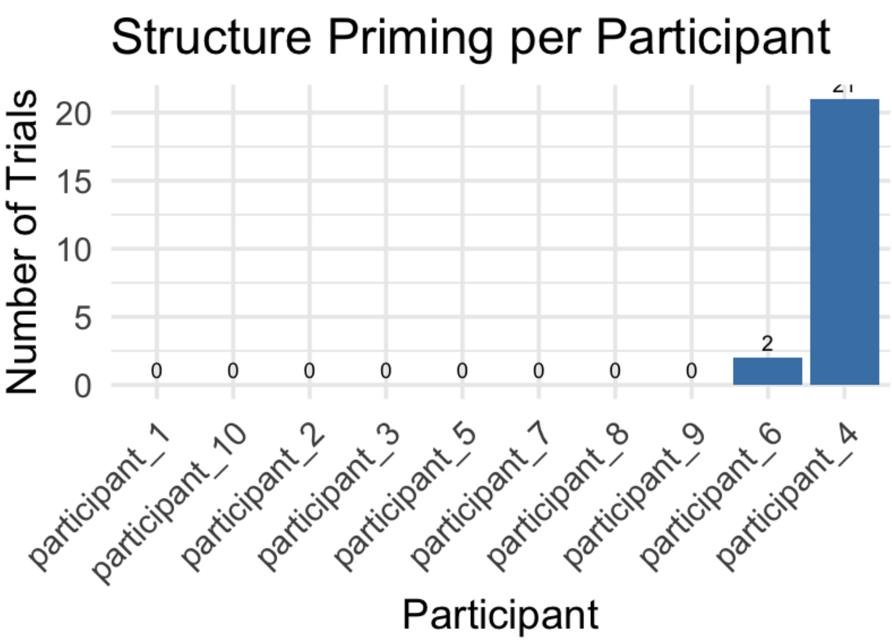


Figure 13. A bar chart of number of trials per participant where structural priming was done.

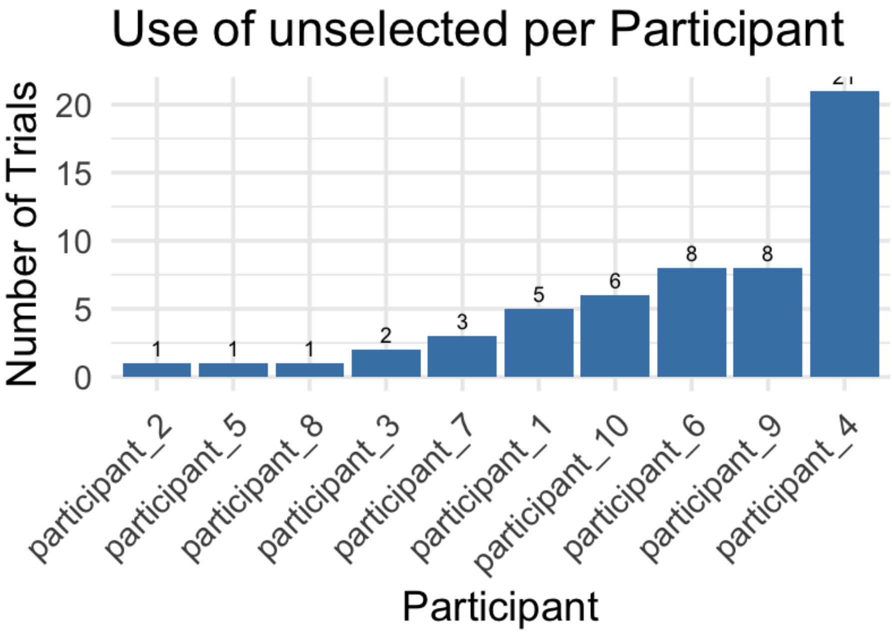


Figure 14. A bar chart of number of trials per participant where the Estonian/foreign alternative to the selection was also used in the response.

3.3. Other aspects of participants' responses

To explore whether the priming effect was used in the produced utterance following the lexical selection, the third research question focuses on three aspects: (i) whether the *A ehk B* structure was used; (ii) whether other foreign words and foreign morphemes were used; (iii) whether code-switching was used. Firstly, in the discussion regarding the model aimed to tease out predictors and reaction time, the first aspect has been covered — priming the structure *A ehk B* occurred rarely. Hence, no priming effect has manifested itself in the data; nor is there any need to perform further statistical analysis.

The second aspect is the use of other foreign words and foreign morphemes. Participants used various foreign words incl. derivations based on various foreign morphemes. The most frequently used foreign morpheme is *normaal-* (6 times). The most frequently used foreign word is *sümptom* 'symptom', used 28 times. Other the second most frequently used foreign words pale in comparison incl. *mehhanism* (7 times), *organism* (6 times), *probleem* (5 times), *infektsioon* (6 times) and *erütrotsüüt* (5 times), *protsess* (5 times). These words are almost commonly used except for *erütrotsüüt*, which was used by four participants in their descriptions of anaemia (the 23rd trial). The alternative to *infektsioon*, *nakkus*, provided by Sõnaveeb, is found in responses by Participant 8, Participant 7, Participant 6, and Participant 1 to the 20th trial *ohatis* vs *herpes*. Interestingly, in the 15th trial *kopsupõletik* vs *pneumoonia*, participant and participant used *infektsioon* (below). By contrast, the alternatives to *sümptom*, *näht* and *haigusnäht* were seldom used.

