Are non-native speakers the drivers of morphological simplification? A Wug experiment on the Dutch past tense system

De Smet, Isabeau; Rosseel, Laura; Van de Velde, Freek

Published in:
Journal of Language Evolution

DOI:
10.1093/jole/lzac008

Publication date:
2022

License:
CC BY-NC

Document Version:
Accepted author manuscript

Are non-native speakers the drivers of morphological simplification? A Wug experiment on the Dutch past tense system.

It has often been suggested that there is an inverse correlation between the number of adult non-native speakers in a language and its morphological complexity. Secluded languages often show more complex morphology, while high-contact languages go through more severe simplifications throughout the ages. One such simplification linked to language contact is the regularization of the Germanic past tense. Yet, a Wug task on the English past tense system by Cuskley et al. (2015) showed that non-native speakers tend to use the irregular past tense even more than native speakers. In this paper, we replicate the Wug experiment for Dutch. Our results show similar evidence for a higher rate of irregularization across non-native speakers. Furthermore, we do not find any other simplification strategies among non-native speakers. Though caution is warranted, these converging results may suggest that non-native speakers are not the drivers of morphological simplification.

1. Introduction

Lately more and more researchers have argued that the morphological complexity of languages is co-determined by its number of adult non-native speakers (i.a. Kusters 2003; Miestamo et al. 2008; Trudgill 2004, 2011; McWhorter 2005, 2007). Indeed, a vast amount of SLA research shows that (especially low-proficiency) L2-learners tend to simplify inflectional paradigms (i.a. Oldenkamp 2013; Brezina & Palotti 2019; De Clerq & Housen 2019). As languages adapt to their communities, a large influx of non-native speakers is thought to reduce the morphological complexity of a language. Evidence for this theory can be found in the difference in morphological complexity between a high-contact language such as English and more secluded Germanic languages such as Faroese and Icelandic (Trudgill 2002), but has also been borne out in cross-linguistic studies (Nichols 2009; Sinnemäki 2009) and especially large-scale ones, such as those by Lupyan & Dale (2010) and Bentz & Winter (2013). Lupyan & Dale (2010) found that languages with more speakers tend to show simpler morphological

inflection than languages with fewer speakers. They argue that the size of the speakers' community can be used as a proxy for the amount of L2-speakers, so that the inverse correlation between morphological complexity and speakers' community size reduces to L2-learning. This is confirmed by Bentz & Winter (2013), who provide evidence that languages spoken by a higher proportion of L2-speakers tend to have smaller case systems than languages spoken by relatively few L2-speakers, and that the community size does not add extra explanatory power. Though there are studies that nuance or contradict the precise role of L2-learning (Sinnemäki & Di Garbo 2018; Raviv et al. 2019; Koplenig 2019; Berdicevskis & Semenuks 2020), there is now an abundant amount of literature testifying to the effects of L2-contact on language change.

One particular morphological simplification that has often been linked to language contact is the regularization (or the 'weakening') of the Germanic strong preterite. The weak variant, where a phonotactically predictable dental suffix is added to the stem (e.g. *play-played*) can be said to be simpler than the strong variant where, depending on the coda and vowel of the present stem, a variety of different, partially unpredictable ablaut changes can be used to form the preterite (e.g. *drive-drove*). Furthermore, it can be argued that the weak inflection (with the addition of the dental suffix) is more analytic than the strong inflection and in that way also less complex (Haspelmath & Michaelis 2017; Haspelmath 2018).

Carroll et al. (2012) have argued that peaks and valleys in the regularization rate of the German preterite are likely to be correlated to peaks and valleys in the amount of language contact German has known throughout the course of time. Similarly, De Smet et al. (2017) suggest that the differences in regularization rates between English, German and Dutch are a result of the amount of contact these languages have been exposed to, with English showing both the highest regularization rate and highest degree of language contact, German showing the lowest regularization rate and lowest degree of language contact and Dutch taking an in between position, both morphologically and contact-wise (see also De Smet & Van de Velde 2019). Finally, Dale & Lupyan (2012) show that speakers of American

English (which has been subject to a higher degree of language contact) have a higher preference for regular past tense forms than speakers of British English.

