Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Defectiveness, overabundance, and standard language ideologies

Maria Copot

m.copot@surrey.ac.uk

University of Surrey

Mari Aigro

University of Tartu

Andrea D. Sims

The Ohio State University

Virve-Anneli Vihman

University of Tartu

Research Article

Keywords: morphology, defectiveness, overabundance, language ideology, preemption, frequency

Posted Date: June 2nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9760604/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Defectiveness, overabundance, and standard language ideologies

First Author[1,2*], Second Author[2,3†] and Third Author[1,2†]

[1*]Department, Organization, Street, City, 100190, State, Country.
[2]Department, Organization, Street, City, 10587, State, Country.
[3]Department, Organization, Street, City, 610101, State, Country.

*Corresponding author(s). E-mail(s): iauthor@gmail.com;
Contributing authors: iiauthor@gmail.com; iiiauthor@gmail.com;
†These authors contributed equally to this work.

Abstract

Morphological defectiveness and overabundance are opposite phenomena in some sense: defective lexemes have "too few" forms (at least one paradigm cell is left empty), while overabundant lexemes have "too many" (at least one cell has multiple, competing options). Social stigmatisation is implicated as a cause of defectiveness in French verbs (Copot and Sims 2025), raising the question of whether, conversely, social tolerance of variation underpins overabundance, thereby relating the two phenomena. This paper compares defective French verbs to overabundant Estonian verbs using parallel judgement tasks. The results show effects of differences in French and Estonian language ideologies, but Estonian speakers also evaluated overabundant words in surprisingly parallel ways to how French speakers evaluated defective words. A followup experiment on Estonian suggests preemption is at the root of the similar judgements. We consider the implications for the relationship between defectiveness and overabundance.

Keywords: morphology, defectiveness, overabundance, language ideology, preemption, frequency

1 Introduction

Inflectional morphology is often idealised as a deterministic system in which the grammar outputs exactly one wordform for each of a given lexeme's paradigm cells. Yet

deviations from this conceptualisation have long been recognised to exist, including in the form of zero-to-one (defective) and many-to-one (overabundant) form-meaning mappings. For instance, in French, a few dozen verbs lack a number of inflected forms: CLORE 'to close' lacks 1PL and 2PL present indicative, imperative plural, and all of the imperfect and simple past forms (Boyé 2000; Morin 1987, 1995). These missing forms are paradigm gaps. Conversely, in Estonian, an extensive parallel form system is described in the main grammars, especially for nouns (Viht and Habicht 2019). For instance, MAJA 'house' has parallel forms available in nine of its twenty-eight cells (e.g. *maju˜ maja-sid* 'house.PAR.PL'); though less widespread, overabundance also occurs in verb inflection, as attested by the active past participle of MÕTLEMA 'think', which can be expressed as either *mõel-nud* or *mõtel-nud* 'think-PST.PTCP'. In recent years, defectiveness and overabundance have garnered interest for what they can reveal about the structure of grammatical systems (e.g. Sims 2015; Thornton 2011). In this paper we explore a complementary line of investigation, testing a hypothesis that social evaluation and language ideology can incline a language toward defectiveness or overabundance.

We take as a starting point Copot and Sims (2025), who argue that in French, social stigmatisation caused by prescriptive discourses works in tandem with grammar-internal factors to delineate the class of defective verbs. Here we pose the question of whether overabundance can derive from the converse social situation. In other words, does social *tolerance* of inflectional variation work in tandem with grammar-internal factors to delineate the class of *overabundant* lexemes? France has a strongly conservative, top-down and anti-variationist language planning culture, while Estonia's language planning culture, though conservative and top-down in some respects – for instance, in rejecting Anglicisms – is generally accepting of morphological variation. This offers the opportunity to examine whether standard language ideologies underpin defectiveness in French and overabundance in Estonian in parallel ways.

We implement parallel experimental studies of defective verbs in French and overabundant verbs in Estonian. We compare speakers' judgements of defective/overabundant forms in each language to Anglicisms, which are socially stigmatised in both languages, and to anchoring controls. If the presence of defectiveness versus overabundance in a language is underpinned by how accepting each language's normative standard is of variation, we expect defective verbs in French – but not overabundant ones in Estonian – to show hallmarks of stigmatisation.

We find evidence that language ideologies do affect how language users react to overabundant and defective forms. However, some of the predictions about the specific manifestation of such influences were not upheld. While the results in French largely follow predictions, some Estonian results are unexpected. Estonian speakers evaluated overabundant verbs – specifically, the less frequent of competing overabundant forms – with high variability, in ways that are surprisingly parallel to how French speakers evaluated words that are candidates to fill paradigm gaps. In light of the fact that overabundance is much more restricted in usage compared to its theoretical availability in the standard language (Aigro and Vihman 2023, 2024), we interpret this finding as a preemption effect, also demonstrated by a follow-up experiment. This brings us to the conclusion that defectiveness and overabundance may share unexpected similarities

in the extent to which people find them acceptable. This finding is in tension with narratives about defectiveness and overabundance as opposite kinds of inflectional phenomena.

2 Standard language ideologies

In this section, we lay some groundwork for our theorizing that a difference in the structural properties of French and Estonian morphology – namely, exhibiting defectiveness or overabundance – derives from the ideologies imbued in the language cultures.

2.1 Different flavours of standard language ideology

Starting with the Herderian nationalist movements of 19th century Europe, structural uniformity has generally been a central goal in language standardisation processes, since an ideology of linguistic uniformity helps to imagine desired ethnonational unity (in the sense of Anderson 1983). Nonetheless, standard languages exist in particular times and places, and are constructed and reconstructed by everyday practices. Treating structural uniformity as an *inherent* goal of standard language ideology (Milroy and Milroy 2012) thus risks an overly homogenising view of standardisation. In particular, from such a perspective it may seem incongruous, if not outright contradictory, for a standard language to codify variation (Coupland and Kristiansen 2011), yet this occurs, for example in the form of tolerance of overabundance in Estonian. To understand why, it is necessary to examine standard languages' ideological underpinnings.

To begin with, standard languages are socially effective only if speakers believe them to be authoritative and legitimate by their nature. Yet as linguistic anthropologists have long observed (e.g. Gal and Woolard 2001), linguistic authority and legitimacy are not inherent properties, but rather, are culturally constructed ones. Relevantly here, the beliefs from which standard languages derive their authority and legitimacy can vary.

Broadly speaking, two types of ideology underpin the authority and legitimacy of standard languages in modern Western societies: anonymity and authenticity (Woolard 2021). An ideology of *anonymity*, rooted in the Enlightenment ideal of a disinterested, universal truth, presents a standard language as a "voice from nowhere". The standard can be identified as authoritative and legitimate exactly because it is purportedly an objective fact of the natural world, identified with nobody in particular (Gal 2001). An ideology of *authenticity*, on the other hand, is rooted in the Romantic ideal according to which particular groups within society are stereotyped as "truer" examples of the national type. It presents a standard language as the voice of those speakers and projects their essentialised traits onto the nation as a whole. Coupland (2003) argues for two types of authenticity: one which flows from language practices of elites (*establishment authenticity*), for example via iconisation of the language as having a venerable literary tradition, and one that flows from valorisation of the language as it is perceived to be used by the "everyday person" (*vernacular authenticity*).

As we lay out in §2.2, a foundational observation for the experimental work that follows is that French and Estonian reflect different underpinning ideologies and different contexts for standardisation, which inform their approaches to structural variation.

2.2 Comparing French and Estonian

Comparison across French and Estonian is interesting exactly because these two languages have been subject to many of the *same* social forces. French standardisation began already in the 17th century,[1] but French ultimately participated in and served as one of the models for 19th-century language standardisation movements, when Estonian codification gained momentum. The two languages thus offer a good basis for comparison. Nevertheless, their trajectories differ in important respects, and it is these differences that we find compelling for the relationship between language ideology and structural properties of the languages.

2.2.1 French standard language ideology

French language planning is top-down in nature, with the French Academy (*Académie Française*) making decisions about the standard language, publishing the official dictionary, and acting as the ultimate language authority. The Academy, whose members typically do not include linguists, represents a highly conservative approach to language planning. Hence, language planning includes an emphasis on the implementation of a centralised, standardised language, taught in schools and rigidly defended by the official authority. Unsurprisingly, this has resulted in a high level of metalinguistic awareness in the population as a whole. Woll (2019) found that a third of French speakers in her sample scored at ceiling on a metalinguistic knowledge questionnaire.

The legitimacy of this standard rests in part on an ideology of establishment authenticity, of which the long and valorised history of high French literature is indexical. It also reflects a belief in the inherent clarity and logic of the French language itself. This authority from anonymity serves to reify the standard language (which is then equated with the language as a whole) as an independent entity. Establishment authenticity and authority from anonymity thus both ideologically disconnect the standard language from its speakers. They also present the language as immutable: high literature is fixed in written form, while logic is unchanging by its nature. Correspondingly, the Academy has consistently expressed a low tolerance for structural variation. This applies to local languages and varieties, but also to variation in lexical and especially morphosyntactic usage. Anglicisms are categorised either as useful loans or as vogue words, the latter of which ought to be avoided. The Academy famously proposes French-based neologisms in place of new Anglicisms (such as *courriel* for 'email').

As with many European languages (Auer 2005; Kristiansen 2025), use of Standard French expanded in the 20th century both to spoken contexts and to broader populations within society. This democratisation of usage positioned the standard variety as a competitor to vernacular varieties, expanding it into contexts previously reserved for vernaculars and leading for the first time to native speakers of the standard.

[1] Neustupný (2006) identifies French standard language development as beginning in the Premodern period, defined by replacement of exoglossic written norms with endoglossic ones. In contrast, the nationalist movements that spurred language standardisation across Europe define the Early Modern period.

At the same time, democratisation of standard language usage positions speakers as arbiters of the standard through their own individual alignment (or lack thereof) with the standard language ideology, and through their corresponding language practices (Coupland and Kristiansen 2011, 28).[2] Previous research shows that French speakers vary in how much they individually agree with discourses reflecting standard language ideology (Copot and Sims 2025; Gadet 2005).

French thus presents a situation in which the standard language ideology promotes strict structural uniformity. Speakers, in turn, can either accept or reject the authority and legitimacy of the standard – and by extension, its structural properties – but the language's relative lack of (non-stigmatised) variation, combined with an ideology that explicitly disconnects the standard from speakers, leaves little option to valorise different ways of speaking within the boundaries of the standard language itself.

In the following section we will see that Estonian presents a more complicated situation.

2.2.2 Estonian standard language ideology

Estonia also has a central language authority, the Estonian Language Institute (*Eesti keele instituut*), but due to its position and approach, it plays a different role to the French Academy. Estonian language planning has its roots in the European nationalist movements of the late 19th century. In the Soviet period (from World War II to 1991), language planning was seen as a top-down tool for resisting Russian influence on Estonian. The Dictionary of Correct Estonian Usage (*Eesti õigekeelsuse sõnaraamat*, or ÕS) is considered to document the language standard. The Estonian Language Institute, mostly comprising linguists, is responsible for publishing and updating it with decisions regarding forms included in the standard language.

Unlike with French, the legitimacy of this standard rests on vernacular authenticity. Starting even in the first half of the 20th century, some scholars promoted the dialects as source material for the standard language (Saareste 1933), and since the 1970s, language planners have explicitly attempted to bring the standard language closer to actual language use (Lindström et al. 2023), thereby validating users' linguistic intuitions and in turn gaining authenticity. In a survey of language attitudes, Algvere (2025) finds that many Estonian speakers ideologically position the standard language under the umbrella of the mother tongue. In Estonian, then, the standard language has long been ideologically connected to how speakers use the language.

This has included increasing tolerance for structural variation (Lindström et al. 2023, 23). Since the 1970s, the decisions regarding morphosyntax have mostly involved the piecemeal acceptance of overabundance for a growing number of wordforms (for a list of such decisions for the period 2008-2025, see Emakeeleselts 2025). These decisions typically reflect already existing variation in usage. All in all, the Estonian language standard currently includes a great deal of parallel morphological forms (Erelt 2002; Lindström et al. 2023; Raag 1998, 2008). This is strikingly different from the approach of the French Academy.

[2] A hallmark of late modernity and postmodernity has been push back against the experiment in shared identity that language standardisation embodies (Ayres-Bennett 2021; Coupland and Kristiansen 2011). As a social movement, postmodern humanism promotes a more decentralized and individualistic understanding of humanity.