The 15th trial:

Participant 6: *Pneumoonia on eh … eh ...* ***inflektsioon*** *haigus*, *eh ... mis haarab kopse*. *Võib harrata väiksemat või suuremat osa või ka mõlemat kopsu*, *eh ... ja sümptomiteks tavaliselt on palavik, köha, torkiv valu rinnus, nõrkus*.

Translation: Pneumonia is eh … eh … infectious disease, eh ... that affects the lungs. It can affect a smaller or larger part or both lungs, eh ... and the symptoms usually include fever, cough, stabbing chest pain, weakness.

Participant 7: *Pneumoonia*. *Kirjeldati bakteriaalset viiruslikku* ***infektsiooni*** *kopsudes*, *mis võib eh ... mitu või ainult ühte sagarat*, *eh ... hõlmata*.

Translation: Pneumonia. A bacterial viral infection of the lungs is described, which can involve, eh ... several or just one lobe, eh ...

The 20th trial:

Participant 6: *Herpes simpleks eh ... on viirus**nakkus***, *millal on kaks alamvormi*. *Eh ... võib sisse peamiselt haarata limaskesti kas näol või genitaalidel ja sümptomiteks on sügeluskipitus*, *valu*, *villid*.

Translation: Herpes simplex eh ... is a viral infection that has two subtypes. Eh ... can mainly affect the mucous membranes of the face or genitals and symptoms include itching, pain, blisters

Participant 7: *Herpes* ***nakkus****viirushaigus*. *Kirjeldati eh ... kahte tüüpi*, *mis on olemas*, *HSV-üks ja HSV-kaks*. *Samuti räägiti*, *milline on sümptom tavaliselt ilmnevab*.

Translation: Herpes, infectious viral disease. two types that exist, HSV-one and HSV-two, were described. What the symptoms usually appear to be is also described.

Finally, the only instance of code-switching in the data occurred in the response by participant 10 to the trial *komedoon* vs *mustpea*:

(*eh* ...) *Komedoon. Komedoon on* (*eh* ...) *nii-öelda* ***blackhead***, *mis tekib* (*eh* ...) *nahale*, *ei oska midagi öelda* (*eh* ...).

Translation: (*eh* ...) *Comedo. Comedo is* (*eh* ...), *so to speak*, *blackhead*, *which appears* (*eh* ...) *on skin*, *that's what I can say* (*eh* ...).

The code-switching is used in the expression *A on nii-öelda B* 'A is, so to speak, B' instead of *A ehk B*. The definition of *nii-öelda* on Sõnaveeb is 'as usually said/spoken (*nagu* (*tavaliselt*) *öeldakse*)' (Eesti keele instituut, n.d.). I translated the expression *A on nii-öelda B* into 'so to speak' based on definitions provided by multiple tools (Oxford University Press, n.d; Sanseido, n.d.). This expression is used to explain or rephrase something in another way, in an unusual or metaphorical way. However, why is it *blackhead* rather than *mustpea*, or why is it code-switching that was used for rephrasing? At first glance, no significant distinctions can be found between *blackhead* and *mustpea* as not only do they refer to the same concept, but their word formations correspond with each other — *must* means *black*; *pea* means *head*. They appear to be a pair of words that can be commonly interchangeably used. However, the data of lexical selections per participant informs us that in the trial *komedoon ehk mustpea*, *mustpea* is the less acceptable for most participants — only participant 7 selected *mustpea* (**Fig. 7**). Despite limited evidence, *mustpea* may be a less perfect calque of *blackhead*, which has been coined and entered Estonian via promotion of cosmetics and skincare products. It is noteworthy that *Head* in English means the suppuration part (of the point of swelling, boil, pimple or tumour) (Oxford University Press, n.d.); this can be categorised into one extended meaning of head — *tip*, *point* or *upper part* — based on metaphor. However, comparable use may not be common in Estonian; no evidence is demonstrated on *Sõnaveeb* (). Hence, it is possible that the less acceptability of *mustpea* should compel Participant 10 to rephrase by using code-switching to highlight the metaphorical meaning.

4. Discussion

Via an audio stimuli-based elicitation experiment, questions were addressed about how medical students make lexical selections between the Estonian word and its foreign equivalent, and whether they prime syntactic and lexical representations in the speech following the lexical selection if the acoustic stimuli are lexically and syntactically homogeneous except for synonyms set in the expression *A ehk B*. The first research question is centred on whether corpus-based frequency ratios of foreign words can predict lexical selections or whether the structural foreignness of the item plays the most important role in selection. The point regarding whether initial latency shows effects of foreignness and/or frequency, thus reflecting the lexical competition before the utterance, is also covered. In total, 10 participants finished 300 trials. After transcription and coding, descriptive statistics, a mixed-effect logistic model, and a mixed-effect linear model were used to analyse 299 responses.