Yet, not all research found evidence that fully supports this trend: in a Wug experiment, a type of experiment in which participants have to inflect a non-word, named after a non-word used in the original experiment by Berko (1958), Cuskley et al. (2015) uncovered a conflicting effect. Their experiment showed that non-native speakers tend to use *more* irregular past tense forms than native speakers of English. Furthermore, when looking at non-native speakers only, their study showed that an increase in self-rated proficiency of English led to a lower proportion of irregular past tense forms. Cuskley et al. (2015) hypothesized that this might be due to the frequency of the irregular past tense category: as irregular verbs are more frequent than regular verbs and non-native speakers encounter mostly these very high-frequent verbs in their input, they tend to over-estimate the proportion of irregular verbs. Native speakers (or non-native speakers with a higher proficiency), by contrast, encounter more of the less frequent regular verbs in their input and therefore estimate the proportion of irregular verbs to be lower. At the same time, Cuskley et al. (2015) did find some indication of morphological simplification among the non-native speakers. Non-native speakers showed a higher preference for the level category (e.g. *hit-hit*) than native speakers, which could be said to be an even simpler past tense formation than the regular *-ed* preterite.

Yet, even though Cuskley et al's (2015) lack of input hypothesis seems plausible and they did find a certain type of simplification tendency among the non-native speakers, their results remain surprising in light of the literature on the effects of language contact on morphological simplification. Hence, the question arises whether the higher rate of irregularization among non-native speakers is a general, cross-linguistic pattern or whether these results are due to a peculiarity in the English irregular verb system. Can we observe the same trends for other Germanic languages? Replication is much needed here, as elsewhere in linguistics (Sönning & Werner 2021), and indeed in science in general (Ritchie 2020).

We set out to replicate the study by Cuskley et al. (2015) by conducting a similar Wug experiment on Dutch. Our study aims to establish whether we can observe the same higher rate of irregularization for non-native Dutch speakers and, if this is indeed the case, whether we can then observe a simplification in the use of the strong strategies in Dutch and how this would manifest itself. In §2 we first present the Dutch preterital system and in §3 discuss our research questions and hypotheses. §4 lays out the methodology of the study including the experimental design, materials, participants, and procedure. In §5 we describe the different steps of our analysis and the results of each step. Discussion of our results and comparison to Cuskley et al. (2015) is the subject of §6. In §7 we review our results in the light of the language contact hypothesis. §8 concludes this paper.

2. Dutch preterital system

In Dutch, like in English and most other Germanic languages, preterites can be formed using either the strong or the weak inflection. The strong inflection is characterized by root-vowel change, also known as 'ablaut' (e.g. *rijd-reed* 'drive-drove'), the weak inflection is characterized by a dental suffix (e.g. *speel-speelde* 'play-played'). Although the strong verbs are not that numerous, they make up the majority of the most-frequent verbs in Dutch, while the much more numerous weak verbs show a much lower frequency (i.a. Van Haeringen 1940; De Smet 2021). Historically, the strong verbs could be divided into 7 ablaut classes according to their coda and vowel pattern. In Dutch, these 7 historical classes are still more or less distinguishable, more so than is the case in English. Class membership today is largely determined by the vowel the verb displays in the present stem. Table 1 shows an overview of the 7 historical ablaut classes and their number of members, i.e. their type frequency (based on the preterites that still show a strong form in Haeseryn et al. 1997). This type frequency is an important predictor of the productivity of the strong verb patterns: classes with a higher type frequency tend to hold on to the strong inflection longer and tend to attract more new members than classes with a lower type frequency (Knooihuizen & Strik 2014; Bybee 1995).