This inclusive approach to variation can be seen as a response to democratisation. As with French, Estonian underwent democratisation of use of the standard language in the 20th century, leading to native speakers of the standard. Whereas in France the dialects all but disappeared, in Estonia, this led to an extended period of diaglossia, a standard-dialect continuum. The development of diaglossia involves a shortening of the distance between the vernaculars and the standard language (Auer 2005).[3] Eventually, the remaining variation "... can be seen as *variation within the standard language*" (Kristiansen 2025, 6, emphasis ours). In other words, diaglossia can change the boundaries of the standard itself.[4]

At the same time, the situation is more complicated when it comes to lexical variation. The intense influence of English on the lexicon is easily noticed. The Estonian Language Institute has not taken a strong position against the use of Anglicisms, but English borrowings are often remarked upon negatively in the public media. In this space, they have been ideologically positioned as a threat to Estonian since the 1990s, replacing the threat from Russian (Lindström et al. 2023, 29). Intolerance of Anglicisms is compatible with an ideology of vernacular authenticity, since English is an external influence and thus one that can be positioned in contrast to the idea of a "true" (read: untainted) Estonian language. This demonstrates that attitudes towards variation are generally more permissive for standard language variation that stems from Estonian dialects, and less so for influences from outside of the language, and especially for Anglicisms.

Finally, with regard to everyday Estonian, destandardisation can be observed starting in the 2010s (Lindström et al. 2023, 29). In *destandardisation*, "...the established standard language loses its position as the one and only 'best language"' (Coupland and Kristiansen 2011, 28). In other words, the authority of the standard language in society is weakened and the standard is to some degree replaced by prestigious spoken norms that are perceived to have more legitimacy.

Thus, whereas French speakers are positioned to accept or reject the standard language within a classic prescriptive binary, Estonian exhibits a more complicated dynamic. Estonian language planning in recent years exhibits an accepting attitude towards variation, at least when motivated by principles of Estonian grammar. However, a shift toward spoken norms and a more general weakening of the position of the standard itself (destandardisation) have to some degree eroded the standard-vernacular dichotomy. As a result, Estonian speakers have greater agency to valorise different ways of speaking.

[3] In Auer's (2005) typology of standard-vernacular relations in Europe, the diaglossic type that includes Estonian is the most common (Auer 2005, 31). French belongs to the second most frequent type (Auer 2005, 27): languages in which the dialects have disappeared entirely.

[4] See Elspaß (2021) for discussion of how and why vernacular variants succeed in becoming established as properties of a standard variety, what he calls standardisation "from below".

3 A hypothesis about the relationship between defectiveness, overabundance, and language ideology

In this section we lay out our hypothesis and predictions about how these issues might relate to the structural properties of each language's grammar, specifically, defectiveness in French and overabundance in Estonian. Our expectations are informed by Copot and Sims's (2025) study of French defectiveness, so we start the discussion there.

Copot and Sims (2025) examine the effect of prescriptivism and language culture on French verbs listed in dictionaries as defective. They propose that these verbs' paradigm gaps "...are not so much examples of ineffable ungrammaticality as they are examples of social stigmatisation" (Copot and Sims 2025, 1293) and implement a judgement experiment in which French speakers were presented with verb forms that are candidates to fill defective paradigm cells. The speakers were asked to rate these verbs (presented in a sentence context) in one of two ways: either normatively (i.e., according to its "correctness") or according to possibility (i.e., whether it would be normal to hear the form used in conversation). These judgements were compared against two types of baseline items: verbs with subject-verb agreement errors (neither correct nor possible) and slang words (not prescriptively correct but possible). The logic of the experiment was that if French verbal defectiveness reflects unresolvable ungrammaticality, judgements about defective words should pattern with judgements of items with subject-verb agreement errors. However, if defectiveness instead reflects social stigmatisation, then judgements of defective words should pattern with judgements of slang words. Participants also rated their agreement with prescriptive discourses in circulation in French society.

Copot and Sims report three major results. First, there was substantial variation in the extent to which participants agreed with prescriptive discourses. Second, participants who reported higher rates of agreement with these discourses rated defective items lower, especially in the normative task. Third, for the defective items, higher frequency lexemes were rated lower. This was also true, to a lesser extent, for slang items. Copot and Sims interpret higher frequency as a likely correlate to higher metalinguistic awareness of the items' prescriptive stigmatisation. Considering the results collectively, they conclude that anti-variationist pressures in French language culture play a role in delimiting the set of defective words by causing speakers who align themselves with standard language ideologies to exhibit a dispreference for forms that are (or may be) prescriptively taboo, especially in normatively-oriented situations, where the risk of social sanction is likely to be high.

We take this idea as a starting point for thinking about the possible relationship between standard language ideologies on the one hand, and defectiveness and overabundance on the other. If stigmatisation arising from an anti-variationist language planning culture is a partial cause of defectiveness in French, does a language planning culture that is accepting of variation play a causal role in overabundance? In both languages, we expect speakers' judgements about normatively overabundant or

defective words to reflect the extent to which they individually orient to language ideologies in circulation in their society's culture. At the same time, since the ideologies of the two languages differ and speakers are positioned differently relative to them, we expect different manifestations of this orientation.

Following Copot and Sims, we expect that when French speakers are asked to judge a "gap-filling" form (i.e. one filling an empty cell) of a normatively defective verb, their responses will reflect how much they view the standard language as authoritative. In a standard language culture that projects the message that there is only one "right" way to say something, speakers who accept that ideology may reject gap-filling forms if they know the verb to be prescriptively defective, or even simply as a kind of risk aversion, if they are uncertain about the verb's form and/or its status as defective. Alternatively, speakers who reject the standard language ideology are predicted to be more accepting of gap-filling forms.

Estonian offers more possible relationships between the standard language ideology and overabundance. Since overabundance is a sanctioned part of the standard language, Estonian speakers who accept the standard language ideology (especially the underlying idea of vernacular authenticity) are expected to be accepting of this morphosyntactic variation. However, as a result of destandardisation, Estonian speakers are more empowered to define for themselves what counts as the "best" language. We expect these speakers to draw on their own linguistic experience, and to evaluate overabundant forms accordingly.

Thus, the languages' differing approaches to variation are expected to produce contrasting effects, mediated by how individual speakers orient to standard language ideology.

4 Properties of defective French verbs and overabundant Estonian verbs

We now turn to a test of these predictions. We start with a summary of relevant structural facts about French defectiveness and Estonian overabundance.

4.1 Defectiveness in French verbs

The structural properties of defective French verbs have been documented in Morin (1987, 1995), Boyé (2000), Boyé and Cabredo Hofherr (2010), and Bach and Esher (2015).[5] The key observation is that the defectiveness of the set of French verbs is rooted in the morphological organisation of verbal stems. The affected cells do not form a natural morphosyntactic class but are expected to have forms built on the same stem, based on Bonami and Boyé's (2002) analysis of stem allomorphy in French verbs. For the defective verbs, the generalisation is thus that if one inflected form is lacking, other forms sharing the same stem or implicated by that stem are absent too.

In some cases this is because the verbs lack any clear analogical model or are pulled between two models, while in other cases the verbs have one clearly expected stem form. For example, Boyé and Cabredo Hofherr (2010) argue that the existing forms

[5] We set aside verbs that are defective for clearly semantic or pragmatic reasons, focusing only on ones in which defectiveness is connected to identifiable structural properties.

Indicative			
Present	**Imperfect**	**Simple past**	**Simple Future**
je cl**os**	-	-	je cl**orai**
tu cl**os**	-	-	tu cl**oras**
il cl**ôt**	-	-	il cl**ora**
-	-	-	nous cl**orons**
-	-	-	vous cl**orez**
ils cl**osent**	-	-	ils cl**oront**
Subjunctive		**Conditional**	**Imperative**
Present	**Imperfect**	**Present**	**Present**
que je cl**ose**	-	je cl**orais**	cl**os**
que tu cl**oses**	-	tu cl**orais**	-
qu'il cl**ose**	-	il cl**orait**	-
que nous cl**osions**	-	nous cl**orions**	
que vous cl**osiez**	-	vous cl**oriez**	
qu'ils cl**osent**	-	ils cl**oraient**	

Table 1 The conjugation of finite non-periphrastic forms of the defective verb CLORE 'to close'. Dashed cells are deemed defective.

of CLORE 'close', shown in Table 1, strongly suggest a particular form to fill most defective cells (e.g., *closez* for 3PL indicative present), but the corresponding stem is lexicalised as being absent. However, they classify the simple past cells of CLORE as ones for which no sufficiently plausible stem exists. As for verbs pulled between two attractors in conjugational space, we analyse verbs like SURFAIRE 'overdo' in the indicative present 2PL as this type: the verb is formed by prefixing a highly irregular verb, FAIRE 'do'. The indicative present 2PL form of the base verb is *faites*. However, on the basis of implicative relationships within the French verbal system, the more likely indicative present 2PL form for this class of verbs would be *faisez*. While this form is not listed in dictionaries and is considered an error by the *Académie française*, in practice it is used.

4.2 Overabundance in Estonian verbs

Potential overabundance – the availability of parallel forms – is rampant in Estonian inflection. While the phenomenon is known to affect the majority of nominals (Viht and Habicht 2019; Aigro and Vihman 2023), it also occurs in the verbal domain (Table 2), where it is largely unexplored.

Verbal inflection in Estonian is organized into fourteen inflection classes, based on formative types and gradation (Viht and Habicht 2019; see also Blevins 2007). Gradation is the pattern underlying stem variation in principal forms – the *ma*-infinitive, *da*-infinitive and 3SG. If all three use the same stem (e.g. *ela-ma*, *ela-da*, *ela-b* 'live'), the verb has no gradation. If the 3SG stem entails phonological loss (assimilation, substitution, elision), the verb has weakening gradation (e.g. *õppi-ma*, *õppi-da*, ***õpi-b*** 'learn'). Loss in the *da*-infinitive stem is known as strengthening gradation (e.g. *värba-ma*, ***värva-ta***, *värba-b* 'recruit'). For instance, one overabundant form in Table 2 (indicative present, 3SG) renders a verb with weakening gradation (*sulge-ma*, *sulge-da*, *sule-b* 'close'), while the other renders a gradationless verb paradigm (*sulge-b*).

Indicative		Conditional	Imperative
Present	**Imperfect**	**Present**	**Present**
ma sule**n**∼sulge**n**	ma sulge**sin**	ma sule**ksin**∼sulge**ksin**	sule∼sulge
sa sule**d**∼sulge**d**	sa sulge**sid**	sa sule**ksid**∼sulge**ksid**	sulge**gu**
ta sule**b**∼sulge**b**	ta sulge**s**	ta sule**ks**∼sulge**ks**	sule**me**∼sulge**gem**
me sule**me**∼sulge**me**	me sulge**sime**	me	sulge**ge**
te sule**te**∼sulge**te**	te sulge**site**	sule**ksime**∼sulge**ksime**	sulge**gu**
nad sule**vad**∼sulge**vad**	nad sulge**sid**	te sule**ksite**∼sulge**ksite**	
		nad sule**ksid**∼sulge**ksid**	

Table 2 The conjugation of the overabundant verb SULGEMA 'to close' - principal forms: *sulgema, sulgeda, suleb*∼*sulgeb*. Cells with two forms separated by tilde are overabundant.

The largest verb inflection class is also one of the only classes without gradation: this includes 62% of all verbs (Viht and Habicht 2019, 180). As an overarching pattern in verb paradigms, verbs from classes *with* gradation also have forms available which follow this dominant, gradationless class. For instance, MÕTLEMA 'think' has weakening gradation (*mõtle-ma*, *mõel-da*, *mõtle-b*) but its parallel forms follow a gradationless pattern (*mõtle-ma*, *mõtel-da*, *mõtle-b*). In these cases, parallel forms result from dual inflection class membership of a lexeme. Overabundance, however, can also result from parallel forms within a single inflection class. Verbs with consonant-final stems, for instance, can include class-internal parallel forms with vowel-final stems (*kõnele-tud* ∼ *kõnel-dud* 'speak.PST.PTCP'), according to the standard treatment (Viht and Habicht 2019).