Statistic results demonstrate that the variation in the frequency ratio is the only strongest predictor (*Estimate* = 7.26573, $p < 0.0001$) (**Fig. 9**, **Table. 3**) of the foreignness of the lexical selection, which, however, has failed to give a full picture of the mechanism of making selections in our small sample. As evidenced by the large confidence interval in **Figure 9**, the prediction starts becoming unstable when the frequency ratios are not skewed towards foreign lexemes; it is also vulnerable to variations in participants' behaviours. On the one hand, the quantified result tallies partially with our hypothesis that the influence of the corpus-based frequency of foreign words is strong yet limited. In a small sample, it is impossible to predict lexical selections based on the corpus-based frequency ratio when the frequency ratio is biased towards Estonian words. In other words, frequency ratios skewed to foreign words foreground the role of frequency ratio in prediction. The restricted influence, on the other hand, is associated with what runs counter to another part of my hypothesis that Estonian words should be preferred due to their advantage in high comprehensibility associated with less structure-based foreignness — across 300 trials, a strong preference is demonstrated for selecting foreign words. Hence, the finally discovered strong yet restricted frequency effect is not the study expected.

The first point to be clarified is the origin of variation. The sample size can account for the restricted frequency effect, whereas the sample size cannot sufficiently account for variation. Recall first that synonyms from specialised language vocabulary are always tricky to deal with due to the difficulty in determining their frequency via picture naming and categorising them into common names and alternative names as in the study by Dylman and Barry (2017). The present study sought to resolve the problem with the frequency ratio of foreign words reflected in written text corpora. This novel solution, inspired by and grounded on findings in both terminological studies (Džuganova, 2019; Erelt, 2007) and experimental studies (e.g., Berman & Nir, 2010; Strömqvist et al., 2002), is the key to disentangling the seeming paradox of the strong yet limited frequency effect.

Given that all pairs of words are ideally *justified parallelism* (Est. *õigustatud rööpsus*) in the word list by Dylman and Barry (2017: 169), the larger the sample size, the firmer the border between common names and alternative names (e.g., *chopper* vs *cleaver*) determined by frequency in picture naming test. However, it is hard to draw the same line between synonymous medical terms. Though the frequency ratio generalised from the written text corpora is seemingly analogical to the single-word frequency used to rank word pairs (e.g., Sarkis & Montag, 2021), the theoretical basis for the experimental design has set restrictions on the interpretation of frequency ratios. It is difficult to quantify the variations in these situations — either the mutual influence of vocabulary between the general language and the specialised language or the long-standing influence of written language on spoken language (Erelt, 2007; Strömqvist et al., 2002). When the sample size increases, the large frequency ratio of a lexeme can secure the high probability that many medical students should be exposed to that word. By contrast, the low-frequency ratio of another lexeme does not necessarily indicate the low probability that fewer students select it. It is more justifiable and cautious to state that the likelihood that medical students familiarise themselves with these terms should decrease yet not work for all students due to other factors at the individual and socio-cultural levels.

The variation would increase when the sample size shrinks. As shown in **Figure 9**, our model still predicts a high probability of the selection of foreign words but with increasing uncertainty when the frequency ratio drops from 0.6. On principle, due to the frequency effect, the likelihood of selecting foreign lexemes increases when frequency ratios are biased towards foreign words. However, this effect does not necessarily work in the opposite direction for a small sample when frequency ratios are skewed towards Estonian words. In other words, the decreasing frequency ratio lessens the effect gradually, while maintaining the prediction in favour of foreign lexemes — this is what *limited* or *restrictive* means. An embryonic conflict between the preference for foreign words and the preference for Estonian equivalents may surface. Yet the observed result is merely the tip of the iceberg because of sample size; only two participants were found whose selections were skewed towards Estonian lexemes. Thus, the results do not rule out the possibility that a group of medical students characterised by a preference for Estonian words exists. Nor will the study deny that this tendency may be as significant as the trend towards foreign words.

Variations also exist in cases of preferring foreign lexemes, though subtle (see the variations in random effect predictors in **Fig. 10** and **Fig. 11**) and thus, not affecting the status of frequency ratio in prediction. These insignificant variations are *de facto* distinctions in participants' background data (e.g., age, year of study, and English proficiency). Among these cases, the strong preference for foreign lexemes is contrary to our hypothesis. Why did this opposite tendency emerge from the data in the small sample? It is necessary to recall what Paet (2023) pointed out — that foreign words fixed already — not the Estonian equivalent — in Estonian are more commonly used, and what Erelt (2007: 143) highlighted — that distinctions in structure-based foreignness are insufficiently the base for approving or disapproving of a term. Though Estonian can be considered a highly planned ethnic language (Paet, 2023: 34; Schubert, 1989, 23), nobody has strongly suggested and/or required preferring to use Estonian words since the earliest *EKÕS* (*Eesti keele õigekeelsussõnastik*) *1918* 'orthological dictionary of Estonian language 1918' (Paet, 2023: 46). This can be explained by the long history of language contact between Estonian and other languages (incl. the history of development of modern medicine science in Estonia since the latest the end of 18^{th} century), as well as particularities of the language planning context. Granted, terms whose formation relies on native morphemes facilitate comprehensibility (Jucks and Paus, 2011), but we should not ignore the fact that foreign morphemes were borrowed together as borrowed lexemes entered Estonian. Those morphemes had been adapted and become more morphophonologically close to words in spoken language — though still sounding less Estonian. Also, medical students are more likely to be frequently exposed to medical terms based on foreign morphemes than people outside that domain. According to Bybee's exemplar theory (Bybee, 2007; Divjak, 2019: 42-45), medical students have models of memory representation to store detailed information on how medical terms have been processed via perceptual stimuli in various specific contexts.

