

Functions of English and Latin in the Parliament of Finland 1970-2020

Anna Ristilä^a

^a *University of Turku*

Abstract

The Finnish parliament, *eduskunta*, operates in two national languages, Finnish and Swedish. Other languages, such as English and Latin, are also present in the plenary speeches in trace amounts. This paper explores the functions English and Latin may have in *eduskunta*, and contemplates whether they are used as elite closure, and if so, in what contexts. The questions are approached with quantitative and qualitative methods: first the use cases of English and Latin in the plenary speeches of the Finnish parliament between 1970 and 2020 are quantified, then ten possible intended functions for English and Latin use are identified by close reading the immediate contexts of each use case. Finally, the distributions of the use cases and their functions along topics are examined, and the possibility that English and Latin could be used as tools of elite closure is discussed.

Keywords

English, Latin, *eduskunta*, parliament of Finland, multilingualism, elite closure

1. Introduction

Different languages and language varieties used in a society have different amounts of political capital. The languages/varieties with the most political capital are those used in government, business, and in higher education (Myers-Scotton, 2002: 35). If such varieties are only available to some people, it can practically exclude people without such linguistic skills from positions of power. For the elite who do have the skills it can be an advantageous strategy to maintain the linguistic differences; this type of strategy has been named *elite closure* by Myers-Scotton (1993; 1997; 2002).

Latin used to be a powerful tool for elite closure during its time as the lingua franca of the learned Europe. Today Latin is practically a dead language since no-one speaks it as their mother tongue, but it has some active L2 speakers (Engelsing, 2017; Coffee, 2012). There are few studies concerning contemporary uses of Latin, but the main areas where Latin is still used are science (Roelli, 2021), law (Gałuskina & Sycz, 2013), Catholic Christianity, and the occult (Banner, 2021; Piętka, 2016). In the modern world English has taken more than just the place of Latin; it has become the closest to a true global lingua franca than any language before. It also seeps widely into other languages through borrowing and codeswitching phenomena (e.g. Furiassi et al., 2012). However, even this new lingua franca is still often the language of the learned and the elite, albeit much less exclusively so than what Latin used to be. For example, in many African and Oceanian countries and in India English is the official language despite not being the majority language. Also, varieties containing English codeswitching have been reported as acts of upper-class identity performance (e.g. Sanei, 2022; Jahan & Hamid, 2019; Matus-Mendoza, 2002).

In Finland there are two national languages, Finnish and Swedish. Finnish is the majority language, while Swedish speaking Finns are a historical minority currently estimated at 5.6% of the population (Statistics Finland 11rm, 2023) and with their own political party, the Swedish People's Party of

Digital Parliamentary Data in Action (DiPaDA 2024) workshop, Reykjavik, Iceland, May 28, 2024.

EMAIL: ankrris@utu.fi

ORCID: 0000-0003-0236-5401



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).



Digital Humanities in the Nordic and Baltic Countries Publications – ISSN: 2704-1441

Finland. English has no official status and only 0.7% of the population state English as their native language (Statistics Finland 11rm, 2023). However, the use of English in the Finnish society has steadily increased in recent decades and English functions in many ways as a lingua franca (Peterson, 2022; European Commission, 2012). In a 2024 EU Barometer survey 81% of Finnish respondents claimed they could have a conversation in English, which is an 11%-unit increase from the previous survey in 2012 (EC, 2024; EC, 2012). The number will probably keep increasing as English continues to spread outside the traditional contexts of media, business and education (Leppänen & Nikula, 2007).

The Finnish parliament, *eduskunta*, operates in the two national languages, Finnish and Swedish. However, many languages are still present in parliamentary speech in small quantities, often as rhetorical devices. In this paper the occurrences and functions of English and Latin, the current and former *lingua francae*, in the Finnish plenary speeches between 1970 and 2020 are studied. The results will be interpreted from the point of view of a prestige language as political capital and as possible elite closure. The research questions are: 1) What functions do English and Latin play in *eduskunta*? 2) Are they used as elite closure and if so, in what contexts?

Studying how much, with what intentions and in which topics the current and former lingua franca languages are used in the parliament may help us better assess the process of how English is spreading and how it contributes to elite closure. This study will also give a better understanding of English and Latin use specifically in a political context, and of how and why these languages can be used as political tools and devices. As far as the author's knowledge, linguistic elite closure has not been studied with quantitative methods from parliamentary speeches before.

2. Background

2.1. Elite closure

Elite closure denotes "a type of social mobilization strategy by which those persons in power establish or maintain their powers and privileges via linguistic choices" (Myers-Scotton, 1993: 149). There are two forms of elite closure, strong and weak. The strong form is typically exercised through official language policies and very effectively excludes non-speakers of a language from political access; it is typically present in countries where universal education of the elite variety is not available (Myers-Scotton 2002: 35). This is the case e.g. in many African countries and India (Norro, 2022; Hickey, 2019; Bedi 2019). The weak form is more subtle, informal, and common since it is expressed through linguistic repertoires (e.g. sociolects) and use patterns (e.g. code switching) that can potentially be learned by anyone through universal education (Myers-Scotton, 1993:151-152). Weak elite closure is more or less present everywhere, especially in the western countries (*ibid.*), but it is more difficult to address because of its subtlety and links to identity building.

One way to study weak elite closure is to examine speech patterns, such as foreign language use, in contexts where elite identity performance is expected, such as in the parliament. Parliaments are public stages where elite identity performance can help legitimize and elevate the speaker's authority not just in front of their peers but also of the media and the public. On the other hand, parliaments are formal institutions with specific norms, rules, and expectations that shape how identity can be performed. Identity performance in a parliamentary context is very strategic, as the speaker must conform to institutional expectations while still distinguishing themselves for political purposes. Elite identity performance through speech patterns can impact how the speaker is perceived in terms of credibility, leadership, and trustworthiness, but at the same time it may work as a tool of exclusion.

2.2. Foreign language use cases

There are many kinds of linguistic use patterns that could be mapped to study elite closure. Many previous studies have focused strictly on code switching. The focus of this article is more widely on (extra- and intra-sentential) code switches, loanwords, quotes, or any kind of embedded language use clearly distinguished from the matrix language, or the language that provides most of the semantic and morphosyntactic procedures in a bi- or multilingual setting (Myers-Scotton, 1997[1993]: 75). These will be collectively referred to as *embedded language use cases*.

Recognizing what counts as embedded language is not always straightforward. Quotes are easy to recognize, most code switches as well, but loanwords can be more or less integrated into the borrowing language. A helpful classification is to see how fully they have been translated (Haugen, 1950: 214-215): fully, partially, or not at all (loanwords proper). Partial and untranslated loans were evaluated on a case-by-case basis and counted if they clearly expressed 'foreignness' in the context in question. The feature for identifying 'foreignness' was orthography and/or pronunciation: a loanword that has not been adapted and still retains the original orthography has been defined as a foreignism (Haspelmath & Hapdor, 2009: 42) (e.g. *high-tech*, compare to e.g. *skuutti*, "electric/kick scooter"). Foreignisms are close to single-word code switches but are more established (*ibid.*: 43), which means they are easily omitted from studies that focus solely on code switches.

2.3. Functions of use cases

Every embedded language use case has some function – or multiple functions – that it performs in the context it appears in. Because the parliament is a venue for extremely strategic language use, the functions identified in this paper have been formulated from the point of view of the speaker's probable *intention*, either conscious or subconscious, rather than from a general usefulness perspective. In other words, the functions should always answer the question "what seems most likely to be the speaker's intended outcome for this use case" instead of a more general "why is the use case here". Many sources that list functions for any type of embedded language use have not always formulated the functions with (conscious or subconscious) intention as the centre. For example, Hockett (1958 [1964]: 404-405) gives two well-known motives for borrowing: *prestige* and *need-filling*. If rephrased with intention, they would be something akin to *to sound more profound* and *to fill a lexical gap*. Gumperz (1982:75-80) lists five conversational functions for code switching: *quotations*, *addressee specification*, *reiteration*, *message qualification*, and *personalization versus objectivization*, some of which lack an explicit intention part. For example, Gumperz's *addressee specification* does include an explicit intention: the speaker intends to focus their message to a certain addressee, so they use an embedded language for the sake of that focus. But as for Gumperz's *quotations*, which could be either direct quotes or reported speech, the definition as-is does not reveal any specific purpose for embedded language use. A probable intended function for any quote or reported speech could be *clarification*, another *appealing to authority*. Romaine (1995 [1989]: 161-164) elaborates on Gumperz with more attention on intention and adds especially the concepts of topic and comment to explain what Gumperz called *message qualification*. Gumperz's categories with Romaine's commentary were used as a starting point when identifying functions of embedded language use cases in the parliamentary context. The process of defining the function categories is explained in the Materials and methods section, and the final categorization is introduced in the Results section.