The main reference dictionaries (e.g. the online *Sõnaveeb*) present a range of options pertaining to morphological paradigms, outlining a system more extensive than the forms actually used by speakers (Aigro and Vihman 2023). For instance, in a dataset of 5,000 most frequent nouns and 5,000 most frequent verbs based on *Sõnaveeb*[6] (Beniamine et al. 2024), 89% of nominals have parallel forms *available* in at least seven of the twenty-eight nominal cells, though not all of these parallel forms are attested in actual usage. Overabundance in verbs constitutes a narrower system: according to Viht and Habicht (2019, 182), the standard includes 106 verbs with simultaneous membership in several inflection classes, creating overabundant paradigms for them.

5 Methodology

The following sections contain two empirical studies aimed at understanding the link between defectiveness and overabundance, and the unique factors that contribute to the manifestation of each phenomenon. Experiment 1 (§6) compares responses to defective items in French with responses to overabundant items in Estonian, where items are maximally parallel, with the goals of investigating the relationship of each item type to language ideology. To ensure maximal comparability between the two languages, we use the less frequent member of each overabundant doublet in Estonian. Experiment 2 (§8) is a repetition of the Estonian portion of the experiment, this

[6] https://sonaveeb.ee/ - accessed 20 March 2026

time including both overabundant forms for each item instead of just the less frequent one, providing crucial context to interpret the results of the first experiment. The task conditions, language attitudes questionnaire, and procedure are shared across Experiments 1 and 2. In this section we therefore discuss these aspects of both studies. Items and participants are discussed for each experiment separately.

5.1 Tasks

In both experiments, participants heard a recording of the sentence, the verb of which was repeated again after the sentence. They could listen to the recording as many times as they wanted. After the recording was played once, a slider with a continuous scale (underlyingly 100 points, unmarked) would appear on the computer screen under the recording. In the instructions, participants were asked to provide a judgement for the repeated verb by placing a marker somewhere along the scale. For each trial, the marker was initialized in the middle of the scale.

Participants were assigned to one of two task conditions. Half of the participants within each language were tasked with providing a normative judgment of the stimuli: they were instructed to judge the verb form based on whether a teacher would mark the usage of the word as wrong, or whether they would expect not to find this usage in a grammar book or a dictionary. Along with emphasizing the importance of standardized norms, this task condition also carried implications about the formality of the context of use, since formal contexts are where the use of language is more prescriptively codified. For this task, the extremes of the scale were labelled (in the target language) as "completely incorrect" and "completely correct".

The other half of participants were asked to judge sentences in a possibility condition, which aimed to create contexts that would minimize the application of standard language filters on the judgement, by instead shifting the focus to everyday informal language use. Participants were asked to rely on their intuitions about language use, such as whether they might hear the usage in question in a conversation between friends at a bar or between students hanging out after school. For this task, the extremes of the scale were labelled (in the target language) as "completely misplaced" and "not at all misplaced".

Instructions at the beginning of the experiment outlined how the participant should judge the underlined words by discussing the idea of normativity or possibility explicitly, as well as by giving heuristics and examples, depending on the task condition. The full task instructions are available on OSF; see the data availability statement at the end of this paper. Each participant saw only one task condition, as described in §5.3.

5.2 Language attitudes questionnaire

In order to assess whether speakers' evaluations of defective or overabundant items are connected to their beliefs about the standard language, participants responded to a language attitudes questionnaire. The nine questions asked participants to rate their agreement with statements reflecting a range of specific beliefs under the umbrella of standard language ideology: anti-variationism and the idea of rules defining the "right" way to speak, negative attitudes about language change, and beliefs about the

superiority of written over spoken language, the authority of language institutions, and the aesthetic beauty of the language used by important literary figures. A sample of these statements:

- Young people today speak impoverished Estonian/French.
- The mission of standardizing the Estonian/French language carried out by the Estonian Language Institute/French Academy is commendable.
- Proper, rule-following standard language allows speakers to express themselves more clearly than the Estonian/French language used in online chat apps.

Participants rated their agreement with these statement on a 5-point Likert scale (0-4). The scores were summed to obtain a score for each participant, ranging from 0 to 36. The full set of questions can be found on OSF.

As already noted, French and Estonian language ideologies operate in somewhat different social contexts. We chose to use parallel questionnaires in the two languages in order to maximize comparison across the languages, but this difference raises the question of whether the statements measure the same underlying construct(s) in each language. We therefore looked at correlations among the responses to the statements; this suggested, unsurprisingly, that what the questionnaire measures is more complex for Estonian than for French.

For French, responses to the statements suggest that the questionnaire can be interpreted as measuring a single underlying construct – adherence to standard language ideology. A higher score on the questionnaire indicates greater alignment with this suite of ideologies, all of which are aligned.

For Estonian, responses to the statements were less strongly correlated. While the responses suggested that high language attitude scores are associated with linguistic conservatism in general, there seemed to be (at least) two distinct underlying constructs: one reflecting negative attitudes towards language change and another positively oriented towards the notion of the standard language and role of language institutions. The pattern of responses to the questionnaire suggests that in the Estonian social context these two dimensions of language ideology are separable, whereas for French they are not. This multi-dimensionality is an interesting issue for future investigation, but to facilitate comparison across the languages, we move forward with a single score for each participant, with high scores interpreted as conveying conservative language attitudes.

5.3 Procedure

The experiments were deployed on PcIBEX (Schwarz and Zehr 2018) and participants were recruited through the Prolific platform.[7] The experiments were administered entirely in the target language. Participants were first presented with a consent form, followed by the language attitudes questionnaire. Each participant was then randomly assigned to one of the two task conditions (normativity vs possibility). Before starting the main task, participants were presented with three practice items where the key

[7] https://www.prolific.com/

item was a noun, to familiarise them with the use of the scale. In the main part of the experiment, each participant heard test items in randomised order.

In order to exclude judgments for lexemes with which participants were not familiar, after the assessment of the main stimuli, each participant was presented with a lexeme check that was used to identify both words that participants were unfamiliar with, and also whether participants were paying sufficient attention during the experiment tasks. Specifically, participants were presented with a list of verbs in their citation forms (in both languages, the infinitive). In addition to verbs used in the experiment, the list also included pseudoverbs. Participants were asked to select only the verbs for which they knew the meaning. All test items that they did not report knowing were excluded from analysis. Additionally, participants who reported knowing more than twenty percent of the pseudoverbs were deemed to have failed the attention check. All responses from these participants were excluded.

At the end of the experiment, participants were asked for demographic information, were offered the chance to give feedback on the experiment, and were redirected to a debriefing about the goals of the research.

6 Experiment 1

The first experiment was designed to test the extent to which defectiveness in French and overabundance in Estonian are sensitive to standard language ideology, in particular, its acceptance or rejection of structural variation. We measure language ideology on different levels: first, at the level of national language culture, with Estonian standardisation processes encoding a more pro-variationist stance than French; secondly, at the level of the individual, by collecting language attitude scores for each participant; and finally, as a contextual difference under the guise of the two task conditions. The purpose of this experiment was to provide maximally comparable data on the behaviour of defective and overabundant items. Median experiment completion time was thirteen minutes for French participants and fourteen minutes for Estonian participants. Participants were compensated €2.50.

6.1 Participants

A total of 160 participants were recruited on the Prolific platform, 80 for French and 80 for Estonian. All participants reported being native speakers of one of the two languages, having grown up speaking French or Estonian as their only language in most day-to-day situations, and residing at the time of the experiment in France or Estonia, respectively. Following the lexeme knowledge attention check described in §5.3, 19 participants were excluded (8 Estonian, 11 French). For the remaining participants, the median age was 26 in French (range: 22-30) and 33 in Estonian (range 23-47). The gender breakdown for French was 22 females, 42 males and 4 nonbinary people, and for Estonian 46 females, 25 males, and 1 nonbinary person. Five people reported an elementary diploma as their highest educational attainment (1 French, 4 Estonian), 30 had a high school diploma (8 French, 30 Estonian), 57 had an undergraduate diploma (30 French, 27 Estonian), and 49 had a masters' diploma or above (37 French, 12 Estonian).

The individual language attitudes questionnaire described in §5.2 revealed that French participants' median score was 21 out of 36 (range: 0–33; Quartile 1: 15, Quartile 3: 25). Estonian participants scored a median result of 22 (range: 6–33, Quartile 1: 19; Quartile 3: 26).

6.2 Items

6.2.1 Item conditions

Participants for each language were shown words belonging to four different item conditions, as follows.

Defectiveness/overabundance: French participants were presented with conventionally defective verbs and Estonian speakers with conventionally overabundant verbs.

Subject agreement errors: In both languages, participants were presented with items in which the verb agreement marker did not match that required by the subject of the sentence. This condition provides a yardstick for the acceptability of word forms that are not licensed by the grammar. Items in this condition were expected to be rated at floor in both task conditions for both languages.

Invariant word forms: In both languages, participants were presented with items for which only one verb form per paradigm cell is considered acceptable according to the standard language's norms. The acceptability of this class of items provides an upper bound for participants' ratings when the word of interest does not have any remarkable qualities. Items in this condition were expected to be rated at ceiling in both task conditions for both languages.

Anglicisms: In both languages, participants were presented with items in which the verb was a recent English borrowing. Anglicisms are stigmatised by the prescriptive culture of both French and Estonian and provide a standard for how much participants in each language penalize stigmatised words.

Within each language, nine items were presented from each condition. The relatively low number of items per condition is a consequence of the low number of viable verbs both in French and in Estonian. Items were matched for lexeme frequency between conditions, and across languages, to the extent possible.

6.2.2 French items

Items in the defective condition were obtained by extracting defective lexeme+cell combinations from Flexique (Bonami et al. 2014), in which word forms are marked as defective if they are cited as such by two major dictionaries. Lexeme+cell combinations that were marked as defective were filled with the most likely form, in consultation with a French morphologist and native speaker. These were the forms used in the

experiment. French has about fifty verbs that are reported to be defective. However, we excluded all lexemes which were listed as archaic or heavily restricted in their usage, either by register or stylistically (e.g. words pertinent to the law or bureaucracy). (See Copot and Sims 2025 for further discussion.) This filter left us with only nine lexemes, all of which were used in the experiment.

Despite this attempt to remove stylistically restricted words, most verbs categorized as defective by French grammars and dictionaries are associated with formal registers. The set includes no words characteristic of informal registers alone. To alleviate the risk of register interference in our study, we performed a norming study to test for any remaining formality bias compared to the other items used in the study. The thirty-six verbs used in the study, including nine defective verbs, were shown via Prolific to a pool of forty native French speakers who did not participate in the main experiment. The lexemes were presented in isolation in their infinitive form, which is not defective for any of the lexemes. The participants first had to indicate whether they were familiar with the lexeme or not, and if they were familiar with it, they were asked to rate how formal the lexeme was, on a Likert scale from 1 to 5, anchored on "very informal" on the left and "very formal" on the right. We transformed the formality judgements into z-scores calculated within participant in order to account for participants using the scale in different ways. We investigated whether there was any leftover formality bias in the defective lexemes we used in the experiment.

condition	mean	median
defective	0.40	0.67
agreement error	0.33	0.24
prescriptively invariant	0.30	0.28
Anglicism	-1.15	-1.25

Table 3 Summary statistics for formality judgements by item condition (z-scores)

Table 3 shows that even after we excluded defective lexemes listed in dictionaries as being associated with formal register, remaining defective lexemes were rated as more formal than the other three verb groups. Lexemes used for the agreement error condition and lexemes that are invariant were comparable in their formality, while Anglicisms were rated as informal, as expected. The inherent association between French defective verbs and formality is of intrinsic interest. However, it also presented a problem for the experiment design. Standard language norms are strongly, perhaps even inherently, tied to formality, since what counts as "good" language tends to be defined in terms of formal, written norms (Curzan 2014). Among the two task conditions in our experiment, the normativity condition was designed to invoke such contexts, and thus likely formality. We return to this issue in discussion of the results of Experiment 1.

French ungrammatical items featured errors in subject-verb agreement. French Anglicisms were obtained based on various lists of common English borrowings into French. The French invariant items were matched to defective items according to how

much uncertainty (in the information-theoretic sense) was associated with the inflected form of the lexeme, as reported by Qumin (Beniamine 2018). They were obtained by finding the average local entropy for French defective words and finding word forms that were non-defective in the same cell that had similar average local entropy. The input lexicon was Vlexique 2.0 (Beniamine et al. 2024).