If we go back to the debate between Rätsep (2002 [1964]) and Saari (1981), the assumed connection between native morphemes and the advantage of conveying precise meanings needs further investigation. As shown in **Figure 13**, some foreign terms dominate such as *alkoholism*, *difteeria*, *leukeemia*, *aneemia*, *menstruatsioon* etc., and Estonian equivalents will be hardly or not used. Some Estonian equivalents such as *joomatõbi* 'alcoholism, literally drinking disease' and *kurgutõbi* 'diphtheria, literally throat disease' convey only ambiguous meanings, though *tõbi* is a productive morpheme in forming compounds regarding medical concepts. Some fail to provide accuracy at the semantic level such as *verevähk*. Leukaemia is a typical blood cancer, yet other types of blood cancer exist. I will not detail much here about the reason why other examples pale in the precision of definition in comparison with their foreign equivalents. The point to highlight is that the word-formation productivity of Estonian words as an advantage (Est. *moodustusvõime*) indicates its association with comprehensibility but not with semantic precision. As indicated in Rätsep's (2002 [1964]: 176) argument, the foreign word's ability to convey a precise meaning is associated with its usage across languages (*rahvusvaheline arusaadavus* 'international comprehensibility') and the external effort (nt adaptations in language planning) to strengthen its fixation to the donor language.

Besides querying what predicts lexical selections, the study has also tried to answer whether the mechanism is competitive or non-competitive. No significant distinction in initial latency was demonstrated. Nor have any main effects of possible predictors been found. Based on these results, the mechanism can prove to be non-competitive, but we must bear in mind the limitations imposed by the small sample size. As one lexeme reaches the threshold and gets activated, another on the way to activation ceases (Oppenheim, 2018). A duration until the unselected to be used would also exist. Nevertheless, considering the influence of sample size and the lack of a control group for comparison, the interpretation must be addressed with rather more cautiousness.

Undeniably, the absence of a control group due to the time limitation is a serious weakness of the present study. It is impossible to probe into the essence of the lexical-selection mechanism without reference to trials where another type of stimuli is presented and/or groups consisting of participants of different backgrounds. Moreover, the study explored speech production following the lexical selection. No valuable evidence to prove the assumed priming effect has been found. The only extreme case in which one participant primed the *A ehk B* structure in 21 trials can be ascribed to individual preference. In addition, only one case of using code-switching to rephrase a foreign word exists. The occurrence of code-switching resulted from the quandary over rephrasing with a word more comprehensible yet less acceptable. However, it is difficult to develop possible avenues for future study on these aspects of speech production based on sporadic cases.

Finally, the interpretability of our data analysis is limited to the negative influence of the sample size; however, the finding makes a first step toward addressing a gap in research on lexical selection of specialised vocabulary with novel empirical methods. Also, the true state of language users in medical science is undoubtedly what the present study has not been settled satisfyingly. If the future study is centred on both the prediction and the mechanism of lexical selections, the sample size must be first enlarged. To observe the accurate prediction and distinctions in selections, it is necessary to compare results between control groups and test groups. Designing experiment conditions in the present study — though not performed — relied on the essential features on which the group of medical students are separated from other people outside medicine science. It could also refer to the differences in the background across participants such as age, year of study, and English proficiency, as revealed in the analysis of random-effect predictors (**Fig. 10**, **Fig. 11**). The comparison inside medical group may contribute to analysing the time-related variations in the role of word formation, with frequency effect controlled. For instance, by roughly dividing students who all study in the third year or above into two groups according to the length of practice, one idea is to design experiments in future research to examine the hypothesis that significant changes in lexical selection should emerge when participants are exposed to more communication with patients outside medical science but inside the health care system as an indispensable member. This comparison not only resolves the issue of simplifying the relation between frequency and word formation but also covers the question regarding the competitiveness of the mechanism. Moreover, as mentioned in the ***Method*** section, the approximation of students' exposure to medical terms is imperfect in sampling corpora. In the future study, in preparing sticmuli for an investigation into lexical selection, a thorough analysis of how variation between documents influences word frequency must be conducted (e.g., dispersion measures in Gries & Hilpert, 2010) and how word frequency varies with other words in the same collocations (Baayen, 2010).

Conclusion

The present study investigates how medical students make lexical selections between synonymous terms and what their speech is following the selection features. The previous literature has not thoroughly addressed issues as to (i) what other factors interacting with frequency effect are; (ii) how people make lexical selection between two synonymous abstract concepts such as terms in a speciality; (iii) whether feasible and valid methods exist to elicit lexical selections between synonymous medical terms. From a usage-based perspective (Bybee, 2007; Langacker, 1987; as cited by Divjak, 2019: 17, 41), the frequency effect plays a significant role yet interacts with other influences (Baayen, 2010; Divjak, 2019). Previous research covered this point (Baayen et al., 2007; Baayen et al., 2016; Blumenfeld et al., 2016); however, investigation into the selection between terms has not drawn sufficient attention. The most investigated topic in the existing literature is the selection between concrete words (e.g., Dylman & Barry, 2017; Oppenheim, 2018). The difficulty in finding valid methods of measuring frequency of abstract words makes it hard to extend the same topic to terms. In terminological research, the heated debate revolves around whether foreign words pale in comparison with Estonian words in the advantage of conveying precise meaning and facilitating comprehensibility (Erelt, 2007; Rätsep, 1964/2002; Saari, 1981). Nevertheless, empirical evidence (e.g., Jucks & Paus, 2011) to support either side of this discussion is quite rare.