3. Materials and methods

All the plenary speeches in the Finnish parliament since 1907 have been collected and made available online and in computer readable format in the Semantic Parliament project (Parlamenttisampo, Hyvönen et al., 2024). The dataset used in this paper consists of years 1970-2020 and includes the actual speeches – born-digital after 2010, converted from originals with optical character recognition before that – and metadata which contains e.g. date and speaker information. The used subset included approximately 650'000 speeches and over 100 million words. English and Latin use cases were detected from the subset, and the most probable intended functions were defined for each detected use case.

3.1. Language use case detection

In order to analyse the embedded language use cases, they first had to be identified. Both automatic detection and manual verification was used for the detection of English and Latin, but in different ways. The automated tool used for English detection was a Python library called Lingua (Lingua 2.0.2, 2023). Since the embedded language could be either Finnish (ca. 97% of the corpus) or Swedish, a separate

Swedish sentence corpus was created along the way. First, speeches were iterated one sentence at a time through a language detector object that could differentiate between Finnish, Swedish and English. If the confidence value of a sentence for Swedish was over 0.5 it was accepted as Swedish. Then the collected sentences were manually browsed through and the few ones that were clearly not Swedish were removed. This resulted in 8695 Swedish sentences.

Since English appeared typically in smaller units than sentences, to automatically catch as many use cases as possible a two-word sliding window was used to feed word pairs to Lingua's language detector object. Again, the differentiation was made between Finnish, Swedish and English, but this time if the language with the highest confidence value was English, the word pair was accepted as English. This resulted in approximately 1000-2000 hits per year which were browsed through manually. Only about 3% of these were considered as truly English use cases and were picked out separately (N=2337). Typical false positives from Lingua were either Swedish language or contained inflected versions of Finnish words *komissio* ("commission") or *oppositio* ("opposition"). Setting a detection probability threshold for English in Lingua could have improved the low percentage of true positives and sped up the process, but any threshold would have potentially increased false negatives. Since English use cases were dispersed very sparsely, it was not possible to reliably control for false negatives. Because of this, and because it was considered important to catch as many use cases as possible, the manual reading was seen as necessary to ensure accuracy.

In case of Latin the process was very different due to automatic language detection not working well with Latin. A list of 181 common Latin phrases was compiled and these phrases were searched with a simple regex search with Python (see Appendix 1 for the list of searched and found Latin phrases). Only 59 of these phrases appeared in the corpus at least once and all found instances (N=795) were collected. Since the list of searched phrases was finite, it is possible that some Latin use cases were not caught, but again due to extreme sparsity it was not possible to approximate how many.

3.2. Topic modeling

The topics present in the corpus were defined with an LDA (latent Dirichlet allocation) topic model created in a previous study (Ristilä & Elo, 2023). Since the details of the model creation have been described in the aforementioned article, only the method and the main characteristics of the model are explained here.

Topic modeling is an unsupervised machine learning method that uses corpus vocabulary to define 1) what topics are present, 2) how probable each topic is in a text passage, and 3) how probably each word is related to any of the topics. Simply put, the model sees all topics in any given text, just in different proportions. The topics provided by the model are abstract collections of words that computationally "go together" and a human needs to interpret what the topic is actually about. For example, one of the topics defined was interpreted as *education* as it was strongly associated with words such as "university", "education" and "school".

The original topic model was trained with the same parliamentary speeches as used in this study but only from years 1980-2010. A longer time span was not modeled because creating a model from a large dataset is very time consuming and the supercomputer used for the training had rather strict time constrictions. Also, since the topics of the chosen model were very broad, they were considered to be similarly present also in the preceding and succeeding decades of 1970s and 2010s in a way that the model would still provide reliable results for texts from these decades. The model defined 26 topics: *administration, agriculture, budget, commerce, crime, democracy, development cooperation, education, employment, energy, foreign and security policy, general, housing, legislation, law proposals, parliamentary factions, pensions, public sector, question time, regionality, social and health care, social benefits, social problems, taxation, traffic and transport, and voting.*

In order to accurately define the topics where an embedded language use case appeared, some amount of text from around every use case had to be fed into the topic model. The topics were pre-defined for entire speeches, but since speech lengths varied a lot (from single word remarks to multi-page monologues), a five-sentence passage (later "context passages") – the sentence where the use case appeared, plus two sentences from both sides of the use case sentence – was decided to be enough context to represent the immediate surroundings of any use case. If the use case appeared in the

beginning or end of a speech, five sentences worth of immediate context was included. Since the topic model was only trained in Finnish, any Swedish sentences were first translated with the European Commission's online machine translation tool (eTranslation v13.3, 2024, with "General Text" as the setting) and lemmatized with Turku Neural Parser Pipeline (Kanerva et al., 2018), before being processed through the monolingual topic model. Machine translation has been shown to be a viable method to enable use of multilingual texts with topic models (e.g. Maier et al., 2022).

Both machine translation and lemmatization were affected by errors in optical character recognition of the original paper documents. Especially older documents had suffered from low quality paper (data provided by the Semantic Parliament project). This resulted in vocabulary mistakes that could potentially affect the topic model. However, this was mostly the case before 2011, after which the data was digital-born and significantly less mistakes were noticed.

3.3. Defining functions

Only the most probable intended function was decided for each identified embedded language use case of English (N=2337) and Latin (N=795). This was done by close reading the context passages surrounding the use cases and considering what the speaker probably wanted to signal with the language choice. The focus was on the embedded language vocabulary and the immediate context, but metadata (e.g. speaker, party affiliation, year) was also taken into consideration when applicable. Since no existing categorization was fully compatible with the focus on intention, Gumperz's system (described in Gumperz 1982 and Romaine 1995) was used as a basis for a new one that was constructed by close reading and suggesting categories that would suit the context of the dataset. Assigning the categories for each use case was a semi-iterative process. If a use case could be categorized with an existing label (Gumperz's original categories or modified with Romaine's commentary), it was labeled as such, otherwise a new label would be thought of. Whenever at a later stage a use case was found not to fit in any of the previous labels, the categorization system was revised and either a new label was added or a previously used label was renamed, in which case the use cases with the redundant labels were re-evaluated and re-labeled.

It should be underlined that the functions represent the author's *interpretation* of the most probable functions intended by the speakers in the context of the parliament, not actual functions. It would not be possible to be certain of the actual intended functions, as they may be subconscious as well as conscious and even interviewing the speakers would not give fully reliable answers. Any listener's interpretation would also be different, subjective, and perhaps even more difficult to average out. Furthermore, usually multiple parallel functions could have been designated for any given use case, but only the most probable ones were marked, which was likely to oversimplify the situation. Lastly, though suitable category labels were discussed with other researchers, since the author was the sole annotator the resulting categorization of the data is still highly subjective. Planned subsequent studies with similar data will use multiple annotators.

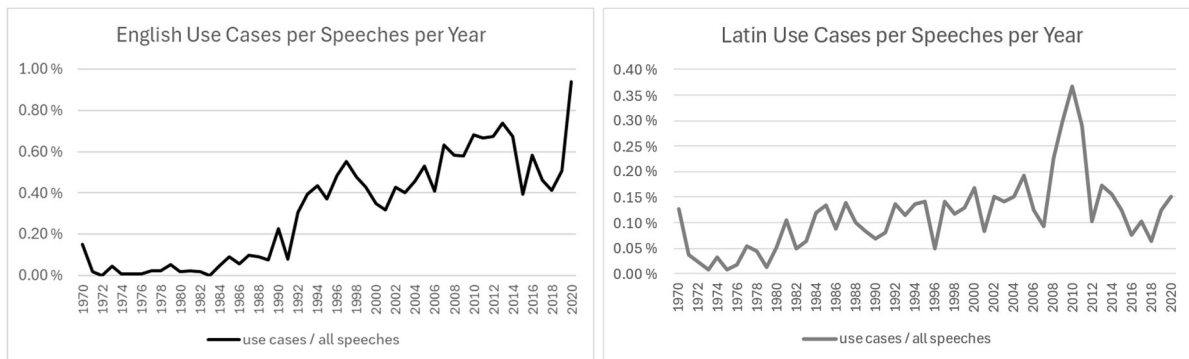
4. Results

After identifying the use cases of English and Latin, interpreting their functions, and running the five-sentence context passages through the topic model, distributions of the use cases, functions and topics were examined. In the following subsections the main findings will be presented: First, the use case frequencies of English and Latin; Second, the topic distributions around English and Latin use cases; Third, the function definitions and the distributions of the functions for each language; Fourth, the topic distributions around different functions.