Table 1. The seven historical ablaut classes[1]

<insert Table 1 here>

Though historically, there were different classes of weak verbs as well, they have collapsed in one overall class in Dutch. Weak forms uniformly take a dental suffix, which may be realised as /tə/ or /də/, depending on the coda. Yet, there are a few weak verbs that, as a result of sound changes, show other changes in the preterite than the addition of a dental suffix, e.g. *zeg-zei* 'say-said' or *zoek-zocht* 'search-searched'. Often both these aberrant weak verbs and the strong verbs are labelled as 'irregular', as opposed to all weak verbs that only show a dental suffix (e.g. Haeseryn et al. 1997) which are called 'regular'. While this terminology is not unique to the Dutch preterital system (cf. English), we prefer not to use it: not only does the label 'irregular' cover two historically different groups of verbs, the strong verbs, especially in Dutch, can hardly be called 'irregular'. We can distinguish very clear patterns in Dutch, which are still relatively productive (as shown by Knooihuizen & Strik [2014] in a Wug experiment). Instead, we will use the more neutral terminology 'strong' versus 'weak' verbs and reserve the term 'irregular' for strong verbs that do not only show ablaut (e.g. *houd-hield* 'hold-held') and weak verbs that do not only show a dental suffix (e.g. *zoek-zocht* 'search-searched').

Although Table 1 presents many different patterns, recent research has shown that Dutch language users show a preference for one particular vowel to form the preterite: /o/ <oo>. In a Wug experiment by Knooihuizen & Strik (2014), vowel patterns that usually call for an /o/ <oo> in the preterite showed very high productivity. Furthermore, /o/ <oo> was even chosen for verbs that did not belong in any of the existing vowel patterns (e.g. *beuk-book*). Additionally, Nowak (2010) has

[1] A few strong verbs can no longer easily be categorized in one of the historical ablaut classes. These are *schep-schiep* 'create-created', *hef-hief* 'lift-lifted', *zweer-zwoer* 'swear-sweared', *tijg-toog* 'go-went', *kom-kwam* 'come-came', *word-werd* 'become-became', *ben-was* 'am-was', *doe-deed* 'do-did', *ga-ging* 'go-went', *sta-stond* 'stand-stood' and *zie-zag* 'see-saw'.

shown that multiple verbs in the history of Dutch have taken this increasingly popular preterite vowel (e.g. *weeg-woog* < *weeg-wach* 'weigh-weighed', *zweer-zwoor* < *zweer-zwar* 'fester-festered', *scheer-schoor* < *scheer-schar* 'shave-shaved'). In non-standard language this vowel is on the rise as well (e.g. *breek-brook* < *breek-brak* 'break-broke', *steel-stool* < *steel-stal* 'steal-stole' etc.), as De Smet (2021) points out. Knooihuizen & Strik (2014) suggest /o/ has become a sort of default vowel for the strong preterites in Dutch. The adoption of /o/ <oo> in the preterite holds a reduction in complexity in the Dutch verb system, both intra- and interparadigmatically. Verbs adopting an /o/ in the preterite, usually already showcased an /o/ in the past participle (e.g. *steel-stool-gestolen* 'steal-stole-stolen', *breek-brook-gebroken* 'break-broke-broken'). Intraparadigmatically, the adoption of /o/ <oo> in the preterite thus means a reduction in the number of different vowels in the verb paradigm from three (a different vowel in present, preterite and past participle, e.g. *breek-brak-gebroken* 'break-broke-broken') to two (a different vowel in the present on the one hand and the preterite and past participle on the other, e.g. *breek-brook-gebroken* 'break-broke-broken'). Interparadigmatically, the number of different strong categories decreases as the category with an /ɑ/-preterite and /o/-past participle (i.e. class IV) joins the category with an /o/-preterite and /o/-past participle (i.e. class II, e.g. *ruiken-rook-geroken* 'smell-smelled-smelled', cf. Table 1).

3. Research questions

The aim of this paper is to establish whether a higher rate of strong forms in past tense morphology among non-native speakers is a general pattern across Germanic languages. Based on our discussion of the effects of language contact on morphology (cf. §1), this would be unexpected. As the weak preterite could be said to be the 'simpler' variant, we would expect non-native speakers to use more weak forms than native speakers. Yet, as mentioned, Cuskley et al. (2015) did find a higher rate of irregularization across non-native speaker. Can the same trend be uncovered for Dutch? To find out, we will largely ask the same research questions as in Cuskley et al. (2015). The key research question is:

- Do native speakers and non-native speakers differ in their choice for the weak versus the strong variant when forming a preterite in a Wug task?