The present research has tried to address a gap in the lexical selection of specialised vocabulary with corpus-based analysis and experimental methods. It is impossible to measure the frequency of terms by picturing naming tests. However, corpus-based analysis was aimed at sifting medical terms for stimuli by measuring the frequency of medical terms in written text corpora, considering experimental evidence that the instantaneous and continual influence of written language on spoken language (Berman & Nir, 2010; Strömqvist et al., 2002). Also, the proposal of the mutual influence of vocabulary between general and specialised languages in the terminological study (Erelt, 2007) contributes to reinforcing the ground for the method. Another difficulty to address is how to elicit responses after presenting these stimuli. With an adapted 'telephone-chain' study (Grosjean, 2008), this study managed to elicit 10 medical students' responses (lexical selections and following speech) in 300 trials involving 30 pairs of stimuli (each consisting of lexical stimuli and description).

The results of the inferential statistics via a mixed-effect logistic model contribute to answering the first research question as to whether the corpus-based frequency ratio can predict medical students' lexical selections. Despite the consistency of a frequency effect found in the existing literature, the study has evinced a different picture in making selections between synonymous medical terms. The effect of frequency represented in frequency ratios is the only and the most significant main effect among the mixed-effect predictors. Its estimation demonstrates a bias towards foreign lexemes. However, due to the small sample size, the inferential model cannot be certain when the frequency ratio drops from 60 per cent. This result is congruent with the assumption grounded in the usage-based theory that the frequency effect plays a significant yet limited role in language change and language use. Running counter to my hypothesis, a strong preference for foreign terms was found among students regardless of inherent distinctions between students. The result indicates that the structural foreignness in Word formation may have no direct influence on the precision of meanings; nor may it hinder the comprehensibility for certain language user groups. At the same time, it is necessary to be cautious to deny the likelihood that the proportion of medical students who favour Estonian words is not significant due to the small sample size.

Due to time constraints, the study failed to recruit a control group and proceed with a focus on the behaviours in an opposite condition. As a result, no significant results were demonstrated to answer other questions regarding the competitiveness of the mechanism and the rest of speech production. The future study must consider dealing with influences of sample size, inherent distinctions among participants, and variations in word frequency in specific contexts before designing experiments.

Reference

Adamska-Salaciak, A. (2013). Equivalence, synonymy, and sameness of meaning in a bilingual dictionary. *International Journal of Lexicography*, *26*(3), 329–345. https://doi.org/10.1093/ijl/ect016

Baayen, R. H. (2010). Demythologizing the word frequency effect. *The Mental Lexicon*, *5*(3), 436–461. https://doi.org/10.1075/ml.5.3.10baa

Baayen, R. H., Milin, P., & Ramscar, M. (2016). Frequency in lexical processing. *Aphasiology*, *30*(11), 1174–1220. https://doi.org/10.1080/02687038.2016.1147767

Baayen, R. H., Wurm, L. H., & Aycock, J. (2007). Lexical dynamics for low-frequency complex words. *The Mental Lexicon*, *2*(3), 419–463. https://doi.org/10.1075/ml.2.3.06baa

Baethge, C. (2008). The languages of medicine. *Deutsches Ärzteblatt International*. https://doi.org/10.3238/arztebl.2008.0037

Baus, C., Costa, A., & Carreiras, M. (2013). On the effects of second language immersion on first language production. *Acta Psychologica*, *142*(3), 402–409. https://doi.org/10.1016/j.actpsy.2013.01.010

Belhoul, R. L., & Benmoussat, S. (2023). Switching between two foreign languages in English for Medical Purposes course: Impact on Medical terminology and vocabulary Comprehension. *ALTRALANG Journal*, *5*(3), 242–258. https://doi.org/10.52919/altralang.v5i3.365

Berman, R. A., & Nir, B. (2010). The lexicon in writing–speech-differentiation. *Written Language & Literacy*, *13*(2), 183–205. https://doi.org/10.1075/wll.13.2.01ber

Bieliaieva, O., Lysanets, Y. V., & Melaschenko, M. P. (2017). Latin as a language of scientific progress: Medicine of the 16th-17th centuries. *Georgian Med News*, *262*, 120–124. http://elib.umsa.edu.ua/handle/umsa/1090

Blumenfeld, H. K., Bobb, S. C., & Marian, V. (2016). The role of language proficiency, cognate status and word frequency in the assessment of Spanish–English bilinguals' verbal fluency. *International Journal of Speech-Language Pathology*, *18*(2), 190–201. https://doi.org/10.3109/17549507.2015.1081288

Bybee, J. (2007). *Frequency of use and the organization of language*. Oxford University Press.