4.1. Use case frequencies

English was used much more than Latin, which was unsurprising. Yearly differences can be seen from figures 1a and 1b. The most striking points for English are the steep increase in use around the 1990s, when Finland joined the EU (1995) and the Internet became more easily available to the public. The

impact of EU may be seen in the use of phrases such as “moral hazard”, “peace making”, and “joint implementation”, whereas the impact of the Internet may show in phrases such as “high-tech” and “free net”. For Latin, 2010 is a clear peak year which can be explained with heated discussions about *de minimis* EU aid.



Figures 1 a and b. English and Latin use cases per number of speeches per year. Notice the difference in scale between y axes.

Table 1.

The most used English phrases in select peak years.

1993	1997	2010	2013	2020
moral hazard (9)	high tech (14)	no bail out (12)	cleantech (21)	take away (36)
high tech (7)	peace making (4)	must carry (7)	pooling and sharing (10)	no bail-out (14)
free net (4)	working poor (3)	not in my backyard (4)	private label (6)	carry back (6)
fifty-fifty (3)	masking factor (3)	so what (3)	total return swap (4)	label (5)
we cannot safely leave politics to politicians (2)	joint implementation (3)	fresh start (3)	no bail-out (3)	end of waste (4)

Table 2.

The most used Latin phrases in select peak years.

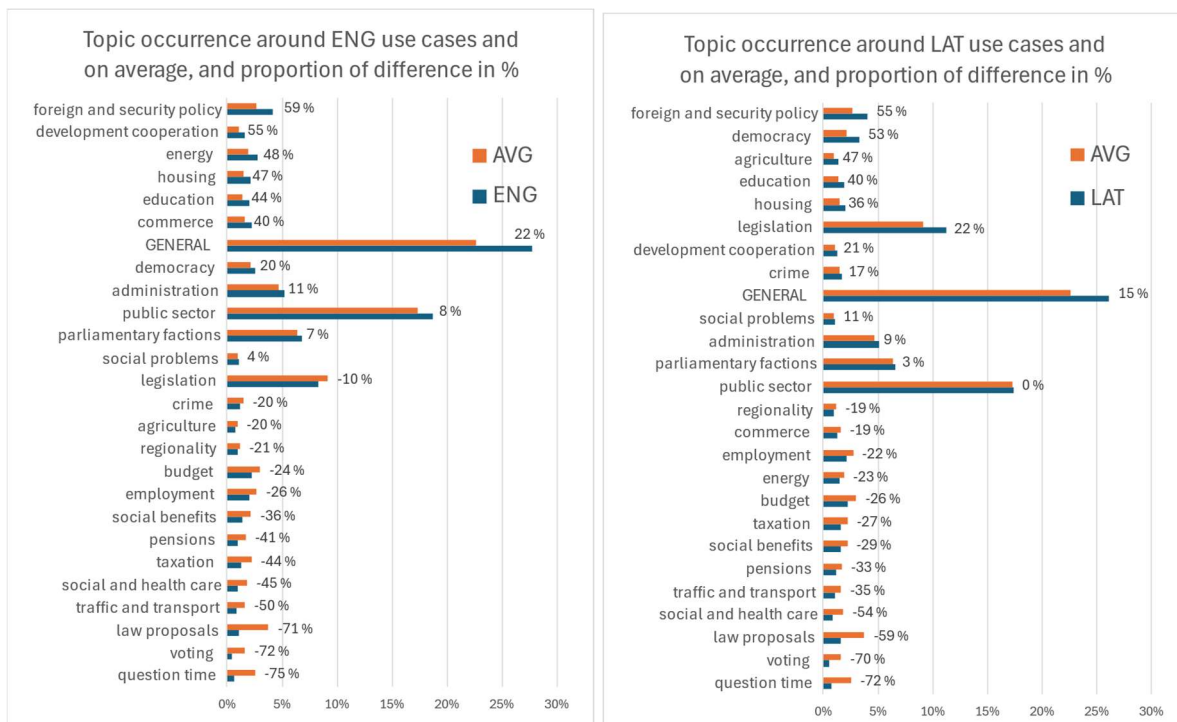
1993	1997	2010	2013	2020
de facto (7)	de facto (13)	de minimis (31)	ne bis in idem (9)	de facto (13)
sui generis (4)	bona fide (3)	de facto (12)	de facto (7)	de minimis (5)
in dubio pro reo (4)	vestigial terrent (1)	lex in casu (4)	status quo (2)	sui generis (1)
bona fide (1)	nomen est omen (1)	ad hoc (4)	de minimis (2)	ad hoc (1)
ad hoc (1)	ad hoc (1)	bona fide (2)	ad hoc (1)	ceterum censeo (1)

Making clean frequency lists of the found English phrases was difficult because of various forms the phrases could take: the words could appear with (English or) Finnish inflection (e.g. *take awayn*, *take awaysta*), be used as compound words (e.g. *fifty-fifty-periaatteella*), and also exhibit variation in spelling (e.g. *know how*, *know-how*, *knowhow*). Removing variations is not an easy task, so it was done only for five peak years of 1993, 1997, 2010, 2013 and 2020. The top five most frequent use cases for each peak year are shown in table 1, and similar frequencies for Latin are given in table 2.

To give an idea of typical English phrases found without addressing variation, the ten most frequent English phrases with all different forms left as-is are listed: *cleantech* (43), *cleantechin* (26), *fifty-fifty* (22), *so what* (14), *high techin* (14), *know-how* (11), *masking factor* (10), *high tech* (9), *fair play* (9), *deadline* (9). The total frequencies of found Latin phrases can be found in Appendix 1, the most frequent overall being: *de facto* (329), *de minimis* (130), *ad hoc* (52), *ex tempore* (32), *status quo* (27), *ne bis in idem* (25), *sui generis* (20), *bona fide* (14), *nomen est omen* (9), *pacta sunt servanda* (9).

4.2. Topics around languages

After defining the topic distributions around every use case, the distributions were summed together per language to form two comprehensive distributions that would represent the English (ENG) and Latin (LAT) "subcorpora". These distributions were then compared to the average topic distribution of the full Finnish-language corpus of years 1970-2020. In figures 1a and 1b the sub and full corpora distributions and their proportional differences are given. Positive difference means the topic is more likely to appear around English or Latin than in average text, negative difference means the topics is less likely to appear around the studied languages than in average text. For example, the topic *energy* is proportionally almost 50% more likely to appear around English use cases than it would appear on average, though its average prevalence is only around 2%.



Figures 2a and 2b: Topic distributions around English and Latin use cases and on average. The x axis shows percentages for the orange and blue topic bars, the numbers after each bar pair are the differences between the bars (ENG or LAT and AVG) as proportions from the average (AVG).

The topics that mostly preferred English were related to international affairs: *foreign and security policy* (59%) and *development cooperation* (55%). Other topics that also clearly preferred English (>40%) were all fundamental pillars of societal infrastructure and development: *energy*, *housing*, *education* and *commerce*. The topics that preferred Latin the most were probably connected through juridical discussions regarding them: *foreign and security policy* (55%) and *democracy* (53%). *Agriculture* (47%) appeared strongly with Latin due to one extremely frequently used term, *de minimis*, referring to certain type of EU aid. Three topics, *agriculture*, *legislation* and *crime*, that avoided English, preferred Latin. On the other hand, topics *energy* and *commerce* which preferred English avoided Latin. The three topics that strongly avoided both English and Latin represented quite formal

contexts: *question time* (ENG -75%, LAT -72%), *voting* (ENG -72%, LAT -70%), and *law proposals* (ENG -71%, LAT -59%).

4.3. Functions of languages

For both English and Latin ten functions were identified during the close reading of the use cases and their surrounding context passages. These were divided into two categories: practical functions, which represent intention to communicate efficiently and in a way that most people can understand, and rhetorical functions, which represent intention to accent and enhance expression. It was assumed that practical functions did not take as much part in elite closure as the rhetorical ones, since a willingness to be practical does not logically match well with a willingness to differentiate from others, while intentional linguistic (self-)enhancement can be more easily interpreted as a willingness to differentiate and exclude. Explanations and examples of functions are given below but, again, it is important to remember that any given use case usually employs several functions at once.