6.2.3 Estonian items

From the dataset of inflectional paradigms of the most frequent Estonian verbal lexemes (Beniamine et al. 2024, N = 6,555), we extracted 145 lexemes with overabundant cells. We removed verbs with several distinct homonymous meanings in reference form (e.g. KARJATAMA 'herd'/'scream') and also excluded thirty-one verbs with parallel forms distinguishable only by the phonological length of stem syllable (as in the long vs "overlong" vowel in /sa:stab/ ∼ /sa::stab/ 'pollute.3SG.PRS').[8] The rest of the lexemes were mostly verbs with parallel forms resulting from dual inflection class membership of the verb. Of these, nine were selected for the overabundant condition, each featuring a unique verb in indicative form (*sulgeb* ∼ *suleb* 'close.3SG.PRS').

The nine verbs were chosen to match as closely as possible the lexeme frequency of the French defective verbs. We used ELEXIS (Martelli et al. 2022), a parallel corpus that contains subcorpora for both Estonian and French. For maximal comparability with the French defective data, only the less frequent member of each overabundant doublet features in our items. This choice was driven by the aim to minimise the confounding effect of user experience with the word form: the experience of French speakers with defective forms in usage is unlikely to be frequent, and as such, the closest match in the Estonian overabundant item condition is provided by the less frequent doublet.

Anglicisms used in Estonian stimuli were selected from the Estonian Teen Language Corpus (Vihman et al. 2023), for instance, TŠÄTTIMA 'chat'. Ungrammatical items include non-overabundant verbs from the Estonian National Corpus 2021 (Koppel and Kallas 2022). Their ungrammaticality stems from verb agreement errors. To find invariant verb forms, we ran Qumin (Beniamine 2018) on Eesthetic (Beniamine et al. 2024), an inflected lexicon of Estonian, and selected items that matched as closely as possible the average local entropy associated with the overabundant items.

6.3 Predictions

Our hypothesis was that Estonian's pro-variationist language culture is what allows overabundant forms to thrive across contexts, while the French anti-variationist language culture's aversion to multiplicity leads to word forms being deemed defective. We therefore expected defective and overabundant forms to be sensitive to language ideology in at least some of its operationalisations.

Our predictions for the behaviour of defective items in French were informed by the results of Copot and Sims (2025), which we expected to reproduce. We expected high between-item and between-participant variability in the rating of defective words,

[8] These were excluded because their length-based distinction might be hard to distinguish, might be unnaturally emphasised in audio stimuli, and is not reflected in the orthography.

language	task	ungrammatical	invariant	Anglicism	DEF/OA
Estonian	possible	14	88	49	57
	normative	9	91	28	54
French	possible	23	86	73	55
	normative	13	85	45	51

Table 4 Means by task, item condition and language

reflecting differences in normative pressure by item and in participant awareness of normative targets. We also expected that a formality bias may confound results, with defective words receiving lower ratings in the possibility task due to their unsuitability for informal contexts (see §6.2.2).

Based on acceptance of overabundance in Standard Estonian, we expected high rates of acceptance of overabundance, despite one form in overabundant pairs often being used more often (Aigro and Vihman 2023, 2024). At the same time, recall that destandardisation empowers Estonian speakers to evaluate for themselves what counts as the "best" language. This is a potential source of variability in judgements.

To facilitate comparison across the languages, we use Anglicisms as a baseline for forms that are grammatically generated but proscribed in both. We predicted that French defective forms would be rated on par with or worse than Anglicisms, while Estonian overabundant forms would be rated better than Anglicisms. To the extent that standard language ideology is at play in how participants rated items, these effects should be more pronounced in the normative condition for both languages.

6.4 Results

We first discuss descriptive data for the experimental results, then go on to fit a model to the data.

6.4.1 Descriptive statistics

As can be seen from the raw data in Figure 1, with the means available in Table 4, ungrammatical lexemes were rated at floor in both languages and both tasks, and invariant items were rated at ceiling. In both French and Estonian, Anglicisms were rated worse in the normativity condition than in the possibility condition. High variability is present for both defective and overabundant forms in both task conditions, following a somewhat bimodal distribution. The high variability for French defective lexemes reproduces the findings of Copot and Sims (2025). Figure 2 shows that this variability is not an effect of aggregation and is instead true of most of the individual defective items. A negative relationship between the frequency of the lexeme and the average rating of the item can also be observed. The low average ratings for overabundant items (Figure 1) and the high variability of these ratings are unexpected.

This high variability can be observed within the ratings for each overabundant form (Figure 3) and is therefore again not an artefact of aggregation. We hypothesize this unexpected result is tied to the decision to use the less frequent member of each overabundant doublet. Unlike for French, there does not appear to be a correlation

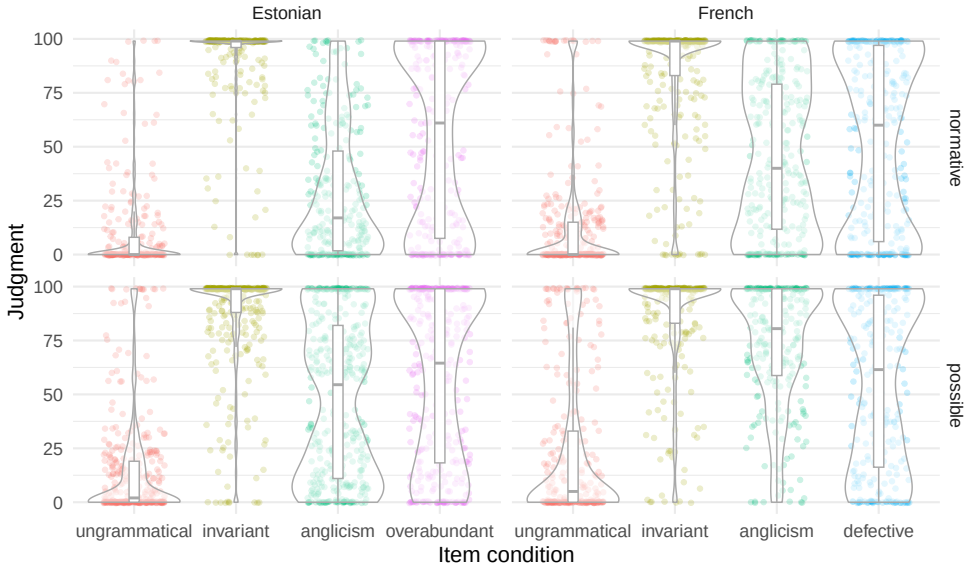


Fig. 1 Raw scores by language, item and task conditions

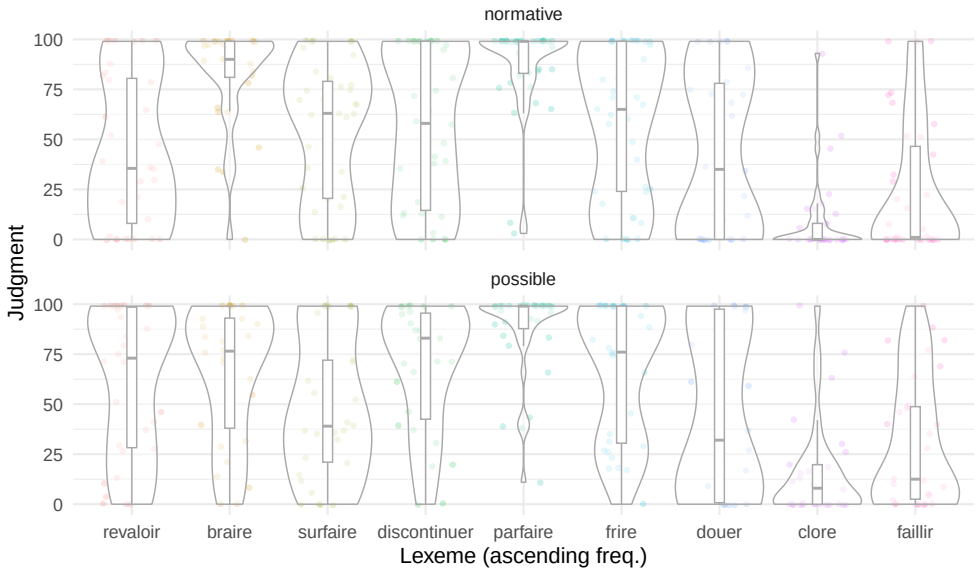


Fig. 2 Defective French lexemes, ordered by ascending lexeme frequency

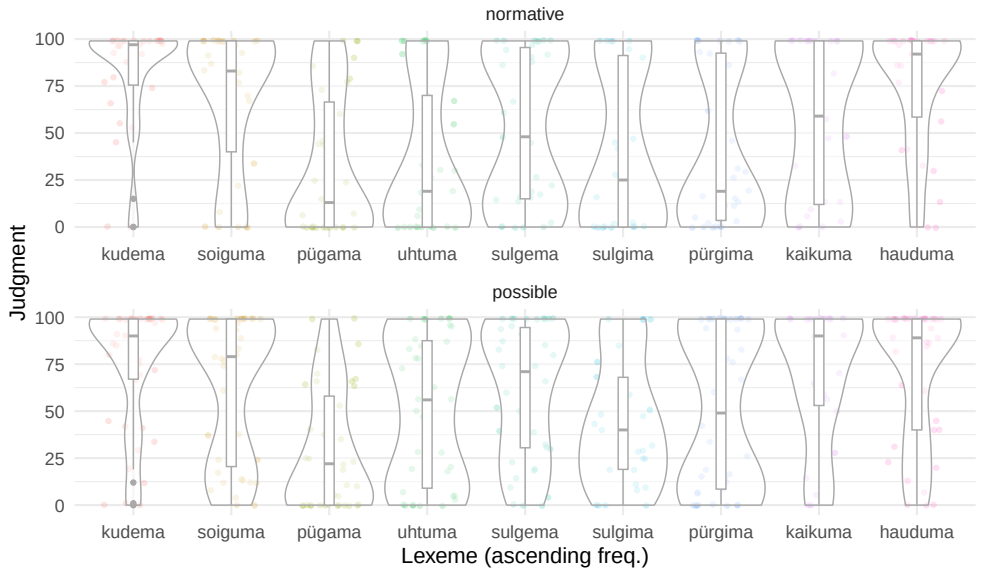


Fig. 3 Overabundant Estonian lexemes, ordered by ascending percentage frequency of the doublet presented in the experiment, relative to its cell mate.

with lexeme frequency, but there does appear to be a correlation with the percentage frequency of the doublet, i.e. what percentage of occurrences of the lexeme in the cell are accounted for by the doublet presented in this experiment. Especially visible in the normative condition, there appears to be a U-shaped trend, with both very high and very low percentages receiving high ratings, and percentages in the mid range receiving lower ratings. This observation is further discussed in §7. §8 provides validation that the low ratings and high variability are linked to the choice of the less frequent doublet.

6.4.2 Inferential statistics

We trained a zero-one-inflated bayesian beta regression with BRMS (Bürkner 2017) on the data. The zero-one-inflation is included to account for participants' tendency to use the extremes of the scale. The choice of the beta regression reflects the fact that the variable we wish to model, acceptability score, is a positive continuous variable bounded between two extremes that represent complete acceptability (100) and complete unacceptability (0). The variable was rescaled to a range between 1 and 0 to fit the requirements of a beta regression.

The model includes the four key variables that characterize the experiment: task condition, item condition, participant language attitude score, and language, in interaction with each other.