Chafe, W. (1994). *Discourse, consciousness, and time: The Flow and Displacement of Conscious Experience in Speaking and Writing*. University of Chicago Press.

Christoffels, I. K., Firk, C., & Schiller, N. O. (2007). Bilingual language control: An event-related brain potential study. *Brain Research*, *1147*, 192–208. https://doi.org/10.1016/j.brainres.2007.01.137

Clyne, M. G. (2003). *Dynamics of language contact: English and Immigrant Languages*. Cambridge University Press.

Computer linguistics. (n.d.-a). https://www.cl.ut.ee/korpused/grammatikakorpus/

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology Learning Memory and Cognition*, *26*(5), 1283–1296. https://doi.org/10.1037/0278-7393.26.5.1283

de Leeuw, J.R., Gilbert, R.A., & Luchterhandt, B. (2023). jsPsych: Enabling an open-source collaborative ecosystem of behavioral experiments. *Journal of Open Source Software*, *8*(85), 5351, https://joss.theoj.org/papers/10.21105/joss.05351

Deuchar, M. (2020). Code-Switching in Linguistics: A position paper. *Languages*, *5*(2), 22. https://doi.org/10.3390/languages5020022

Divjak, D. (2019). *Frequency in language: Memory, Attention and Learning*. Cambridge University Press.

Dylman, A. S., & Barry, C. (2017). When having two names facilitates lexical selection: Similar results in the picture-word task from translation distractors in bilinguals and synonym distractors in monolinguals. *Cognition*, *171*, 151–171. https://doi.org/10.1016/j.cognition.2017.09.014

Džuganová, B. (2019). Medical language – a unique linguistic phenomenon. *JAHR*, *10*(1), 129–145. https://doi.org/10.21860/j.10.1.7

Eesti keele instituut. (n.d.). Sõnaveeb. In *Sõnaveeb*. https://sonaveeb.ee/

Erelt, T. (2007). *Terminiõpetus* [Terminology]. Kirjastus Tartu Ülikool

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis☆. *Journal of Memory and Language*, *58*(3), 787–814. https://doi.org/10.1016/j.jml.2007.07.001

Gotti, M. (2015). Code-switching and plurilingualism in English-medium education for academic and professional purposes. *Language Learning in Higher Education*, *5*(1). https://doi.org/10.1515/cercles-2015-0005

Gries, S. T., & Hilpert, M. (2010). Modeling diachronic change in the third person singular: a multifactorial, verb- and author-specific exploratory approach. *English Language and Linguistics*, *14*(3), 293–320. https://doi.org/10.1017/s1360674310000092

Grosjean, F. (2008). *Studying bilinguals*. OUP Oxford.

Grosjean, F. (2013). Speech production. In P. Li & F. Grosjean (Eds.), *The Psycholinguistics of Bilingualism* (pp. 50–69). John Wiley & Sons.

Haspelmath, M. (2009). Lexical Borrowing: Concepts and issues. In M. Haspelmath & U. Tadmor (Eds.), *Loanwords in the world's languages: A comparative handbook* (pp. 35–54). De Gruyter Mouton. https://doi.org/10.1515/9783110218442

Jucks, R., & Paus, E. (2011). What makes a word difficult? Insights into the mental representation of technical terms. *Metacognition and Learning*, *7*(2), 91–111. https://doi.org/10.1007/s11409-011-9084-6

Kasik, R. (2015). *Sõnamoodustus* [Word Formation]. Tartu Ülikooli Kirjastus.

Kootstra, G. J., Dijkstra, T., & Van Hell, J. G. (2020). Interactive alignment and lexical triggering of Code-Switching in bilingual dialogue. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.01747

Krieger, J. L., & Gallois, C. (2016). Translating science. *Journal of Language and Social Psychology*, *36*(1), 3–13. https://doi.org/10.1177/0261927x16663256

Kroll, J. F., Bobb, S. C., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism Language and Cognition*, *9*(2), 119–135. https://doi.org/10.1017/s1366728906002483

Langacker, R. W. (1987). *Foundations of cognitive Grammar: Volume I: Theoretical Prerequisites*. Stanford University Press.

Levelt, W. J. M. (2000). Producing spoken language: a blueprint of the speaker. In C. M. Brown & P. Hagoort (Eds.), *The Neurocognition of Language* (pp. 82–122). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198507932.003.0004

Levelt, W. J. M. (2001). Spoken word production: A theory of lexical access. *Proceedings of the National Academy of Sciences*, *98*(23), 13464–13471. https://doi.org/10.1073/pnas.231459498

Lexical Computing CZ s.r.o. (2025, January 16). *Timestamped corpora – time annotated corpora | Sketch Engine*. Sketch Engine. https://www.sketchengine.eu/jozef-stefan-institute-newsfeed-corpus/

Myers-Scotton, C. (2006). *Multiple voices: An Introduction to Bilingualism*. Wiley-Blackwell.

Oppenheim, G. M. (2023.). Lexical selection in language production. [Manuscript Submitted for Publication]. Department of Psychology, University of Texas at Austin.

Oxford University Press. (n.d.). Oxford English Dictionary. In *Oxford English Dictionary*. https://www.oed.com/

Paet, T. (2023). *Võõrainese kinnistumine eesti keeles: keelekorralduslik ja leksikograafiline vaade* [Fixation of loanwords and foreign words in Estonian from the perspective of standardisation and lexicography]. University of Tartu.