Practical functions included four functions:

- TRAN; to discuss **translations** or meanings of foreign words.
Example: *Oli muuten vaikea löytää sopivia suomenkielisiä vastineita sellaisille sanoille kuin ruotsin "lagarbete" ja englannin "team work"*.
"Let me tell you, it was difficult to find suitable Finnish equivalents for words such as 'lagarbete' of Swedish and 'team work' of English."
- CLAR; to **clarify**, often hinting that the English or Latin version is better known or at least may help better understand the concept or phenomenon.
Example: *Tätä kutsutaan tieteessä ja ekologiassa masking factor -ilmiöksi*.
"In science and ecology this is called the 'masking factor' phenomenon."
- ESTA; **established** use, often no Finnish equivalent in common use.
Example words: *skinhead, hightech, cleantech*.
- EFFI; to be **efficient**, when the concept in English is shorter or quicker to say than the Finnish equivalent.
Example: *Siitä tehdään parhaillaan virkamiesten think tank –työskentelyä*.
"Civil servants are currently doing 'think tank' work on it."
(Finnish equivalent for 'think tank' is *ajatushautomo*, which is six syllables, while the English term is only two syllables.)

Rhetorical functions included six functions, as shown below:

- ATTI; to express **attitude** or emotion.
Example: *...tutkaillaan, sopiiko se sen rahoituksen piiriin, all right, ei sovi, nakataan jälleen seuraavaan...*
"...let's see if it is suitable for this funding, 'all right', no it isn't, let's throw it to the next one..."
(Here the English part is taken as an expression of 'someone doesn't care', or 'someone shows disregard', so this use case also functions as DIST; see below.)
- EMPH; to **emphasize** the contents of the message.
Example: *Kun hän on hyvässä kunnossa, in good order, kun hän on eläväinen, henkisesti valveutunut ja ruumiillisesti hyvässä kunnossa...*
"When they are in good condition, 'in good order', when they are lively, spiritually enlightened and physically in good condition..."
- HUMO; to add **humour**.
Example: *USA:ssa on Bill Clinton, Bob Hope ja Johnny Cash, Suomessa taas Esko Aho, no hope and no cash*.
"In the USA there is Bill Clinton, Bob Hope and Johnny Cash, in Finland there is Esko Aho, 'no hope and no cash!'"
(It should be noted that it is irrelevant whether the audience thinks of the use cases as funny or not, just that the speaker intends their utterance to be funny.)

- SELF; to **self-emphasize**, attempting to emphasize personal linguistic and cultural knowledge, to be trendy or to show off.
Example: *Think globally, act locally*.
- AUTH; to appeal to **authority**, either by quoting/paraphrasing a person or by using a proverb or saying.
Example: *...englantilaiset ovat useissa yhteyksissä toistaneet vanhaa brittiläistä viisautta "right or wrong - my country"...*
"The English have in many contexts repeated an old British wisdom 'right or wrong – my country'".
- DIST; to **distance** oneself from the topic or reported speaker, usually in negative contexts.
Example: *...kun matkustivat Kiinaan, ilmoittivat, että business is business and human rights are human rights...*
"When they travelled to China, they stated that 'business is business and human rights are human rights'".

The time difference between the original utterances and the time this paper is written raises some small issues, since the examined period reaches more than 50 years into the past from the time of writing. Many foreign language elements may have become practical necessities over time through original unwillingness to find or use suitable domestic language variants due to elite closure. Diachronic changes in word uses are also constantly taking place, meaning that an annotator making an interpretation about a use case today might have made a completely different interpretation 50 years ago. It was not possible to address these issues in detail, but they have been considered when interpreting especially the older speeches.

Figures 3a and 4a show that English was used mostly for practical functions and that there has been a clear rise in the use of English since the 1990s, which is the decade when Finland joined the EU and when internet connections started to become common in households. There is also a steep spike in 2020, with “take away” being the most popular phrase during that time (i.e. take-away food portions during the Corona pandemic, see example (a)) but not enough to fully explain the spike. English’s functions spread more evenly than Latin’s (2a), but still concentrated more on EFFI, CLAR, ESTA and SELF. The most used function for English was EFFI, but CLAR was also surprisingly prominent, hinting that English is often seen as the better-known version for many concepts and thus given as a clarifying aid.

Latin use concentrated strongly on two functions, EFFI and SELF (Figure 3b). *De facto* was the most prominent Latin phrase overall (329 use cases) and was typically categorized as either EFFI or SELF. Both functions almost always overlapped, which is demonstrated in examples (b) and (c). The first example (b) was categorized as SELF, on the grounds that though *de facto* has probably replaced the Swedish word *faktiskt* making the sentence that much easier to pronounce (EFFI), the sentence overall is quite long and “inefficient”, which supports the interpretation that Latin was used more for the sake of raising the register and consequently the status of the speaker. In example (c) *de facto* was used the “correct” way as an antonym to *de jure*, both of which could be said in Finnish terms but using the well-known Latin terms is more accurate, quicker, and thus more efficient (EFFI).

(a) *Ravintolat pystyvät jossain määrin sopeutumaan tilanteeseen **take away** -myynnillä ja kotiintoimituksilla. Kotiintoimitukset eivät kuitenkaan ole mahdollisia alkoholijuomien osalta (...)*

” Restaurants can to some extent adapt to the situation with **take away** sales and home deliveries. However, home deliveries are not possible for alcoholic beverages (...)

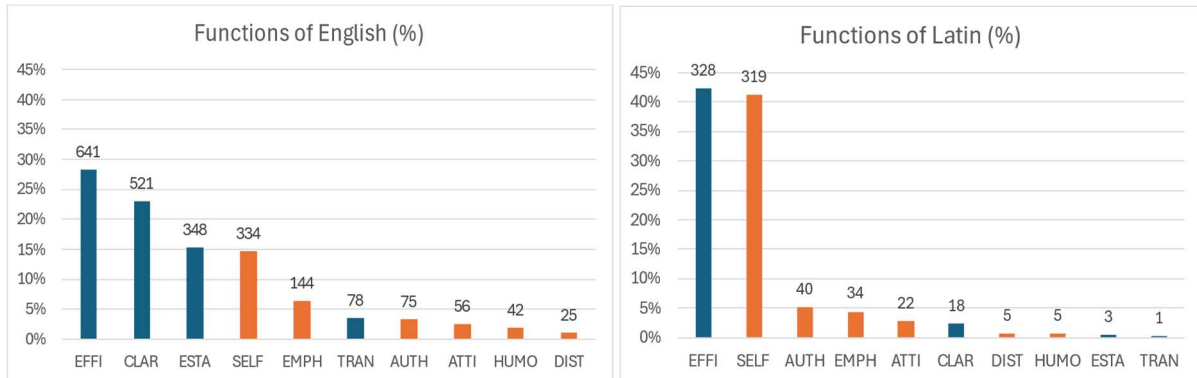
(b) *Fler och fler konsumenter har dock sakta men säkert börjat ifrågasätta och fråga sig vad maten som finns att köpa i dagligvaruhandeln innehåller, om den **de facto** är så trygg, hälsosam och välsmakande som det sags.*

“However, more and more consumers have slowly but surely begun to question and wonder what the food available for purchase in the grocery store contains, if it is *de facto* as safe, healthy and tasty as it is said.”

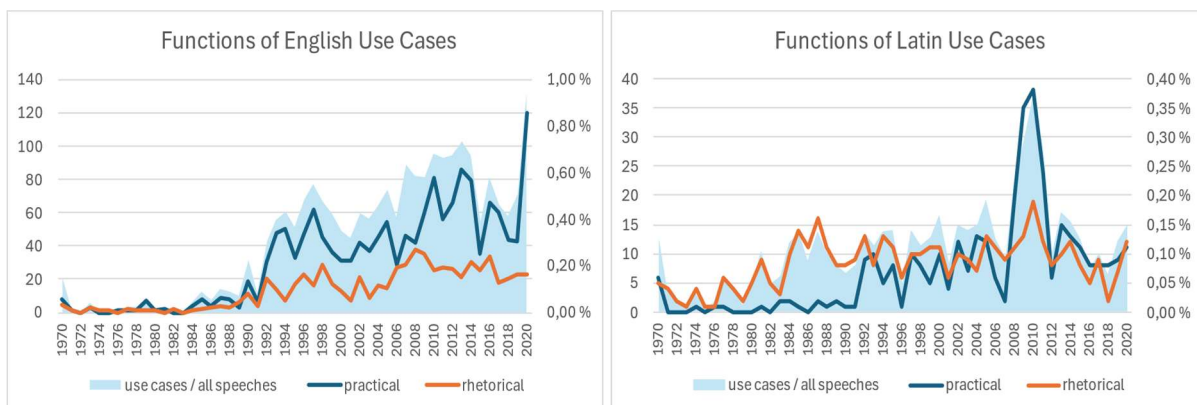
(c) *Krimin niemimaa on tällä hetkellä de facto venäläisten joukkojen valvonnassa. Venäjä piirittää alueella olevia Ukrainan tukikohtia ja rajavartioasemia.*

“The Crimean peninsula is currently *de facto* under the control of Russian forces. Russia is besieging Ukrainian bases and border guard stations in the region.”