A second research question we aim to answer through this experiment, again following Cuskley et al. (2015) is:

- Which factors influence non-native speakers' choice for the weak versus the strong variant when forming a preterite in a Wug task?

More specifically, differences between speakers in age of acquisition of Dutch, self-rated proficiency of Dutch and the proficiency in another Germanic language will be considered. Starting from the assumption that non-native speakers choose the 'simpler' variant, one would expect this effect to increase as the age of acquisition of Dutch increases or as the proficiency of Dutch decreases. However, Cuskley et al. (2015) again found the reverse trend in their experimental data. With regards to a high proficiency in another Germanic language, one would expect this to boost the probability of the use of the strong variant, as such proficiency may entail more experience with the use of ablaut.

Finally, we aim to shed light on the use of strong preterite formation strategies in native versus non-native speakers. Our third research question then is:

- Do native and non-native speakers differ in their choice of strong preterite formation strategies when forming the preterite in a Wug-task?

We would expect native speakers to be better acquainted with the existing strong verb patterns in Dutch and therefore to diverge less from existing verb patterns than non-native speakers. If non-native speakers indeed diverge from the existing verb patterns, we would expect them to do this in a way that simplifies the morphology of the verbal system. A drastic reduction in the complexity of the verb system was also observed in Cuskley et al.'s (2015) experiment where non-native speakers showed a preference for the level category (e.g. *hit-hit*). In Dutch, however, the level category is not a grammatically sanctioned strategy for simplification. Yet, another possible strategy for simplification can come to mind: the extrapolation of the /o/ <oo> preterite. Not only is this a highly frequent vowel

among the strong verbs and one which has lately been on the rise in Dutch (like the level category is in English), the adoption of /o/ <oo> also reduces the number of strong verb categories interparadigmatically and the number of different vowels intraparadigmatically, thus simplifying the Dutch verbal system (cf. §2).

4. Method

4.1. Design & materials

Our study consisted of a Wug experiment in which participants were asked to form the preterite of a non-verb. Each participant was randomly assigned 32 trials (out of a total of 60, one for each non-verb). For each non-verb, they were given the prompt *Nu [non-verb] ik bijna nooit* 'Now I [non-verb] almost never' and were asked to fill in the blank in the following sentence *Vroeger … ik elke dag* 'Before I … every day'. An example of the Wug task can be found in Figure 1.

<insert Figure 1 here>

Figure 1. Example Wug task. Participants were asked to fill in the blank using the preterite of the non-verb. The non-verb was given in the infinitive and used in the present tense in an short sentence as an example.

Following Cuskley et al. (2015), we constructed our non-verb stimuli using the existing Dutch strong verbs (showing a strong preterite) and the 500 most frequent Dutch weak verbs (extracted from SUBTLEX-NL, Keuleers et al. 2010) as a basis. From this list, we excluded multisyllabic stems (e.g. *begin* 'begin'), stems with a coda ending in a vowel (e.g. *ga* 'go') and loans (e.g. *douch* 'shower'). However, unlike Cuskley et al. (2015), we did not generate an exhaustive list of non-verbs with all possible onset-vowel-coda combinations as this rendered an unworkable amount of output. Instead we opted to

follow Knooihuizen & Strik's (2014) method who also conducted a Wug experiment for Dutch and generated their stimuli with Wuggy (Keuleers & Brysbaert 2010). Wuggy is a non-word generator that takes into account which subsyllabic combinations are possible in a language. The generator was used to create one pseudoword for each existing word on our list of existing strong and weak verbs, mimicking their syllabic structure (same length of subsyllabic elements and same transition frequency between subsyllabic elements).