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin*, *134*(3), 427–459. https://doi.org/10.1037/0033-2909.134.3.427

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*(02). https://doi.org/10.1017/s0140525x04000056

Raymond, W. D., & Brown, E. L. (2012). Are effects of word frequency effects of context of use? An analysis of initial Fricative reduction in Spanish. In S. T. Gries & D. Divjak (Eds.), *Frequency effects in language: Learning and processing* (Vol. 1, pp. 35–52). Mouton de Gruyter: The Hague. https://doi.org/10.1515/9783110274059.35

Rätsep, H. (2002). Sõnavara rikastamise allikatest [About the sources of enriching vocabulary]. *Sõnaloo Raamat*, 173–190. (Original work published 1964)

Saari, H. (1981). Omasõna ja võõrsõna paarid eesti oskussõnavaras (2): Omast vaatenurgast [Pairs of native word and foreign word in Estonian vocabulary]. *Keel Ja Kirjandus*, *4*, 210–211.

Sanseido. (n.d.). *Wisdom English-Japanese Dictionary*.

Santesteban, M., & Schwieter, J. W. (2020). Lexical selection and competition in bilinguals. In R. Heredia & A. Cieślicka (Eds.), *Bilingual lexical ambiguity resolution* (pp. 126–156). Cambridge University Press. https://doi.org/10.1017/9781316535967.007

Sarkis, J. T., & Montag, J. L. (2021). The effect of lexical accessibility on Spanish-English intra-sentential codeswitching. *Memory & Cognition*, *49*(1), 163–180. https://doi.org/10.3758/s13421-020-01069-7

Schubert, K. (1989). Interlinguistics – its aims, its achievements, and its place in language science. In K. Schubert (Ed.), *Interlinguistics: Aspects of the Science of Planned Languages* (pp. 7–44). Walter de Gruyter. https://doi.org/10.1515/9783110886115.7

Shi, Y., & Lei, L. (2021). Lexical use and social class: A study on lexical richness, word length, and word class in spoken English. *Lingua*, *262*, 103155. https://doi.org/10.1016/j.lingua.2021.103155

Shulman, H. C., Dixon, G. N., Bullock, O. M., & Amill, D. C. (2020). The effects of jargon on processing fluency, Self-Perceptions, and scientific engagement. *Journal of Language and Social Psychology*, *39*(5–6), 579–597. https://doi.org/10.1177/0261927x20902177

Sinisalu, V., & Soosaar, A. (2013). Meditsiiniline erialakeel vajab süsteemset arendamist ja kirjasõna suuremat tähelepanu. *Eesti Arst*. https://doi.org/10.15157/ea.v0i0.11445

Soosaar, S.-E. (2020). Eesti tüvevara ajalugu [The history of Estonian roots]. In E. Velsker (Ed.), *Eesti keele ajalugu* (pp. 337–375). Tartu Ülikooli Kirjastus.

Strömqvist, S., Johansson, V., Kriz, S., Ragnarsdóttir, H., Aisenman, R., & Ravid, D. (2002). Toward a cross-linguistic comparison of lexical quanta in speech and writing. Written Language & Literacy, 5(1), 45–67. https://doi.org/10.1075/wll.5.1.03str

Tadmor, U. (2009). Loanwords in the world's languages: Findings and results. In M. Haspelmath & U. Tadmor (Eds.), *Loanwords in the World's Languages A Comparative Handbook* (pp. 55–75). Berlin, New York: De Gruyter Mouton. https://doi.org/10.1515/9783110218442

Van Weijen Dr, D. (2012). *The language of (future) scientific communication*. Research Trends. https://www.researchtrends.com/researchtrends/vol1/iss31/3?utm_source=www.resea rchtrends.com%2Fresearchtrends%2Fvol1%2Fiss31%2F3&utm_medium=PDF&utm_campaign=PDFCoverPages

Viitso, T.-R. (2003). Structure of the Estonian language. In M. Erelt (Ed.), *Estonian Language* (Vol. 1, pp. 9–92). Estonian Academy Publishers.

Kokkuvõte

Käesolevas magistritöös on uuritud, kuidas valivad arstiteaduse õppekaval õpilased sünonüümsete terminite vahel ja mida iseloomustab nende valiku järgnevat kõne. Varasemad uuringud ei ole põhjalikult käsitlenud, (i) millised on muud sagedusemõjuga vastastikku mõjutavad tegurid; (ii) kuidas inimesed teevad leksikaalset valikut kahe sünonüümse abstraktse mõiste, nt eriala terminite vahel; (iii) kas on olemas teostatavad ja kehtivad meetodid sünonüümsete meditsiiniterminite vahel leksikaalsete valikute esilekutsumiseks. Kasutuspõhiselt seisukohalt (Bybee, 2007; Langacker, 1987; Divjak, 2019: 17, 41) mängib sageduse mõju olulist rolli, kuid on koostoimes teiste mõjudega (Baayen, 2010; Divjak, 2019). Mõned varasemad uuringud on seda punkti hõlmanud (Baayen et al., 2007; Baayen et al., 2016; Blumenfeld et al., 2016); kuid praeguseks ei ole terminitevahelise valiku uurimine piisavalt tähelepanu pööranud. Varasemas kirjanduses on enim uuritud teema endiselt konkreetsete sõnade valik (nt Dylman & Barry, 2017; Oppenheim, 2018). Leksikaalse valiku teemana terminitele laiendamise raskus ei lahene, kui kehtivaid meetodeid pole leitud ega rakendatud. Termoinoloogilises uurimistöös käib tuline vaidlus selle üle, kas võõrsõnad kahvatuvad eestikeelsete sõnadega võrreldes täpse tähenduse edasiandmise ja arusaadavuse tugevdamise eelise ees (Erelt, 2007; Rätsep, 1964/2002; Saari, 1981). Sellegipoolest on empiirilised tõendid (nt Jucks & Paus, 2011), mis toetavad selle arutelu kumbagi poolt, üsna haruldased.