Diachronic use of Latin has been rather steady for both practical and rhetorical functions (Figure 4b), except for the spike around 2010 which can be explained with intensive discussions about *de minimis* aid, a type of European Union funding. The small rise in Latin practical functions during the 1990s can be in part explained with the same phenomenon. It is also possible that Latin use has increased because of English use, since English texts often employ a lot of Latin.



Figures 3a & 3b: Proportions of functions of English and Latin. Practical functions are in blue, rhetorical functions in orange.



Figures 4a & 4b: Frequencies of practical and rhetorical functions of English and Latin use cases per year (blue and orange lines). Note the difference in scale between the left y axes. The light blue area shows the proportion of use cases per all speeches given every year and it uses the secondary y axis on the right.

Three different translation strategies were observed for functions: either a use case had some kind of a translation or explanation (after the use case), had no translation, or was used as a translation for a concept or idea which was first explained in Finnish or Swedish. Example (d) shows how a single word equivalent is given as a translation for an English term, and (e) exemplifies a longer explanation. Example (f) shows a use case of Latin with no translation, suggesting that the speaker either thinks everyone knows the concept or deliberately wants to speak abstrusely. Example (g) shows an English (ATTI) use case with no translation, and (h) shows how an English concept is used as a translation or explanation for a concept first given in Finnish.

(d) *Siellä sillä on aivan oma terminsä. Se on looting elikkä ryöstö, ihan puhtaasti sanottuna ryöstö.*

“There it has its own term. It's **looting**, that is, robbery, clearly said, robbery.”

(e) *Meillä on tiukemmat säännöt euroalueella, meillä on **fiscal compact** eli tämmöinen taloussopimus, jossa on muitakin kuin euromaita (...)*

“We have stricter rules in the euro area, we have a **fiscal compact**, i.e. an economic agreement which also includes non-euro countries (...)”

(f) *Kristillinen liitto vaatii 0,0 promillen rajaa tieliikennelakiin. Lopuksi, arvoisa puhemies, kristillisen liiton **ceterum censeo**: Alkoholijuomat on poistettava elinkustannusindeksistä.*

“The Christian alliance demands a limit of 0.0 per mille in the road traffic law. Finally, honorable Speaker, **ceterum censeo** of the Christian alliance: Alcoholic beverages must be removed from the cost-of-living index.”

(g) *Det finns nu ett **window of opportunity** till en fredlig lösning enligt det som kallas medelvägskonceptet, inte självständighet utan självstyrelse, och (...)*

“There is now a **window of opportunity** for a peaceful solution according to what is called the middle way concept, not independence but self-government, and (...)”

(h) *Erityisesti minua entisenä neuvottelijana kiinnostaa, pystytäänkö tässä poistamaan niin sanottuja näkymättömiä kaupan esteitä, **invisible barriers to trade**, ja sitten sellaisia, jotka liittyvät tavalla tai toisella turvallisuuspolitiikkaan (...)*

“In particular, as a former negotiator, I am interested in whether it is possible to remove so-called invisible barriers to trade, **invisible barriers to trade**, and then those that are related in one way or another to security policy (...)”

The strategies were marked for each use case, and the results are given in figures 5a and 5b. Overall, Latin use cases were explained less than English use cases, though close reading revealed that rare cases were usually translated. ESTA cases typically had no translations, as expected, since established words and phrases, e.g. *high tech*, were well-known. EFFI and ATTI also usually had no translations, but probably for very different reasons; many use cases what were typically used for the sake of efficiency, such as *win-win* or *de facto*, are also usually well-known, but explaining an expression of attitude would water down the effect on those who understood the use case (example f). Similarly, HUMO in English also had little translations probably for the same reason as ATTI, as explaining a joke takes away the humour. Latin only had five cases of HUMO, three of which came with translations, which reflects the fact that Latin as a little-known language may not work well as humour. SELF and DIST had mostly no translations but were also used a little bit with and as a translation. AUTH had almost half of the cases translated in English, probably because the function included longer quotes; most Latin AUTH cases also had translations. CLAR, EMPH and TRAN were the most evenly distributed functions among the three translation strategies for both English and Latin (except for TRAN which only had one occurrence in Latin).

English					Latin				
Function	AT%	NT%	WT%	total # of use cases	function	AT%	NT%	WT%	total # of use cases
ATTI	0 %	95 %	5 %	57	ATTI	0 %	95 %	5 %	22
AUTH	4 %	49 %	47 %	75	AUTH	10 %	23 %	68 %	40
CLAR	28 %	25 %	47 %	521	CLAR	39 %	28 %	33 %	18
DIST	4 %	84 %	12 %	25	DIST	0 %	80 %	20 %	5
EMPH	20 %	58 %	22 %	144	EMPH	35 %	47 %	18 %	34
ESTA	1 %	96 %	3 %	348	ESTA	0 %	100 %	0 %	3
HUMO	2 %	83 %	14 %	42	HUMO	0 %	40 %	60 %	5
EFFI	2 %	90 %	8 %	641	EFFI	0 %	92 %	7 %	328
SELF	7 %	75 %	18 %	334	SELF	3 %	88 %	9 %	319
TRAN	15 %	28 %	57 %	79	TRAN	100 %	0 %	0 %	1
total	10 %	68 %	22 %	2266	total	12 %	83 %	4 %	775

Figures 5a and 5b: Placement of explanations or translations around use cases of English and Latin. AT = as translation, NT = no translation, WT = with translation (or explanation).

Finally, a brief estimation of phrase occurrence rates per function for both languages was also conducted. Latin was easier to estimate because the searched phrases were predefined: EFFI use cases included dominating cases such as “de minimis” (121), “de facto” (58), and “ad hoc” (48), the other functions mostly had more than five occurrences each. SELF was dominated by “de facto” (256). English was more difficult to estimate because the use cases varied from single words to multi-sentence quotes and had varying orthographies. By browsing an alphabetically ordered list of all use cases most English functions (CLAR, TRAN, EMPH, AUTH, ATTI, HUMO and DIST) seemed to have almost exclusively use cases with only 1-5 occurrences each. SELF use cases had mostly 1-5 occurrences each, with some exceptions of around 6-10 occurrences, but “fair play” alone had 20 occurrences. EFFI had a large portion of use cases with 5-20 occurrences, the most frequent being “no bail-out” with 70 occurrences (all after 2010 and related to the global financial crisis). ESTA was dominated by few extremely frequent use cases, especially “cleantech” (91 occurrences), “high tech” (89), and “know-how” (40).

4.4. Topics around functions

To see the differences between functions more clearly, the summed topic proportions of each function class were subtracted from the average topic proportions. The remainder showed in which contexts the functions were more or less likely to be used. A one-sample t-test was used to identify differences that were statistically significant ($p < 0.05$, degrees of freedom for each function were $n_f - 1$ where n_f was the total number of use cases in each function given in figures 5a and 5b), and the significant results are presented for English in figure 6 and for Latin in figure 7. It should be noted that all the differences listed are still quite small, and since the point of view given by the used topic model is only one of many possible ones, these results should be taken as preliminary and possible trends only. To have more reliable results would require the inspection of results from several different models.