Language was contrast-coded, with Estonian as the baseline. Item condition was contrast-coded as a four-level variable. We combined defective and overabundant items into one condition, which we will call target items (overabundant items in Estonian,

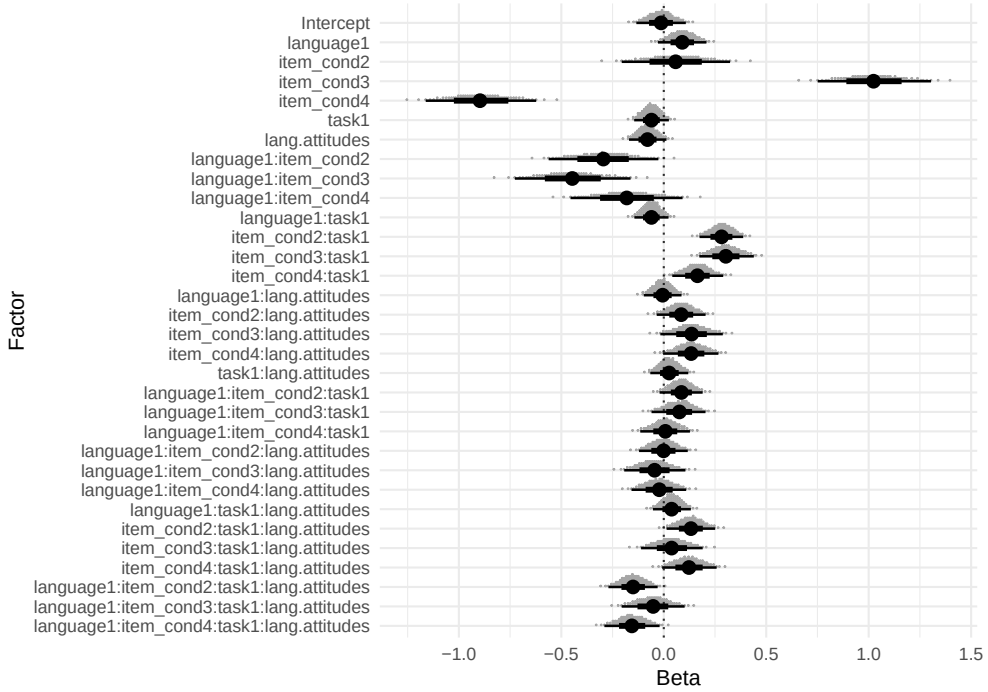


Fig. 4 Plot of coefficient draws for the model presented in Table 6

defective items in French). The effect for the two morphological phenomena will be differentiated by the interaction of this item condition with the language variable. Item condition is treatment-coded as shown in Table 5.

	2	3	4
Anglicism	-0.25	-0.25	-0.25
targets	0.75	-0.25	-0.25
invariant	-0.25	0.75	-0.25
ungrammatical	-0.25	-0.25	0.75

Table 5 Treatment coding scheme for item condition

Anglicisms have been chosen as the baseline for this condition because of their status as normatively stigmatized in both French and Estonian. The scores for this condition serve as a yardstick for how poorly speakers of each language rate items that are grammatical but proscribed. The primary comparison we are interested in is that between target items and Anglicisms in the two languages. Task condition is contrast-coded, with the possibility condition acting as the baseline. Participant language attitude score has been normalised.

	Estimate	Est.Error	-95%CI	+95%CI	BF
Intercept	−0.01	0.06	−0.13	0.11	1.43
lang1	0.09	0.06	−0.03	0.21	13.79
item_cond2	0.06	0.14	−0.21	0.32	2.04
item_cond3	1.03	0.14	0.75	1.30	Inf
item_cond4	−0.89	0.14	−1.16	−0.62	Inf
task1	−0.06	0.04	−0.14	0.02	11.25
lang.attitudes	−0.08	0.05	−0.17	0.01	21.41
lang1:item_cond2	−0.30	0.13	−0.56	−0.03	62.49
lang1:item_cond3	−0.44	0.14	−0.73	−0.16	665.67
lang1:item_cond4	−0.18	0.14	−0.45	0.09	9.22
lang1:task1	−0.06	0.04	−0.14	0.02	11.16
item_cond2:task1	0.28	0.06	0.17	0.39	Inf
item_cond3:task1	0.30	0.07	0.17	0.44	Inf
item_cond4:task1	0.16	0.06	0.04	0.29	209.53
lang1:lang.attitudes	−0.01	0.05	−0.10	0.09	1.22
item_cond2:lang.attitudes	0.09	0.06	−0.03	0.20	11.88
item_cond3:lang.attitudes	0.14	0.08	−0.02	0.29	23.39
item_cond4:lang.attitudes	0.13	0.07	−0.00	0.27	35.20
task1:lang.attitudes	0.03	0.05	−0.07	0.12	2.50
lang1:item_cond2:task1	0.09	0.05	−0.02	0.19	17.31
lang1:item_cond3:task1	0.08	0.07	−0.06	0.20	6.69
lang1:item_cond4:task1	0.01	0.06	−0.11	0.13	1.22
lang1:item_cond2:lang.attitudes	−0.00	0.06	−0.12	0.12	1.03
lang1:item_cond3:lang.attitudes	−0.04	0.08	−0.19	0.11	2.56
lang1:item_cond4:lang.attitudes	−0.02	0.07	−0.16	0.11	1.70
lang1:task1:lang.attitudes	0.04	0.05	−0.05	0.13	3.82
item_cond2:task1:lang.attitudes	0.13	0.06	0.01	0.25	67.38
item_cond3:task1:lang.attitudes	0.04	0.08	−0.11	0.19	2.25
item_cond4:task1:lang.attitudes	0.12	0.07	−0.01	0.26	29.30
lang1:item_cond2:task1:lang.attitudes	−0.15	0.06	−0.27	−0.03	139.35
lang1:item_cond3:task1:lang.attitudes	−0.05	0.08	−0.21	0.10	2.99
lang1:item_cond4:task1:lang.attitudes	−0.15	0.07	−0.29	−0.02	79.81

Table 6 Coefficients for the model in Experiment 1.

The coefficients for the model are shown in Table 6. Along with the coefficients, we also report the Bayes Factor (BF), indicating the amount of evidence available for the coefficient having the reported monotonicity,[9] shown visually in Figure 4.

6.4.3 Main effects and two-way interaction

The main effect of language suggests French speakers rated items somewhat higher than Estonian speakers overall. As expected, invariant items (item_cond3) received dramatically higher ratings than Anglicisms (β = 1.03, 95% CI: 0.75, 1.32, ER ¿ 100), while ungrammatical items (item_cond4) were rated substantially lower (β = -0.89, 95% CI: -1.16, -0.62, ER ¿ 100). Target items (item_cond3) showed a modest positive effect relative to Anglicisms (β = 0.07, 95% CI: -0.20, 0.34), though with limited

[9]To obtain this figure, we divide the proportion of the posterior with monotony corresponding to that of the estimate by the proportion of the posterior with the opposite monotony. This is equivalent to an evidence ratio between positive and negative proportions of the posterior. We rely on Kass and Raftery (1995) for guidelines about the correspondence between BF magnitude and amount of evidence for an effect and discuss the effects in line with this.

evidence (ER = 2.29). The main effect of the target item condition is nevertheless not really interpretable since it merely provides the combined average difference from Anglicisms for both defective and overabundant items. The task condition main effect (task1 = -0.06, 95% CI: -0.14, 0.02) indicates that ratings were generally lower in the normative condition, consistent with increased normative pressure in that context.

Critically, the two-way interaction lang1:item_cond2 (β = -0.29, 95% CI: -0.56, -0.02, ER = 57.39) reveals that French defective items were rated substantially lower than Estonian overabundant items relative to their respective Anglicism baselines. Combined with the small positive coefficient for the target item condition, this indicates a reversal of the relative ratings of Anglicisms and target items in the two languages, where overabundant forms are generally rated higher than Anglicisms in Estonian, but defective items are generally rated lower than Anglicisms in French. This pattern aligns with our hypothesis that antivariationist French culture leads to more negative evaluation of defective forms compared to pro-variationist acceptance of overabundant forms in Estonian.

A comparison of the coefficients involving item_cond2 and item_cond3 shows that in Estonian, despite overabundant items being rated higher than Anglicisms, the size of this difference is much smaller in magnitude than that between Anglicisms and invariant items. This is a surprising finding, as discussions of overabundance in the literature assume that overabundant doublets are just as acceptable and well-formed as any other wordform in the language's lexicon. We address this finding in §7, and use it as the basis for the experiment in §8.

6.4.4 Three-way interactions: task context effects

On average, target items are rated better in the normative condition compared to the possibility condition (item_cond2:task1, β = 0.28, 95% CI: 0.17, 0.39, ER ¿ 100). This is surprising, on one hand, because no such effect was expected in Estonian (since we expected overabundant forms to be rated at ceiling), but also because we expected a negative effect for defective items. The three-way interaction language1:item_cond2:partic_condition1 (β = 0.08, 95% CI: -0.02, 0.19, ER = 15.74) suggests this unexpected increase in scores may actually be driven by French defective items, though the evidence is only moderate. Copot and Sims (2025) find the same unexpected positive effect of task condition for defective items, decreasing the possibility that the effect we find is a false positive.

Our explanation for this effect lies in the formality scores reported in Table 3, which confirmed that lexemes that are deemed defective by dictionaries are more likely to be associated with formal contexts. Because formality and normativity are non-orthogonal dimensions, our interpretation is that the positive effect on judgements of seeing defective lexemes in a formal context compared to an informal one is greater in magnitude than any additional negative effect of thinking the forms themselves are ill-formed.

6.4.5 Language attitudes and the critical four-way interaction

Participants' language attitude scores showed a general tendency toward more negative ratings (lang.attitudes = -0.08, 95% CI: -0.17, 0.01, ER = 20.22), providing strong

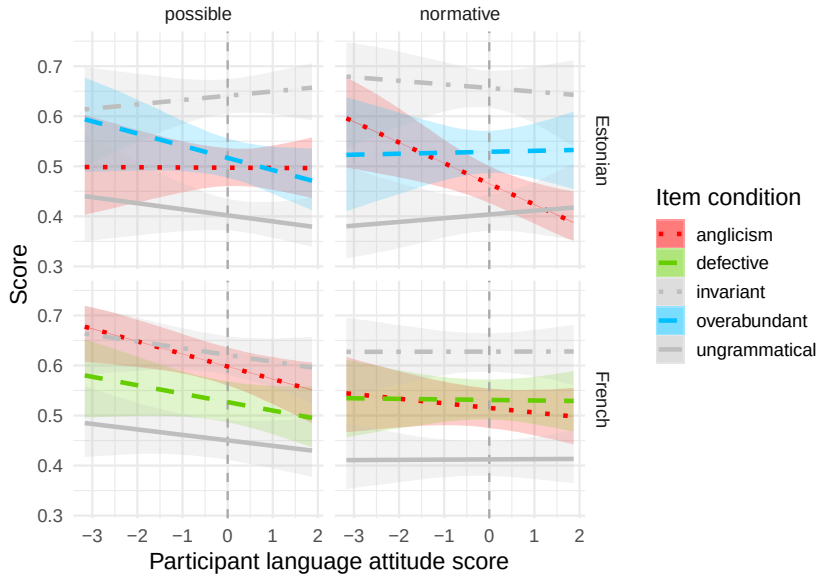


Fig. 5 Conditional effects plot for the model trained on data from Experiment 1

evidence that individuals who are more oriented towards linguistic conservativism gave lower acceptability ratings overall. The interaction item_cond2:lang.attitudes (β = 0.09, 95% CI: -0.03, 0.21, ER = 11.84) suggests that a higher language attitude score was associated with relatively more positive ratings of target items compared to Anglicisms, with substantial evidence. The three-way interaction item_cond2:task1:lang.attitudes (β = 0.13, 95% CI: 0.01, 0.25, ER = 61.02) indicates that people with higher language attitude scores tended to rate target items as better than Anglicisms in the normative task, demonstrating evidence for contextual modulation of scores for target items.

The four-way interaction language1:item_cond2:task1:lang.attitudes (β = -0.15, 95% CI: -0.27, -0.03, ER = 120.21) – visualised via a conditional effects plot in Figure 5 – provides decisive evidence for our central hypothesis. This negative coefficient indicates that the positive effect of language attitude score on target items compared to Anglicisms in the normative condition was significantly attenuated in French compared to Estonian. The emerging picture suggests that overabundant and defective items are sensitive both to task effects manipulating the engagement of linguistic knowledge and to the language ideologies held by the participant.

However, the pattern of influence of contextual and metalinguistic factors is different to what was predicted but surprisingly similar across both: as can be seen in Figure 5, both defectiveness and overabundance exhibit a negative effect of language attitude score in the possibility condition and a flat effect in the normative condition. This suggests that people more attuned to standard language ideologies are less likely to think that defective and overabundant forms are used (this is surprising for overabundant forms), but when people are asked to think metalinguistically about

what their language deems correct, participants judge the acceptability of defective and overabundant forms to be similar regardless of their own language attitudes (both phenomena are given middling ratings on average). For French, this pattern matches that for Anglicisms, but for Estonian, Anglicisms have inverted behaviour with regard to language attitude score in the two task conditions: the slope is flat in the possibility task, and negative in the normative task.