Käesolev uurimus on püüdnud lahendada probleeme oskussõnavara leksikaalses valikus korpusepõhise analüüsi ja eksperimentaalsete meetoditega. Terminite sagedust on võimatu mõõta pildi nimetamise abil. Seega oli korpusepõhise analüüsi eesmärk meditsiiniliste terminite väljasõelumine stiimulite jaoks, mõõtes meditsiiniliste terminite esinemissagedust kirjalikes tekstides, võttes arvesse eksperimentaalseid tõendeid selle kohta, et kirjakeel avaldab kõnekeelele kohest ja pidevat mõju (Berman & Nir, 2010; Strömqvist et al., 2002). Samuti aitab meetodi alust tugevdada ettepanek sõnavara vastastikusest mõjust üld- ja oskuskeelte vahel terminiõpetus (Erelt, 2007). Teine probleem, mida tuleb lahendada, on see, kuidas pärast nende stiimulite esitamist vastuseid esile kutsuda. Kohandatud „telefoniahela" uuringuga (Grosjean, 2008) õnnestus selles uuringus kutsuda esile 10 arstitudengist vastuseid (leksikaalne valik ja sellele järgnev kõne) 300 katses, mis hõlmasid 30 paari stiimuleid (igaüks koosneb leksikaalsetest stiimulitest ja kirjeldusest).

Segamõjulise logistilise mudeli abil saadud seletava statistika tulemused aitavad vastata esimesele uurimisküsimusele, kas korpusepõhine sagedussuhe suudab prognoosida arstitudengite leksikaalseid valikuid. Vaatamata olemasolevas kirjanduses leitud sagedusmõju järjepidevusele, on uuring sünonüümsete meditsiiniterminite vahel valiku tegemisel näidanud teistsugust vaadet. Sagedussuhetes avalduv sageduse mõju on segaefektiga tegurite seas ainus ja kõige olulisem põhiefekt. Sel teguril põhinev hindamine on üle ootuste suur, mis näitab suundu võõrlekseemide poole. Kuid väikese osalejate valimi tõttu ei saa logistiiline mudel olla kindel, kui sagedussuhe langeb 60 protsendilt. See tulemus on kooskõlas kasutuspõhisest teooriast lähtuva eeldusega, et sagedusmõjul on keelemuutuses ja keelekasutuses oluline, kuid piiratud roll. Vastupidiselt minu hüpoteesile leiti üliõpilaste seas tugevat võõrterminite eelistamist, sõltumata õpilaste sisemistest erinevustest. Tulemuses selgub, et sõnamoodustuse struktuurivõõrus ei pruugi tähenduse täpsust otseselt mõjutada ega takista arusaadavus teatud kasutaja rühmadele. Samas tuleb olla ettevaatlik, välistades tõenäosust, et arstitudengite osakaal omaterminite eelistamises olla märkimisväärne.

Ajapiirangute tõttu ei õnnestunud uuringus otsida kontrollrühma ja vaadelda vastupidises olukorras käitumisele. Selle tulemusena ei näidatud olulisi tulemusi, mis vastaksid teistele küsimustele leksikaalse valiku mehhanismi konkurentsivõime ja ülejäänud kõne tootmise osas. Tulevases uuringus tuleb enne katsete kavandamist arvestada valimimahu mõju, osalejate kaasasündinud erinevusi ja sõna sageduse muutusi konkreetsete kontekstide järgi.

Non-exclusive licence for the reproduction and publicization of the thesis

I, Zhenyu Wang,

1) grant the Unviersity of Tartu a free permit (non-exclusive licence) for the reproduction of "Lexical selection between synonymous medical terms", supervised by Virve-Anneli Vihman, with the purpose of preservation, incl. for adding to the Dspace digital archives until the expiry of the term of copyright;

2) grant the Unviersity of Tartu a permit to publicise the thesis specified in Point 1 via the web environment of the University of Tartu, incl. via the DSpace digital archives, under the Creative Commons CC BY NC ND 4.0, which allows to reproduce, distribute the work and communicate it to the public, by giving appropriate credit to the author, and prohibits the creation of derivative works and the commercial use of the work until the expiry of the term of copyright;

3) I am aware that the author retains the rights specified in Points 1 and 2;

4) I confirm that granting th enon-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legalisation.

Zhenyu Wang

26/05/2025