The most noticeable feature of English was that the *general* topic was the least likely context for functions EFFI and ESTA, which are practical functions, while it was the most likely for all the rhetorical functions plus TRAN. The division along the *general* topic is interesting because it is the largest topic in the corpus overall and indicates a meaningful division. It seems that English is used for practical reasons mostly in very specific contexts, while its rhetorical use is more freely spread. Other points of notice were that ESTA function favoured topics *public sector* and *energy*, meaning that these contexts probably have accumulated established English vocabulary without good Finnish counterparts, while EFFI concentrated on topics *legislation* and *public sector*, hinting that perhaps the vocabulary used for its efficiency in these contexts is in a process of becoming established. An example of such concept is “no bail out”, which has been used especially during the financial crisis of Greece around 2010 to refer to an idea that no EU Member State is liable for the debts of other Member States, which in turn comes from the common way to refer to article 125 of the Treaty on the Functioning of the European Union as the “no bailout clause”. The concept re-emerged after 2020, meaning that in the past decade no Finnish equivalent has replaced the English concept.

Latin use was expected in topics that in some way correspond to the known use contexts of science and law. Since science did not emerge as its own topic in the model, the focus here was on the topic *legislation*. Also, only EFFI and SELF had had a large number of occurrences, and only EFFI seemed to favour *legislation* as a context while SELF avoided it. This suggests that legislative speech may actually require Latin to be efficient, and does not go well together with rhetorical uses which could add ambiguity. The lack of instances made it difficult to generalize the results in many cases, but the most significant findings are presented in figure 7.

Practical functions						Rhetorical functions													
EFFI (n=641)		CLAR (n=521)		ESTA (n=348)		TRAN (n=78)		EMPH (n=144)		SELF (n=334)		AUTH (n=75)		ATTI (n=56)		HUMO (n=42)		DIST (n=25)	
topic	diff	topic	diff	topic	diff	topic	diff	topic	diff	topic	diff	topic	diff	topic	diff	topic	diff	topic	diff
legisl	0,9 %	for & sec p	1,0 %	pub sec	3,2 %	GEN	4,1 %	GEN	2,0 %	GEN	3,3 %	GEN	3,1 %	GEN	3,0 %	GEN	3,3 %	GEN	3,9 %
pub sec	0,8 %	legisl	0,9 %	ener	3,1 %	educ	2,9 %	taxation	0,6 %	parl fac	1,3 %	parl fac	2,0 %						
hous	0,4 %	dev coop	0,5 %	comm	1,8 %	legisl	2,5 %												
comm	0,2 %	democ	0,4 %	emplo	1,0 %	admin	1,8 %												
				soc prob	0,4 %														
				traf & tran	0,3 %														
						traf & tran	-0,2 %												
						soc & heal	-0,3 %												
						law prop	-0,3 %												
						agric	-0,4 %												
						soc prob	-0,5 %												
				crime	-0,6 %	taxation	-0,5 %					pensi	-0,3 %						
				dev coop	-0,7 %	crime	-0,6 %					educ	-0,6 %						
				democ	-0,8 %	emplo	-0,8 %					emplo	-0,6 %						
voting	-0,1 %			parl fac	-1,2 %	budget	-0,8 %			comm	-0,7 %	hous	-0,9 %	voting	-0,2 %	agric	-0,3 %	democ	-0,9 %
soc prob	-0,3 %	educ	-0,3 %	for & sec p	-1,6 %	comm	-1,3 %			ener	-0,8 %	comm	-1,0 %	emplo	-0,5 %	law prop	-0,4 %	comm	-1,2 %
emplo	-0,4 %	comm	-0,4 %	legisl	-1,8 %	ener	-1,4 %	for & sec p	-0,9 %	legisl	-1,2 %	ener	-1,3 %	dev coop	-1,0 %	comm	-1,0 %	ener	-1,3 %
GEN	-1,1 %	parl fac	-0,8 %	GEN	-3,5 %	pub sec	-5,7 %	ener	-1,2 %	pub sec	-1,3 %	pub sec	-2,8 %	pub sec	-2,8 %	for & sec	-1,7 %	pub sec	-3,9 %

Figure 6. Topics around functions of English use cases, significant differences from average.

Practical functions				Rhetorical functions							
EFFI (n=328)		CLAR (n=18)		EMPH (n=34)		SELF (n=319)		AUTH (n=40)		ATTI (n=22)	
topic	diff	topic	diff	topic	diff	topic	diff	topic	diff	topic	diff
legisl	2,4 %					pub sec	1,4 %	parl fac	3,9 %	GEN	5,9 %
emplo	0,7 %										
		voting	-0,3 %					region	-0,4 %		
		soc & hel c	-0,4 %					comm	-0,5 %	soc ben	-0,6 %
		dev coop	-0,7 %	tra & tran	-0,4 %			ener	-0,6 %	budg	-0,7 %
		law prop	-0,9 %	crime	-0,6 %			emplo	-0,8 %	ener	-0,9 %
		soc ben	-1,0 %	agric	-0,6 %			agric	-0,8 %	law prop	-0,9 %
voting	-0,1 %	agric	-1,1 %	emplo	-0,8 %	emplo	-0,6 %	tax	-0,9 %	agric	-1,0 %
parl fac	-0,9 %	parl fac	-2,0 %	tax	-0,9 %	legisl	-1,9 %	pub sec	-2,7 %	legisl	-2,5 %

Figure 7. Topics around functions of Latin use cases, significant differences from average. The functions that had less than 10 occurrences (DIST, HUMO, ESTA and TRAN) have been left out.

5. Discussion

The English and Latin use cases examined in this paper demonstrate how the parliamentary speakers utilize linguistic resources to fulfil communicative needs and enhance expression. Certain features of the observed patterns imply weak elite closure while others hint at the opposite: inclusion.

English use was found to have increased since the 1990s, especially the practical uses. The topic/function distribution for English suggested that practical uses of English were more common in specific contexts (avoiding the *general* topic), while rhetorical English use was more common in very general contexts. The contexts where English use was the most prominent were all important areas of societal infrastructure and development: *energy*, *housing*, *education* and *commerce*. Practical Latin use also increased somewhat after the 1990s, probably backed by the prominence of English. The topics that preferred Latin the most were probably connected through juridical discussions regarding them: *foreign and security policy* (55%) and *democracy* (53%). However, Latin use concentrated on two functions, EFFI and SELF, which were strongly divided along the topic *legislation*: efficiency was preferred while self-emphasis was avoided in speeches regarding legislation. Both English and Latin use cases included some extremely frequent ones, but also several rare occurrences, which seemed more likely candidates for elite closure especially when left untranslated or unexplained.

5.1. Roles of English and Latin

Based on the findings, Latin seems to have two main roles in the parliament. First, it appears as a practical means in Latin's traditional use context of law. Second, Latin is used to underline the speaker's own eloquence and to achieve a higher register, but this seems less likely to happen in legislative

contexts. The results reflect a perhaps apparent dichotomy for modern use of Latin: there are very specific contexts where Latin use is almost a necessity, and in most other contexts its use appears to be out of place – in an elitist kind of way. The situation of English is a bit similar, since two main roles can be deduced: practical English use was also more common in specific contexts (avoiding the *general* topic), while rhetorical English use was more common in very general contexts. The roles of English are surprisingly close to Latin's roles, though the division is clearer with English and the rhetorical uses of English do not seem nearly as much elitist as Latin's.

Like mental biases, elite closure can be conscious or subconscious, intentional or unintentional. Though field-specific jargon is usually seen as a classic example of elite closure, the practical use cases of Latin in the context of law seem to arise from a need to speak in precise terms, rather than from an intentional attempt at elite closure. Examples of such use are concepts like *de minimis* (a type of EU aid) or *ne bis in idem* (a criminal law principle under which a person cannot be punished and be subject to several procedures twice for the same facts), both of which would take long to explain without Latin every time and might lose some important connotations if translated. Overall, since “legalese” is so well established and deep-rooted, it is probably very difficult to talk about certain subjects without using at least some Latin terminology. Because of this, even without intentional exclusion, the result may still be an exclusive variety that clearly marks an elite identity. Similar process may be taking place with English in its specific practical use contexts. Especially established words and phrases such as *cleantech* or *hands-free* are hard to circumnavigate, and practical choices matter in the parliament where one is under pressure to speak clearly, effectively, and sometimes even with a time limit. Another possibility is that a speaker can misjudge the fluency levels of their listeners and chooses not to explain, which may lead to unintentional elite closure; the small number of translated English use cases supports this possibility, but more detailed and qualitative studies would be required to verify.