As the vowel of the verb stem is known to play a major role in the choice between the strong or weak inflection in Dutch (Knooihuizen & Strik 2014; De Smet 2021), we decided to divide the verbs into different vowel groups, one for each vowel that occurs among the Dutch verbs (i.e. /ɛɪ/ <ij>,[2] /i/ <ie>, /œy/ <ui>, /ɛ/ <e>, /ɪ/ <i>, /e/ <ee> or <e>, /o/ <oo> or <o>, /a/ <aa> or <a>, /ɑ/ <a>, /ɔ/ <o>, /ø/ <eu>, /ʏ/ <uu> or <u>, /ʏ/ <u>, /u/ <oe>, /ɑʊ/ <ou>) rather than into groups based on the distance between the non-verbs and the closest existing verb (as in Cuskley et al. 2015). The distances calculated in Cuskley et al. (2015) give the same weight to changes in onset, coda and vowel, while changes in onset or coda are usually less important in Dutch than changes in vowel for the choice between the weak or strong inflection (Knooihuizen & Strik 2014; De Smet 2021). For each vowel group eight verbs were chosen from the Wuggy output, excluding verbs that turned out to exist in non-standard varieties of Dutch. For some vowels Wuggy did not generate the required eight verbs. In those cases, additional non-verbs were created manually, using onset and coda combinations from verbs with different vowels from the Wuggy output that were not already chosen for a different vowel group. For /ɑʊ/ <ou>, all stimuli had to be constructed manually. Dutch only has verbs where this vowel is followed by either /d/ or /w/ (e.g. *houd* 'hold' or *houw* 'hew').

[2] Verbs with the <ei> orthography, which also corresponds to /ɛɪ/, were not included, because it is unclear whether these verbs would behave in the same way as verbs with the <ij> orthography. The same goes for the orthography of the diphthong /ɑʊ/: only verbs with <ou> were included, not verbs with the alternative corresponding spelling <au>.

In a pretest, the 120 non-verbs resulting from the procedure described above (i.e. 8 verbs for each of the 15 vowel groups) were submitted to an association task. Participants (*N* = 81) were randomly presented with 40 non-verbs and were asked with which existing words they associated the non-verbs. The aim of the pretest was to diminish the chances that participants would associate the non-verb too strongly with an existing verb and as a result would choose the preterite of the non-verb based on the preterite of the existing verb. Based on the pretest, 4 verbs of each vowel group that were least associated with one and the same existing verb (and never by more than half of the participants) were included in the final stimulus set. The final selection of verbs can be found in Table 2 which includes the type frequency for each vowel group (based on the number of strong verbs with the same vowel in the present stem in Haeseryn et al. 1997)[3] and the expected preterite vowel (based on the existing vowel patterns).[4]

Table 2. Non-verbs used in the experiment

<insert Table 2 here>

[3] As the choice for the ablaut vowel in the preterite is not fully predictable from the vowel in the present stem, class membership of nonce verbs is not strictly determined, type frequency is calculated by counting all strong verbs with the same present stem vowel and not by counting the number of members in one of the historical ablaut classes.

[4] Like in Cuskley et al. (2015) we did not add control items (items with a certain guaranteed answer which could be used to identify participants who are not paying attention or are answering randomly, if they answer differently) to our experiment. There are no straightforward items similar to our target items for which we know for sure how an L2-speaker would respond. Even existing strong verbs can be used with the weak inflection or existing weak verbs with the strong inflection. Though there are verbs where the strong inflection would be very unlikely from an L1-perspective, we do not know whether this is the case for L2-speakers as well.

4.2. Participants

The study includes both a sample of non-native and native speakers. Unlike Cuskley et al. (2015) we decided to focus on one homogenous group of non-native speakers: French-speaking Belgians who were likely to have a fair proficiency in Dutch. This decision was made for several reasons. As we know that participants' L1 can influence their L2 skills (Schepens et al. 2013, Van der Slik et al. 2019), we must expect it will affect their use of past tense morphology in Dutch. Therefore we need to take L1 into account in this study, either by studying its influence or by controlling for it. As the position of Dutch as a language in the world can hardly be compared to that of English as the global lingua franca, the latter was the preferred or at least the most feasible option: the number and diversity of Dutch L2-speakers is considerably smaller than that of English L2-speakers which makes participant recruitment more challenging. Participant recruitment through online platforms like Cuskley and colleagues (2015) did, is not an option when looking for a decent sized sample of Dutch L2-speakers. As a result we had to resort to alternative methods of recruitment. The two largest groups of L2-speakers we had access to were French and German learners of Dutch. Given the linguistic similarities between Dutch and German preterite formation, it was decided that French learners of Dutch would provide a more suitable group to include in the study, even though we have to be aware that the many irregular verb paradigms in French could influence the participants' answers in this ask. The reason we avoided recruiting very low proficiency learners is that participants needed to be proficient enough to be able to form the preterite of a verb in order to carry out the Wug task.