To sum up, we observe the predicted asymmetry in the way that French defective and Estonian overabundant items are rated compared to Anglicisms, suggesting differing acceptability of the two phenomena in the respective standard languages. The four-way interaction provides decisive evidence (ER ¿ 100) that modulation of acceptability operates differentially across standard language cultures. However, we observe unexpected parallels in the behaviour of overabundant and defective items, such as a negative slope for language attitude score in the possibility task and a neutral one in the normative task. We return to these unexpected similarities in the rest of the paper.

7 Interlude: preemption

The results in §6.4 strongly point to the ratings of both overabundant and defective words being modulated by the amount of metalinguistic knowledge engaged by the task and to the language ideological inclinations of individual participants, suggesting that a holistic account of these phenomena would benefit from increased attention to extralinguistic factors. However, the predictions we laid out for how the different types of items would be affected were falsified in a couple ways. Particularly surprising are the comparatively low ratings of Estonian overabundant items and the multiple parallels in both raw score and the effect of task and language attitudes score on defective and overabundant items.

An important design choice in Experiment 1 was the selection of the *lower frequency doublet* for the overabundant items in Estonian. This was intended to counteract possible effects of token frequency for acceptability judgments, thus increasing the comparability of defective and overabundant items in this respect. However, the low acceptability ratings given for overabundant items are unexpected in light of the literature describing overabundance as a situation where more than one word form is equally acceptable as a way to express a lexeme in a particular paradigm cell (Thornton 2019; Bermel et al. 2023). We suggest that this result may be interpreted as a case of statistical preemption (Boyd and Goldberg 2011). In such a scenario, the acceptability of the lower frequency doublet is (either consciously or subconsciously) judged against the acceptability of the more frequently encountered and presumably preferred doublet, leading to lower scores for the lower frequency doublet as a result of the implicit comparison. This would mean that making overabundant items more comparable to defective ones in frequency may have led to them being penalized for being the less preferred member of a doublet pair.

Designing Experiment 1 around the less frequent form from each doublet might also explain the peculiar pattern observed with regards to the effect of language ideology and task on overabundant items: participants who have more conservative attitudes

about language may have been more aware of the dispreferred status of the less frequent form in usage, potentially explaining the negative slope for language attitudes score in the possibility task. But all Estonian speakers, regardless of their individual language attitudes scores, are aware that these forms are generated by Estonian grammar, leading to the absence of an effect for individual language attitudes in the normative task.

We therefore propose that the pattern of rating observed for overabundant forms in Experiment 1 is the result of preemption of the less frequent form in each doublet by the more frequent one. This is in tension with the common assumption that, because overabundant forms are equally licensed by the grammar, they should be equally acceptable to speakers. In §8, we perform the same experiment on Estonian only, this time including both overabundant doublets in the experiment, to substantiate our hypothesis.

8 Experiment 2

We performed a follow-up experiment identical to that described in §6, except that we changed the experiment items to expand the set of overabundant items investigated.

8.1 Participants

A total of 67 participants were recruited on Prolific. All participants reported being native speakers of Estonian, having grown up speaking it as their only language in most day-to-day situations, and currently residing in Estonia. Following the lexeme knowledge attention check described in §5.3, 15 participants were excluded. Of the retained participants, 24 completed the normative task and 28 completed the possiblity task. For these participants, the median age was 28 (range: 20-51). The gender breakdown was 25 females, 27 males. Two people reported an elementary diploma as their highest educational attainment, 18 had a high school diploma, 19 had an undergraduate diploma, 12 had a masters' diploma or above, and one person declined to answer. The language attitudes questionnaire described in §5.3 revealed that participants scored a median result of 18 out of a possible 36 (range: 9–29, Quartile 1: 16; Quartile 3: 23). (Note that this median is somewhat lower compared with the scores for Estonian participants in Experiment 1, and Experiment 2 had a slightly narrower range of scores overall.)

8.2 Items

For this experiment, which only involves Estonian, we augmented the overabundant item condition. These overabundant items were shown alongside the same Anglicisms and ungrammatical items from Experiment 1. Invariant items were not included.

For the overabundant condition, we included the nine lexemes used in Experiment 1 (§6.2.3) and also verbs with overabundance in both participial and infinitive forms (*kõnelenud* ∼ *kõnelnud* 'speak.PST.PTCP'; *kõneleda* ∼ *kõnelda* 'speak.INF'). Nine participle examples and nine infinitival examples were selected. In addition, unlike in the first experiment, both forms of these twenty-seven overabundant verbs were included

in the set of stimuli, rather than only the less frequent form. A sentence frame was created for each overabundant doublet, and only one form of each doublet was shown to each participant. Participants on average saw half of the frequent forms and half of the infrequent ones.

In most of the overabundant pairs (23 out of 27), one variant is significantly more frequent than the other, making up 75–99% of uses in the ELEXIS corpus. For four pairs the distribution was more even, the dominant form being used in 57–70% of instances. Detailed frequency information is available in the OSF repository.

8.3 Analysis

We fit a zero-one-inflated beta regression to test the effect of different form proportions on acceptability ratings. Parallel form frequency is coded as distinct item conditions, meaning the model compares ungrammatical items, Anglicisms, infrequent overabundant forms and frequent overabundant forms. We expected that, similarly to Experiment 1, ungrammatical items would be rated at floor in both task conditions, and that Anglicisms would receive higher judgements in the possibility condition. We expected the more frequent overabundant variant to be rated at ceiling, while the less frequent variant was expected to show the same type of variation observed in Experiment 1.

Both item and task condition were treatment coded. Infrequent overabundant items are the baseline for the item condition (see Table 7), and the possibility task is the baseline for the task condition. Language attitudes score is centred and standardised.

	2	3	4
less frequent variant	-0.25	-0.25	-0.25
Anglicism	0.75	-0.25	-0.25
more frequent variant	-0.25	0.75	-0.25
ungrammatical	-0.25	-0.25	0.75

Table 7 Treatment coding scheme for item condition

8.4 Results

Figure 6 shows the raw data pertaining to overabundant forms, broken down by task, whether the overabundant variant is the more or less frequent one, and tense-aspect-mood (TAM) value. TAM value does not appear to matter in the acceptability of forms and will be ignored in the remainder of this article. There is a large, clear difference between the frequent and infrequent member of each overabundant doublet, with the frequent member being rated at ceiling and the infrequent member receiving more variable but generally much lower scores.

The coefficients of the statistical model are shown in Table 8, and the conditional effects are plotted in Figures 7 and 8. The model shows a positive effect of language attitude score; participants with higher language attitude scores give higher ratings on

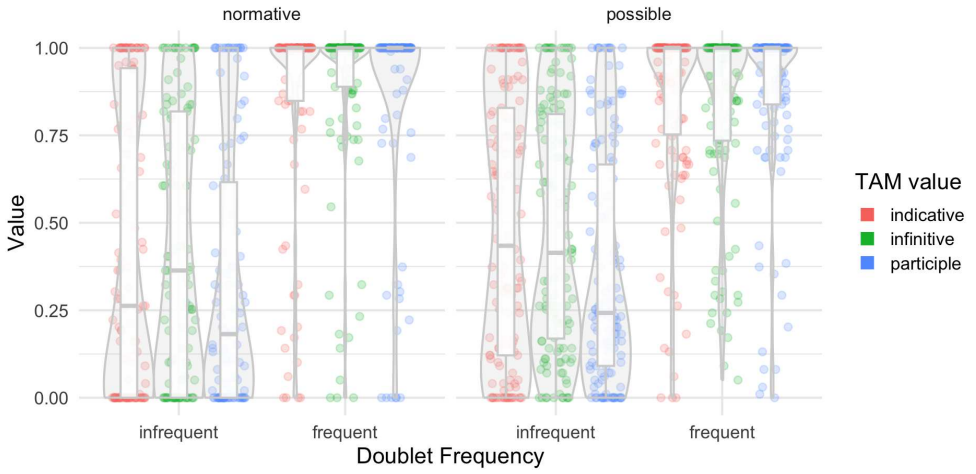


Fig. 6 Ratings for overabundant forms by task condition, grouped by the relative frequency of the doublet in the cell and by the TAM values of the form

	Estimate	Est.Error	-95%CI	+95%CI	BF
Intercept	−0.18	0.08	−0.34	−0.03	101.56
lang.attitudes	0.08	0.07	−0.06	0.23	6.74
oa_freq2	0.14	0.11	−0.08	0.37	9.09
oa_freq3	0.86	0.12	0.63	1.10	*Inf*
oa_freq4	−0.50	0.14	−0.77	−0.23	*Inf*
task1	−0.18	0.15	−0.47	0.11	8.08
lang.attitudes:oa_freq2	0.00	0.08	−0.16	0.17	1.09
lang.attitudes:oa_freq3	0.03	0.11	−0.18	0.23	1.49
lang.attitudes:oa_freq4	0.06	0.10	−0.14	0.27	2.54
lang.attitudes:task1	0.06	0.14	−0.22	0.34	1.88
oa_freq2:task1	−0.61	0.16	−0.93	−0.29	*Inf*
oa_freq3:task1	−0.12	0.20	−0.51	0.28	2.57
oa_freq4:task1	0.39	0.22	−0.04	0.81	24.97
lang.attitudes:oa_freq2:task1	−0.61	0.17	−0.94	−0.27	7999.00
lang.attitudes:oa_freq3:task1	0.04	0.21	−0.38	0.46	1.39
lang.attitudes:oa_freq4:task1	0.36	0.21	−0.05	0.78	22.05

Table 8 Coefficients for the model in Experiment 2

average. Both Anglicisms (item_cond2) and the more frequent overabundant variants (item_cond3) are on average rated better than the less frequent overabundant variants. The only condition that is rated lower on average than infrequent overabundant forms is ungrammatical items (item_cond4). Additionally, ratings in the normative task are lower compared to the possibility task (task1). The relative difference between the infrequent overabundant items and Anglicisms is reversed in the normative condition (item_cond2:task1). This is the result of Anglicisms being rated lower, rather than the

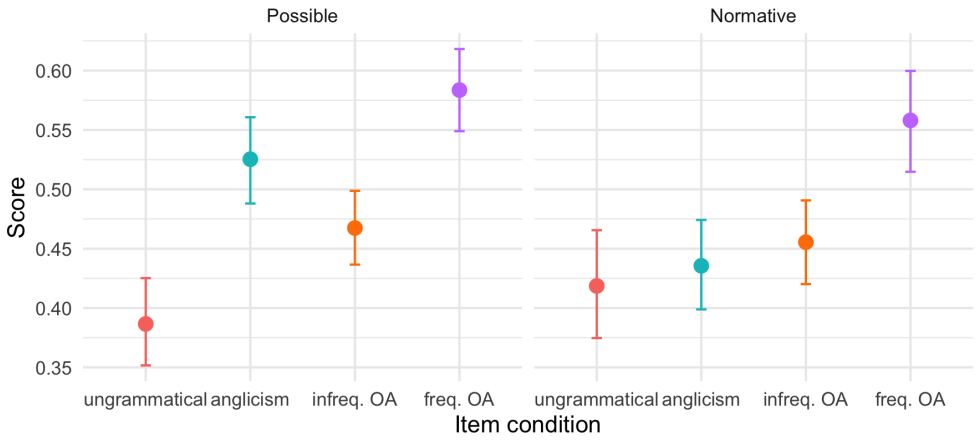


Fig. 7 Conditional effects plot of the model for Experiment 2, showing the estimates for the different item conditions split by participant condition

lower-frequency overabundant form being rated higher, and the difference with the higher frequency overabundant form is somewhat reduced (item_cond3:task1).

There is scarce evidence for any main effect of language attitude score by item condition or task, though it does appear to have an effect when both are considered together. We find a reversal of the language attitude slopes between Anglicisms and the less frequent overabundant items: participants expressing conservative language attitudes are more likely to recognise that Anglicisms are used (positive slope in the possibility task) but not traditionally acceptable (negative slope in the normative task). For the less frequent overabundant items, it is the opposite: they have a flat slope in the normativity task, but the higher a participant's language attitude score, the lower their rating for the infrequent variant in the possibility condition. This suggests that participants aligned with standard language discourses are more likely to know that Estonian grammar normatively accepts the use of all forms generated by its rules, even if alternatives might be more common. There is little evidence that ungrammatical items are sensitive to task condition, being rated at floor both in normative and possibility conditions. Ungrammatical stimuli and Anglicisms do not differ significantly from the lower-frequency overabundant variants, although Anglicisms are more sensitive to task conditions (Figure 8). The higher-frequency overabundant variants receive significantly higher ratings than their lower-frequency counterparts.