There is also a difference between an elite variety that one cannot fully understand, and an elite variety that one can understand but cannot produce. The former is clearly a stronger form than the latter. Given this premise, both extensive Latin and English uses in the parliament mostly fall into the latter category, and only very few use cases can fulfill both roles. Some distinction between cases can be seen from the occurrence rates of the use cases and the amount of explanation surrounding them. Surprisingly, Latin use cases were explained less than English use cases though Latin is a less known language in Finland. In case of Latin, rare phrases left unexplained would be clear attempts at elite closure, since only few people with good knowledge of Latin could fully understand. However, close reading of the excerpts revealed that most of the rare or infrequent Latin use cases *were* often explained, both practical and rhetorical cases. This is probably a strategy of experienced speakers: the speaker is able to mark themselves with an elite identity while avoiding being accused of being too abstruse. Latin use is not easy to replicate for a person who is not used to hearing it and does not know how and where such phrases should be used, so the elite variety remains exclusive even when the use cases are translated. English, on the other hand, is universally taught and acts as a common *lingua franca* for most Finns (EC, 2024), so despite the large variation in fluency levels it generally takes heavier use and more disregard to explanations to produce English use cases that could not be understood by most listeners in the parliament.

English is commonly linked with an international outlook, and similar tendencies were observed in the rapid rise of English use since the 1990s, as Finland joined the EU, and the world started to become even more interconnected via the internet. Interestingly, the last year of observation (2020) hints at a similar elbow in the future frequency of English use. Also, practical Latin use seems to have increased together with the use of English in the 1990s. Integrating Latin words and phrases into English is quite common especially in higher registers, so it could be that the overall increased use of English in the society has also strengthened the role of Latin. Especially the most common Latin phrases with practical functions, i.e. *de facto* and *ad hoc*, may have become more accepted because people have heard them used so many more times because they hear so much more English. Again, replicating features of elite English in Finnish may be another way to mark an elite variety of Finnish.

For both English and Latin the occurrence rates hint at certain words and phrases being popular and widely used during certain times (e.g. “no bail out” during financial crises) and others being rare and used only very few times (e.g. *si vis pacem, para bellum*; “if you want peace, prepare for war”). Especially the rare cases seem likely candidates to be attempts of elite closure. Some of the mid-rate cases may even be political attempts at redefining a concept (e.g. discussions about the differences

between “peace keeping”, “peace making” and “peace enforcement”). But if a phrase is used often by many different people, it is probably well known enough (in the parliamentary setting at least) that it is not possible to see such use as elite closure. But could it be the opposite? Cases where the speaker gives foreign language (usually English) terms for concepts they already explained in Finnish or Swedish is an interesting phenomenon (see example g earlier). The two most typical interpretations for this type of use are that either the speaker thinks that the foreign language term is better known than the Finnish term and wishes to clarify what is being spoken of (CLAR), or the speaker is condescendingly flaunting their language skills (SELF). The CLAR type is basically inclusive, as it can firstly help the listeners to see the spoken topic from multiple angles by providing a concept from another language, and secondly help the listener to follow if the speaker is to use the term more frequently.

5.2. Future trends

As stated by Taavitsainen and Pahta (2008: 3), "forces of globalisation, growing economic interdependence and the ensuing social and demographic shifts, have had a deep impact on the patterns of language use in the world". The changes in the language use of the Finnish parliament are perhaps subtle but still noticeable even in this brief study. For example, practical uses of Latin increased at the same time with English in the 1990s, which was very likely a by-product of increased English use. The fact that the majority of functions of English use were practical reflects the situation where English skills are really becoming a necessity even in the Finnish society. The use of English in Finland has long been seen both as a threat and as a possibility (e.g. Laitinen et al., 2023; SKL, 2018; Leppänen & Pahta, 2012). The results support the idea of English as a language of international communication, but also as a marker of elite or upper-classness in general, not just in specific contexts. There are similarities between the roles of English and Latin, as both can be used as upper-class and elite identity performance. The present study showed that some of this performance can be effectively exclusive.

The present paper leaves many directions for further studies. The most interesting direction in terms of elite closure would be to examine the used vocabulary more closely and with a discourse analytic perspective. Since English use cases varied from single words to multi-sentence quotations and were not pre-defined, which was the case with Latin, phrase or vocabulary rates for English were not easy to estimate. Some phrases can also be used in various contexts, so a deeper qualitative examination would reveal more about the way English (and Latin) is used to either unite or to differentiate. Another avenue for further examination lies in the question of *who* more specifically are using English or Latin in the Finnish parliament, which will also shed light on the question of for who it will be a necessity to use English (or Latin) in the future.

Acknowledgments

I want to thank my supervisors Veronika Laippala and Kimmo Elo for their guidance. I also thank the Finnish IT Center for Science (CSC) for providing computational resources and my colleagues at TurkuNLP, the University of Turku, and Langnet network for support.

References

- Banner, Nicholas. 2021. “Do Not Read the Latin: Latin as Satanic Signifier in Supernatural Horror Cinema”. *Classical Receptions Journal* 13 (3): 399–415. <https://doi.org/10.1093/crj/claa033>.
- Bedi, Jaskiran. 2019. *English language in India: a dichotomy between economic growth and inclusive growth*. New York: Routledge.

- Coffee, Neil. 2012. "Active Latin: Quo Tendimus?" *Classical World* 105 (2): 255–69. <https://doi.org/10.1353/clw.2012.0007>.
- Engelsing, Eduardo. 2017. "Census Latinus 2009: Goals, Data Collected, Importance, Perspectives". *Classical World* 110 (3): 399–421. <https://doi.org/10.1353/clw.2017.0023>.
- "eTranslation v13.3". 2024. European Commission, Directorate-General for Translation. https://commission.europa.eu/resources-partners/etranslation_en.
- [EC] European Commission. 2012. "Europeans and Their Languages". 386. Special Eurobarometer. <https://op.europa.eu/en/publication-detail/-/publication/f551bd64-8615-4781-9be1-c592217dad83>,.
- [EC] European Commission. 2024. "Special Eurobarometer 540 - Europeans and Their Languages". 540. Special Eurobarometer.
- Furiassi, Cristiano, Virginia Pulcini, Félix Rodríguez González, and European Society for the Study of English, eds. 2012. *The anglicization of European lexis*. Amsterdam; Philadelphia: John Benjamins Pub. Co.
- Gałuska, Ksenia, and Joanna Sycz. 2013. "LATIN MAXIMS AND PHRASES IN THE POLISH, ENGLISH AND FRENCH LEGAL SYSTEMS – THE COMPARATIVE STUDY". *Studies in Logic, Grammar and Rhetoric* 34 (1): 9–26. <https://doi.org/10.2478/slgr-2013-0020>.
- Gumperz, John J. 1982. *Discourse strategies*. *Studies in interactional sociolinguistics* 1. Cambridge [Cambridgeshire]; New York: Cambridge University Press.
- Haugen, Einar. 1950. "The Analysis of Linguistic Borrowing". *Language* 26 (2): 210–31.
- Hickey, Raymond, ed. 2019. *English in multilingual South Africa: the linguistics of contact and change*. *Studies in English language*. New York, NY: Cambridge University Press.
- Hyvönen, Eero, Laura Sinikallio, Petri Leskinen, Senka Drobac, Rafael Leal, Matti La Mela, Jouni Tuominen, Henna Poikkimäki, and Heikki Rantala. 2024. "Publishing and Using Parliamentary Linked Data on the Semantic Web: ParliamentSampo System for Parliament of Finland". *Semantic Web*. <https://www.semantic-web-journal.net/system/files/swj3605.pdf>.
- Jahan, Iffat, and M. Obaidul Hamid. 2019. "English as a Medium of Instruction and the Discursive Construction of Elite Identity". *Journal of Sociolinguistics* 23 (4): 386–408. <https://doi.org/10.1111/josl.12360>.
- Kanerva, Jenna, Filip Ginter, Niko Miekka, Akseli Leino, and Tapio Salakoski. 2018. "Turku Neural Parser Pipeline: An End-to-End System for the CoNLL 2018 Shared Task". In: *Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*. Association for Computational Linguistics.
- Laitinen, Mikko, Sirpa Leppänen, Paula Rautionaho, and Sara Backman. 2023. "Englanti Suomen kansalliskielten rinnalla – Kohti joustavaa monikielisyyttä". *Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja* 2023 (59).
- Leppänen, Sirpa, and Tarja Nikula. 2007. "Diverse Uses of English in Finnish Society: Discourse-Pragmatic Insights into Media, Educational and Business Contexts". *Mult* 26 (4): 333–80. <https://doi.org/10.1515/MULTI.2007.017>.
- Leppänen, Sirpa, and Päivi Pahta. 2012. "Finnish Culture and Language Endangered — Language Ideological Debates on English in the Finnish Press from 1995 to 2007". In: *Dangerous*

- Multilingualism, edited by Jan Blommaert, Sirpa Leppänen, Päivi Pahta, and Tiina Räisänen, 142–75. London: Palgrave Macmillan UK. https://doi.org/10.1057/9781137283566_7.
- Maier, Daniel, Christian Baden, Daniela Stoltenberg, Maya De Vries-Kedem, and Annie Waldherr. 2022. “Machine Translation Vs. Multilingual Dictionaries Assessing Two Strategies for the Topic Modeling of Multilingual Text Collections”. *Communication Methods and Measures* 16 (1): 19–38. <https://doi.org/10.1080/19312458.2021.1955845>.
- Matus-Mendoza, Mariadelaluz. 2002. “The English Lexical Loan: A Class Marker”. *Journal of Hispanic Higher Education* 1 (4): 329–37. <https://doi.org/10.1177/153819202236977>.
- Myers-Scotton, Carol. 1993. “Elite closure as a powerful language strategy: the African case”. *International Journal of the Sociology of Language* 103 (1). <https://doi.org/10.1515/ijsl.1993.103.149>.
- . 1997. *Duelling Languages: Grammatical Structure in Codeswitching*. Oxford: Clarendon Press.
- . 2002. *Contact Linguistics*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198299530.001.0001>.
- ”Parlamenttisampo”. 2023. <https://parlamenttisampo.fi>.
- Peterson, Elizabeth. 2022. “The English Language in Finland: Tool of Modernity or Tool of Coloniality?” In: *Finnishness, Whiteness and Coloniality*, edited by Josephine Hoegaerts, Tuire Liimatainen, Laura Hekanaho, and Elizabeth Peterson, 267–89. Helsinki University Press. <https://doi.org/10.33134/HUP-17-11>.
- Piętka, Radosław. 2016. “A Thrill for Latinists: Latin Language in Contemporary Horror Films”. *Teoksessa Antiquity in Popular Literature and Culture*, edited by Konrad Dominas, Elżbieta Wesołowska, and Bogdan Trocha, 255–66. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Ristilä, Anna, and Kimmo Elo. 2023. “Observing Political and Societal Changes in Finnish Parliamentary Speech Data, 1980–2010, with Topic Modelling”. *Parliaments, Estates and Representation* 43 (2): 149–76. <https://doi.org/10.1080/02606755.2023.2213550>.
- Roelli, Philipp. 2021. *Latin as the Language of Science and Learning*. De Gruyter. <https://doi.org/10.1515/9783110745832>.
- Romaine, Suzanne. 1995. *Bilingualism*. 2nd ed. *Language in society* 13. Oxford, UK ; Cambridge, Mass., USA: Blackwell.
- SKL. 2018. ”Suomi tarvitsee pikaisesti kansallisen kielipoliittisen ohjelman. Suomen kielen lautakunnan kannanotto 26.10.2018”. Oct 26, 2018. https://www.kotus.fi/ohjeet/suomen_kielen_lautakunnan_suosituksia/kannanotot/suomi_tarvitsee_pikaisesti_kansallisen_kielipoliittisen_ohjelman.
- Stahl, Peter. 2023. ”Lingua 2.0.2”. Python. <https://github.com/pemistahl/lingua-py/releases/tag/v2.0.2>.
- Statistics Finland. 2023. 11rm: “Language according to sex by municipality, 1990-2023”. https://pxweb2.stat.fi/PxWeb/pxweb/fi/StatFin/StatFin_vaerak/statfin_vaerak_pxt_11rm.px/.
- Taavitsainen, Irma, and Päivi Pahta. 2003. ”English in Finland: Globalisation, Language Awareness and Questions of Identity”. *English Today* 19 (4): 3–15. <https://doi.org/10.1017/S0266078403004024>.

Appendix 1: Searched Latin phrases

Found:

a priori (7)
 ad hoc (52)
 alea iacta est (1)
 alma mater (1)
 ave caesar, morituri
 te salutant (1)
 bona fide (14)
 carpe diem (1)
 ceteris paribus (5)
 ceterum censeo (8)
 de facto (329)
 de jure (5)
 de minimis (130)
 deus ex machina (2)
 dies irae (1)
 ecce homo (1)
 errare humanum est
 (3)
 ex cathedra (6)
 ex officio (3)
 ex oriente lux (3)
 ex post (3)
 ex tempore (32)
 festina lente (5)
 finis finlandiae (3)
 homo homini lupus
 (1)
 ille faciet (1)
 in absentia (1)
 in dubio pro reo (7)
 in memoriam (1)
 ius sanguinis (1)
 ius soli (1)
 lex in casu (6)
 liberum veto (1)
 mare nostrum (3)
 mea culpa (1)
 mens sana in corpore
 sano (2)
 mutatis mutandis (2)
 navigare necesse est,
 vivere non est
 necesse (2)
 ne bis in idem (25)
 nomen est omen (9)
 non scholae sed vitae
 discimus (1)
 numerus clausus (5)
 o sancta simplicitas
 (2)
 o tempora, o mores
 (8)
 pacta sunt servanda
 (9)
 panem et circenses
 (1)
 pater familias (3)

per se (1)
 persona non grata (1)
 primus inter pares
 (6)
 pro forma (3)
 pro patria (3)
 quis custodiet ipsos
 custodes (1)
 si vis pacem, para
 bellum (2)
 status quo (27)
 sui generis (20)
 tabula rasa (2)
 veni, vidi, vici (1)
 vestigia terrent (8)
 vox populi (4)

Not found:

a posteriori
 ab ovo
 ab urbe condita
 acta est fabula, nunc
 plaudite
 ad acta
 ad maiorem dei
 gloriam
 ad usum delphini
 alterego
 amicus curiae
 amicus plato, sed
 magis amica
 veritas
 amor patriae
 an nescis, mi fili,
 quantilla
 prudentia
 mundus regatur
 annus mirabilis
 argumentum ad
 hominem
 argumentum ad
 misericordiam
 argumentum ad
 nauseam
 ars gratia artis
 ars longa, vita brevis
 auri sacra fames
 casus belli
 causa sui
 caveat emptor
 citius, altius, fortius
 cogito ergo sum
 corpus delicti
 cui bono
 cuius regio, eius
 religio
 cum hoc ergo propter
 hoc

damnatio memoriae
 de dicto
 de lege ferenda
 de lege lata
 de mortuis nil nisi
 bene
 de profundis
 de re
 de se
 delictum
 deus vult
 dicto simpliciter
 dis manibus
 e pluribus unum
 ecclesiola in ecclesia
 et al.
 et nunc et semper
 et tu, brute
 etiamsi omnes, ego
 non
 ex more
 ex nihilo nihil fit
 ex nihilo
 ex post facto
 falsa demonstratio
 non nocet
 filioque
 habeas corpus
 hannibal ad portas
 homo sovieticus
 ignoramus et
 ignorabimus
 in hoc signo vinces
 in loco parentis
 in medias res
 in pectore
 in situ
 in vino veritas
 incertae sedis
 inter alia
 ipso facto
 jure uxoris
 larvatus prodeus
 magnum opus
 memento mori
 modus operandi
 naturalia non sunt
 turpia
 ne plus ultra
 nemo me impune
 lacessit
 nemo propheta in
 patria
 noli me tangere
 nomen nescio
 non compos mentis
 non praevalent
 non sequitur

novus ordo seclorum
 obiter dictum
 omnia mea mecum
 porto
 otium
 passim
 pax vobiscum
 pecunia non olet
 penitenziagite
 per aspera ad astra
 per capsulam
 persona grata
 plenum plenum
 plurale tantum
 post hoc ergo propter
 hoc
 pro bono
 pro hac vice
 puer aeternus
 puer robustus sed
 malitiosus
 pulvis et umbra
 sumus
 quid pro quo
 quo vadis, domine
 quod licet iovi, non
 licet bovi
 requiescat in pace
 salva veritate
 semper fidelis
 sic semper tyrannis
 sic transit gloria
 mundi
 sine qua non
 stare decisis
 sub rosa
 summa cum laude
 terra australis
 terra incognita
 terra nullius
 ultima ratio regum
 ultra vires
 urbi et orbi
 uti possidetis
 vade retro, satana
 vagina dentata
 vanitas vanitatum et
 omnia vanitas
 veto exclusionis
 vicarius filii dei
 vidi aquam