The difference between parallel forms shown in Figure 6 is also present when parallel form proportion is treated in a continuous fashion: Pearson correlation between judgements of overabundant forms and form proportion in a given cell was 0.52, expressing moderate correlation.

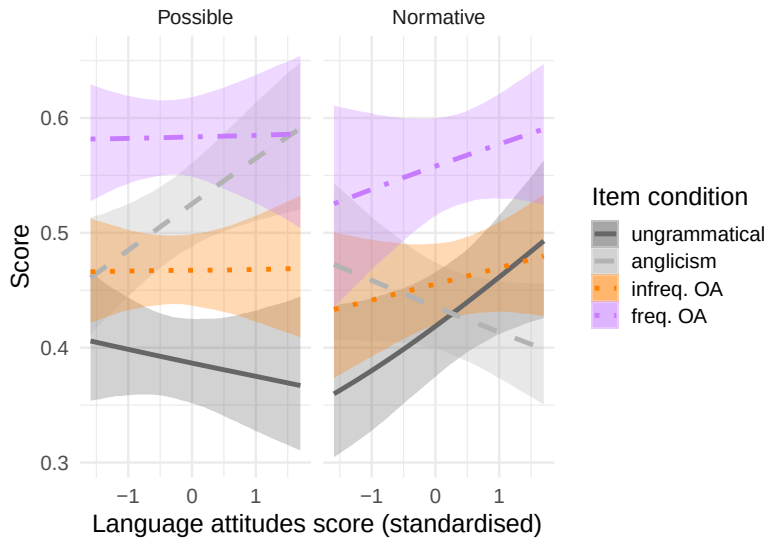


Fig. 8 Conditional effects plot of the model in Experiment 2, showing the interaction of all the terms in the model

9 Discussion

This study set out to investigate whether the difference between two non-canonical morphological phenomena – defectiveness in French and overabundance in Estonian – could be partially explained by differences in the standard language ideologies of the two languages. The results offer partial support for this hypothesis. In what follows, we discuss the findings in turn, beginning with the ideology-related results in light of the frameworks introduced in §2, before turning to what the results reveal about the nature of overabundance, and finally to what the unexpected similarities between the two phenomena imply for how defectiveness and overabundance are theorised in relation to one another.

9.1 The effects of language ideology

The introduction posed a specific question: if stigmatisation arising from anti-variationist language planning partially causes defectiveness in French (Copot and Sims 2025), does the converse hold – does an ideology that accepts variation partially cause overabundance? The language-level asymmetry we observe in Experiment 1 is consistent with this claim: relative to Anglicisms, which are stigmatised in both languages and thus provide a shared yardstick for prescriptive penalisation, overabundant forms in Estonian fare better than defective forms do in French (§6.4). In the condition most likely to activate knowledge of standard language norms (the normative task), Estonian speakers with more conservative language attitudes differentiate overabundant forms from Anglicisms (rating the latter lower), while French speakers who align themselves with the standard language ideology do not differentiate defective forms from Anglicisms. This is the pattern predicted by the ideological contrast laid out in

§1: French language planning, rooted in establishment authenticity, treats variation as error, while Estonian language planning, rooted in vernacular authenticity, codifies it.

At the same time, the way individuals' language ideologies modulate ratings departs from predictions in informative ways. We expected the language attitude scores to negatively affect participants' ratings for French defective items. For Estonian overabundant items, we expected an overall positive effect of language attitude score, but also an effect of participants' own experience with the language. Instead, both phenomena show the same profile: a negative language attitude score slope in the possibility task and a flatter one in the normative task. This parallel is initially puzzling but becomes interpretable once we consider what the attitude score likely indexes in each language culture. In French, alignment with standard language discourses entails alignment with an ideology that proscribes the target forms; more prescriptive speakers are thus less willing to grant that defective forms "could be heard". In Estonian, alignment with standard language discourses may instead index metalinguistic attentiveness – a closer awareness of which forms are dominant in usage – without entailing rejection of the less frequent variant, since the Estonian standard explicitly endorses both. The negative slope in the Estonian possibility task then reflects not stigmatisation but accurate frequency awareness: more metalinguistically attuned speakers recognise the presented form as the minority variant. Meanwhile, the flat slope in the normative task follows naturally, since all speakers can recognise that Estonian grammar licenses both forms regardless of their personal language ideological leanings. In short, language attitude scores from parallel questionnaires may measure partially different stances across language cultures – something that was already suggested by the fact that the Estonian questionnaire seems to measure two distinct underlying ideological constructs (§5.2). This complication for cross-linguistic comparison deserves attention in future work.

This result has a broader implication as well. Within each language, the task manipulation and individual language attitude scores modulate ratings of target items in ways that would be unexpected if the phenomena were purely grammar-internal, providing controlled, granular evidence that social evaluation penetrates inflectional judgements. Moreover, the hypothesis finds independent support in diachronic work: in Italian, prescriptive intervention targeted frequent overabundant lexemes, and overabundance persists today primarily in lower-frequency items that escaped such attention (Thornton 2012); in Russian, language standardisation processes have been implicated in the development of 1SG gaps (Baerman 2008). These cases supply mechanistic evidence — showing how standard language culture can reshape inflectional systems over time — that complements the synchronic snapshot we provide here. Taken together, the evidence is sufficient to warrant systematic investigation of how standard language regimes correlate with the distribution of non-canonical phenomena across typologically varied languages. Languages exhibiting both defectiveness and overabundance, such as Italian and Czech, offer particularly promising test cases, since they allow the ideological asymmetry to be examined within a single system.

9.2 The relationship between overabundance and defectiveness

The introduction framed defectiveness and overabundance as opposite deviations from the canonical deterministic paradigm – too few forms versus too many. The previous section addressed their relationship on the grounds of language ideology. Here we turn to what is revealed about the two phenomena by the profile of their acceptability ratings themselves.

The most striking finding is the convergence between the less frequent member of Estonian overabundant doublets and French defective forms. Both show high between-participant and between-item variability, both are rated higher in the normative than the possibility condition, and both show parallel slopes according to language attitude score (§9.1). These similarities are not predicted by existing, structurally-based accounts. Notably here, threshold models (Pertsova 2016) postulate that there is an important structural distinction between situations in which there is more than one reliable inflectional pattern and ones in which there is no reliable pattern at all. Threshold models treat morphological well-formedness as a gradient property that is predicted, for example, from the reliability with which inflectional rules apply in some conditioning environment (Albright and Hayes 2003). There may be no reliable inflectional rule in some context or multiple ones. This, in turn, predicts that the two situations should feel different to speakers – one is the absence of any sufficiently acceptable form, the other the availability of multiple acceptable alternatives. For example, Pertsova (2016, 7) argues that a threshold approach should allow "... us to distinguish between two conceptually different cases [of form competition]... One can imagine that both scenarios would lead to variation, but in the first case we want the model to predict speakers to be equally satisfied with the variants, while in the second case they should be equally unhappy with the variants, a situation that results in a paradigm gap". Yet in our study, when we compare the acceptability profiles of defectiveness and overabundance, the two phenomena converge. This suggests that the two phenomena cannot be straightforwardly interpreted as reflecting opposite structural conditions.

Experiment 2 sharpens this picture: when both members of overabundant doublets are tested, the frequent variant is rated at ceiling while the infrequent variant shows the same depressed, high-variability profile observed in Experiment 1. The item-level correlation between ratings and corpus form proportion confirms that acceptability tracks usage frequency. We interpret this as statistical preemption (Boyd and Goldberg 2011) operating within inflectional paradigms: the more frequent doublet member, reliably attested in contexts where the target would occur, implicitly signals to speakers that it is the conventional way to express the relevant function, penalising the less frequent alternative.

This extends the preemption mechanism into a new domain, and the extension is not trivial. Previous work has demonstrated preemption effects for argument structure overgeneralisation (Ambridge et al. 2018; Samara et al. 2025) and derivational morphology (Boyd and Goldberg 2011), where the competing forms are structurally distinct (e.g. ditransitive versus PP-dative) and the preempted form is typically ungrammatical or marginal. Inflectional overabundance differs in a theoretically important respect: the competing forms are exact functional synonyms occupying the

same paradigm cell – a purer case of competition than the syntactic alternations where preemption is usually studied. That the mechanism operates here, and does so against explicit normative endorsement of both variants, strengthens the case for its generality.[10]

This finding challenges a tacit assumption in the overabundance literature. Theoretical treatments from Thornton (2011) onward generally assume that if a cell is overabundant, both forms are available to the speaker in a roughly equivalent way. Corpus evidence already suggested otherwise: Aigro and Vihman (2023) show that realised overabundance in Estonian is dramatically more restricted than the potential overabundance posited by reference grammars, with most lexemes showing strong skew toward one variant, and Aigro and Vihman (2024) demonstrate that cell token frequency and inflection class – not semantic factors – predict which variant dominates. Our data provide the acceptability complement to these production findings: the frequency asymmetry documented in corpora surfaces as an acceptability asymmetry even when the task frames the question normatively. Taken together, these findings raise a question about the ontological status of overabundance: if one variant is strongly dispreferred in both production and judgement, is the cell genuinely overabundant in the speaker's grammar, or is it better described as a cell with one entrenched form and one marginal competitor? Reference grammars may describe a property of the linguistic system that does not straightforwardly correspond to a property of speaker knowledge.

Both defectiveness and overabundance also show frequency-mediated acceptability effects, but along different dimensions. For French defective items, *lexeme* frequency drives lower ratings: speakers encounter high-frequency lexemes often enough to have learned their stigmatised status, consistent with the negative frequency effect in Copot and Sims (2025).[11] For Estonian overabundant items, it is the *within-cell proportion* of the doublet that matters, with a U-shaped pattern: both very dominant and very rare variants are rated higher than mid-range ones. The U-shape is noteworthy because simple preemption predicts a monotonic relationship – the more dominant the competitor, the worse the preempted form. That very rare variants escape penalisation suggests that when a form is infrequent enough, speakers lack sufficient evidence to have strong expectations about the cell, and so neither variant is penalised. This parallels the logic of why low-frequency lexemes escape defectiveness in French (Copot and Sims 2025): insufficient exposure cuts both ways, shielding forms from both stigmatisation and preemption. The non-monotonic profile of preemption effects in inflectional paradigms presents questions that warrant further investigation.

Ultimately, defectiveness and overabundance are not straightforwardly opposite phenomena. They share a variability profile and a sensitivity to frequency-driven competition that existing models do not predict. But they are not identical: in the language pair examined here, their presence in different linguistic ideological ecologies produces differences – even the less frequent overabundant variants fare better than Anglicisms in Estonian, while defective forms in French do not – and the frequency dimensions

[10] It also suggests that overabundant paradigms could serve as a productive testbed for preemption research, offering controlled competition within a single functional slot without the confounds of structural or semantic non-equivalence that complicate argument structure studies.

[11] See Rowe et al. (2026) for further probing of word frequency effects in defectiveness.

involved differ. The label "opposite phenomena" accurately describes the grammar's output (zero forms versus multiple forms) but not the speaker's experience, which converges when the speaker confronts a form that is grammatically motivated but not fully entrenched. A more adequate framing may be that paradigm cells vary in the degree of *determinacy* speakers command over them, and that both defectiveness and overabundance occupy regions of low determinacy, arrived at for structurally and socially different reasons.

10 Conclusions

This study set out to test whether standard language ideology is the missing factor that determines whether competition among inflectional patterns resolves as defectiveness or overabundance. The answer is: partially. The macro-level asymmetry between French and Estonian supports the hypothesis, and the within-language sensitivity to language attitudes confirms, that language ideology penetrates inflectional judgements. But the picture is richer than a simple model in which standard language norms act as a filter on grammatical forms. Frequency-driven entrenchment creates preference hierarchies even within normatively sanctioned variation, and task framing modulates which layer of speaker knowledge surfaces. This suggests a complex interaction between grammatical structure, usage-based entrenchment, and the evaluative frameworks that standard language cultures provide.

Declarations

- Funding This work was supported by X1 and X2 grants awarded to Author 1 and by X3 grant awarded to Author 4.

References

Aigro, M. and V.A. Vihman. 2023. Realised overabundance in Estonian noun paradigms: A corpus study. *Word Structure* 16: 154–175. https://doi.org/10.3366/word.2023.0227 .

Aigro, M. and V.A. Vihman. 2024. Preferences in the use of overabundance: Predictors of lexical bias in Estonian. *Cognitive Linguistics* 35: 289–312. https://doi.org/10.1515/cog-2023-0035 .

Albright, A. and B. Hayes. 2003. Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition* 90: 119–161. https://doi.org/10.1016/S0010-0277(03)00146-X .

Algvere, K. 2025. Adult L1 users' attitudes towards their proficiency in written Estonian as evidence of deep-rooted language ideologies. *Eesti ja Soome-Ugri Keeleteaduse Ajakiri 16*(2): 5–50. https://doi.org/10.12697/jeful.2025.16.2.01 .

Ambridge, B., L. Barak, E. Wonnacott, C. Bannard, and G. Sala. 2018. Effects of both preemption and entrenchment in the retreat from verb overgeneralization errors: Four reanalyses, an extended replication, and a meta-analytic synthesis. *Collabra: Psychology 4*(1): 23. https://doi.org/10.1525/collabra.133 .

Anderson, B. 1983. *Imagined communities: Reflections on the origin and spread of nationalism*. London: Verso.

Auer, P. 2005. Europe's sociolinguistic unity, or: A typology of European dialect/standard constellations, In *Perspectives on variation: Sociolinguistic, historical, comparative*, eds. Delbecque, N., J.v.d. Auwera, and D. Geeraerts, 7–42. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110909579.7.

Ayres-Bennett, W. 2021. Modelling language standardization, In *The Cambridge handbook of language standardization*, eds. Ayres-Bennett, W. and J. Bellamy, 27–64. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108559249.002.

Bach, X. and L. Esher. 2015. Morphological evidence for the paradigmatic status of infinitives in French and Occitan, In *Historical linguistics 2013: Selected papers from the 21st International Conference on Historical Linguistics, Oslo, 5-9 August 2013*, ed. Haug, D.T., 135–154. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.334.08bac.

Baerman, M. 2008. Historical observations on defectiveness: The first singular nonpast. *Russian Linguistics* 32: 81–97. https://doi.org/10.1007/s11185-007-9017-9 .

Beniamine, S. 2018. *Typologie quantitative des systèmes de classes flexionnelles*. Ph. D. thesis, Université Paris Diderot, Paris.

Beniamine, S., M. Aigro, M. Baerman, J. Bouton, and M. Copot 2024, May. Eesthetic: A Paralex lexicon of Estonian paradigms. In N. Calzolari, M.-Y. Kan, V. Hoste, A. Lenci, S. Sakti, and N. Xue (Eds.), *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, Torino, Italia, pp. 5526–5537. ELRA and ICCL.

Beniamine, S., M. Coavoux, and O. Bonami. 2024, February. Vlexique 2.0. Zenodo.

Bermel, N., L. Knittl, and A. Nikolaev. 2023. Uncertainty in the production of Czech noun and verb forms. *Word Structure* 16: 258–283. https://doi.org/10.3366/word.2023.0231 .

Blevins, J.P. 2007. Conjugation classes in Estonian. *Linguistica Uralica 43*(4): 250–267. https://doi.org/10.3176/lu.2007.4.02 .

Bonami, O. and G. Boyé. 2002. Suppletion and dependency in inflectional morphology, In *Proceedings of the 8th International Conference on Head-Driven Phrase Structure Grammar*, eds. van Eynde, F., L. Hellan, and D. Beermann, 51–70. Stanford, CA: CSLI.

Bonami, O., G. Caron, and C. Plancq. 2014. Construction d'un lexique flexionnel phonétisé libre du français, In *Quatrième Congrès mondial de linguistique française*, eds. Neveu, F., P. Blumenthal, L. Hriba, A. Gerstenberg, J. Meinschaefer, and S. Prévost, 2583–2596. Les Ulis, France: EDP Sciences. https://doi.org/10.1051/shsconf/20140801223.

Boyd, J.K. and A.E. Goldberg. 2011. Learning what NOT to say: The role of statistical preemption and categorization in a-adjective production. *Language 87*(1): 55–83. https://doi.org/10.1353/lan.2011.0012 .

Boyé, G. 2000. *Problèmes de morpho-phonologie verbale en français, en espagnol et en italien.* Ph. D. thesis, Université Paris Diderot, Paris.

Boyé, G. and P. Cabredo Hofherr. 2010. Defectiveness as stem suppletion in French and Spanish verbs, In *Defective paradigms: Missing forms and what they tell us*, eds. Baerman, M., G.G. Corbett, and D. Brown, 35–52. London: Oxford University Press, in coordination with British Academy Press. https://doi.org/10.5871/bacad/9780197264607.003.0003.

Bürkner, P.C. 2017. brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software 80*(1): 1–28. https://doi.org/10.18637/jss.v080.i01 .

Copot, M. and A.D. Sims. 2025. The influence of prescriptivism on French verbal defectiveness. *Linguistics* 63: 1293–1338. https://doi.org/10.1515/ling-2023-0028 .

Coupland, N. 2003. Sociolinguistic authenticities. *Journal of Sociolinguistics 7*(3): 417–431. https://doi.org/10.1111/1467-9481.00233 .

Coupland, N. and T. Kristiansen. 2011. SLICE: Critical perspectives on language (de)standardisation, In *Standard languages and language standards in a changing Europe*, eds. Kristiansen, T. and N. Coupland, 11–35. Oslo: Novus.

Curzan, A. 2014. *Fixing English: Prescriptivism and language history.* Cambridge: Cambridge University Press.

Elspaß, S. 2021. Language standardization in a view 'from below', In *The Cambridge Handbook of Language Standardization*, eds. Ayres-Bennett, W. and J. Bellamy, 93–114. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108559249.004.

Emakeeleselts. 2025. Keeletoimkond. https://www.emakeeleselts.ee/keeletoimkond/. Accessed 8 July 2025.

Erelt, T. 2002. *Eesti keelekorraldus*. Tallinn: Eesti Keele Sihtasutus.

Gadet, F. 2005. France, In *Sociolinguistics/Soziolinguistik: An international handbook of the science of language and society, vol. 3*, eds. Ammon, U., N. Dittmar, K.J. Mattheier, and P. Trudgill, 1787–1792. Oxford: Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.1036.

Gal, S. 2001. Linguistic theories and national images in Nineteenth-century Hungary, In *Languages and publics: The making of authority*, eds. Gal, S. and K.A. Woolard. London: Routledge.

Gal, S. and K.A. Woolard. 2001. Constructing languages and publics: Authority and representation, In *Languages and publics: The making of authority*, eds. Gal, S. and K.A. Woolard, 1–12. London: Routledge.

Kass, R.E. and A.E. Raftery. 1995. Bayes factors. *Journal of the American Statistical Association 90*(430): 773–795. https://doi.org/10.2307/2291091 .

Koppel, K. and J. Kallas. 2022. Eesti keele ühendkorpuste sari 2013–2021: mahukaim eestikeelsete digitekstide kogu. *Eesti Rakenduslingvistika Uhingu Aastaraamat* 18: 207–228. https://doi.org/10.5128/erya18.12 .

Kristiansen, T. 2025. The rise and current status of standard varieties in Germanic languages, In *Oxford research encyclopedia of linguistics*, ed. Aronoff, M. Oxford: Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.1036.

Lindström, L., L. Risberg, and H. Plado. 2023. Language ideologies and beliefs about language in Estonia and Estonian language planning. *Journal of Estonian and Finno-Ugric Linguistics 14*(1): 7–48. https://doi.org/10.12697/jeful.2023.14.1.01 .

Martelli, F., R. Navigli, S. Krek, J. Kallas, P. Gantar, S. Koeva, S. Nimb, B. Sandford Pedersen, S. Olsen, M. Langemets, K. Koppel, T. Üksik, K. Dobrovoljc, R. Ureña-Ruiz, J.L. Sancho-Sánchez, V. Lipp, T. Váradi, A. Győrffy, L. Simon, V. Quochi, M. Monachini, F. Frontini, C. Tiberius, R. Tempelaars, R. Costa, A. Salgado, J. Čibej, and T. Munda. 2022. Parallel sense-annotated corpus ELEXIS-WSD 1.0. Slovenian language resource repository CLARIN.SI.

Milroy, J. and L. Milroy. 2012. *Authority in language: Investigating Standard English, 4ed.* London: Routledge.

Morin, Y.C. 1987. Remarques sur l'organisation de la flexion des verbes français. *ITL: Review of Applied Linguistics* 77–78: 13–91 .

Morin, Y.C. 1995. De l'acquisition de la morphologie: Le cas des verbes morphologiquement défectifs du français, In *Tendances récentes en linguistique française et*

générale: volume dedié à David Gaatone, eds. Bat-Zeev Shyldkrot, H. and L. Kupferman, 295–310. Amsterdam: John Benjamins. https://doi.org/10.1075/lis.20.26mor.

Neustupný, J.V. 2006. Sociolinguistic aspects of social modernization, In *Sociolinguistics: An international handbook of the science of language and society*, eds. Ammon, U., N. Dittmar, K.J. Mattheier, and P. Trudgill, 2209–2223. Berlin: Walter de Gruyter. https://doi.org/10.1515/9783110184181.3.10.2209.

Pertsova, K. 2016. Transderivational relations and paradigm gaps in Russian verbs. *Glossa* 1: 1–13. https://doi.org/10.5334/gjgl.59 .

Raag, R. 2008. *Talurahva keelest riigikeeleks*. Tartu: Atlex.

Raag, V. 1998. *The effects of planned change on Estonian morphology*. Ph. D. thesis, Uppsala University, Uppsala.

Rowe, C., M. Copot, and A.D. Sims. 2026. Reconciling diverging frequency effects in inflectional defectiveness: A study of Peninsular Spanish verbs. *Proceedings of the Linguistic Society of America* 11 .

Saareste, A. 1933. Meie keele korraldamise põhimõtteist ja nende kohaldamisest. *Eesti Kirjandus* 9: 446—-454 .

Samara, A., E. Wonnacott, G. Saxena, R. Maitreyee, J. Fazekas, and B. Ambridge. 2025. Learners restrict their linguistic generalizations using preemption but not entrenchment: Evidence from artificial-language-learning studies with adults and children. *Psychological Review 132*(1): 1–17. https://doi.org/10.1037/rev0000463 .

Schwarz, F. and J. Zehr. 2018. PennController for internet based experiments (IBEX). Open Science Framework. Accessed 15 January 2023.

Sims, A.D. 2015. *Inflectional defectiveness*. Cambridge: Cambridge University Press.

Thornton, A.M. 2011. Overabundance (multiple forms realizing the same cell): A non-canonical phenomenon in Italian verb morphology, In *Morphological autonomy: Perspectives from Romance inflectional morphology*, eds. Maiden, M., J.C. Smith, M. Goldbach, and M.O. Hinzelin, 358–381. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199589982.003.0017.

Thornton, A.M. 2012. Overabundance in Italian verb morphology and its interactions with other non-canonical phenomena, In *Irregularity in morphology (and beyond)*, eds. Stolz, T., H. Otsuka, A. Urdze, and J. van der Auwera, 251–269. Berlin: Akademie. https://doi.org/10.1524/9783050059587.251.

Thornton, A.M. 2019. Overabundance in morphology, In *Oxford Research Encyclopedia of Linguistics*, ed. Aronoff, M. Oxford: Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.554.

Vihman, V.A., M.L. Pilvik, A. Mandel, A. Kängsepp, M. Aigro, K. Koreinik, K. Praakli, and L. Lindström. 2023. Estonian teen language corpus. https://datadoi.ee/handle/33/596; https://doi.org/10.23673/re-455.

Viht, A. and K. Habicht. 2019. *Eesti keele sõnamuutmine [Inflection in Estonian]*. Tartu: Tartu University Press.

Woll, N. 2019. How French speakers reflect on their language: A critical look at the concept of metalinguistic awareness. *Language Awareness 28*(1): 49–73. https://doi.org/10.1080/09658416.2019.1567743 .

Woolard, K.A. 2021. Language ideology, In *The international encyclopedia of linguistic anthropology*, ed. Stanlaw, J., 1–21. Hoboken, NJ: Wiley Blackwell. https://doi.org/10.1002/9781118786093.iela0217.