These non-natives speakers were recruited among teachers and professors of Dutch in Wallonia, the French speaking southern half of Belgium, university and college students of Dutch in Wallonia and in the personal and professional networks of the authors. We asked specifically for non-native speakers (no bilinguals) with a fair proficiency in Dutch. Most participants learned Dutch in high school, usually from the age of 12 and onwards. Native speakers were recruited through social media

and Dutch speaking Belgian universities. In total 551 participants finished the experiment. 5 participants were excluded because they spoke another native language than Dutch or French. Hence, a total of 546 participants was retained in the sample, of which 442 were native speakers of Dutch and 104 were non-native speakers. Of the native speakers 276 identified as women, 160 as men and 6 as 'other'. Of the non-native speakers 78 identified as women and 27 as men. Although a chi-square test (excluding the speakers identifying as 'other') shows a significant difference between the distribution in gender in both subsamples (χ^2 = 4.513, *df* = 1, *p* = 0.034) and research suggests a discrepancy in proficiency between male and female L2-speakers (Van der Slik et al. 2015), adding gender to our models did not have an impact on the results. Ages of native speakers ranged between 18 and 80 (*SD* = 15.054, *M* = 32.34), ages of non-native speakers ranged between 18 and 67 (*SD* = 13.622, *M* =33.100). A two-tailed *t*-test (*t* = -0.477, *df* = 545, *p* = 0.634) shows there is no significant difference in age distribution among the two subsamples and a Cohen's D (= 0.05) shows the effect size is minimal.

4.3. Procedure

The experiment was conducted using the online software tool Qualtrics. Participants did not receive monetary compensation. The experimental procedure was approved by the KU Leuven Ethics Committee. All questions and instructions were in Dutch and proofread by non-native speakers to make sure the task was explained clearly enough. Each participant was first presented with some demographic questions and questions about their language proficiency. They were asked to give their gender, age, education, native language(s), up to two other languages they spoke best besides Dutch and their native language, and their self-rated proficiency (on a scale of 0-100) and self-rated age of acquisition of these languages. Non-native speakers of Dutch also had to fill in where they learnt Dutch (in school, in contact with family, in contact with friends or at work), their age of acquisition and their self-rated proficiency of Dutch (on a scale of 0-100).

After these questions, the actual experiment began. Participants were explicitly instructed to give the past tense (we used the Dutch terms *onvoltooid verleden tijd* and *imperfectum*) of the non-verb. They were also told the verb did not exist in Dutch and that the meaning of the verb therefore did not matter. Two examples of the Wug-task were given, one with the strong and one with the weak inflection filled in. Participants were told to act intuitively and that no right or wrong answers existed. Finally, after completing all sentences, they were asked to explain which strategy they had used in the task.

5. Analysis & results

5.1. Overview data

The answers given by the participants were often not as straightforwardly 'strong' or 'weak', but rather fell on a continuum, with on the far-left side preterites without a marker (cf. the level category in English) and on the far-right side preterites with both a strong marker and a weak marker and in between variations on this and of course the regular strong and weak inflection itself. In our analysis, we will focus solely on cases that could be categorized straightforwardly as strong or weak, which make up the clear majority. All verb forms that showed only a vowel change were coded as 'strong'. All attestations that only showed an added dental suffix were coded as 'weak'. The following suffixes were assessed as dental suffixes: *-de, -te, -ed* (either a typo or a suffix influenced by the English weak preterites), *-ede, -ete* and *-ege* (weak preterite suffix variants used in East- and West-Flemish varieties of Dutch). Small typos were ignored (i.e. *blluipte* as preterite of *bluip*). We disregarded cases where:

- the participant gave both a strong and weak variant (24 cases).
- the preterite showed both a vowel change and a dental suffix (e.g. *treende* < *truin*) (231 cases).
- the preterite showed no change (e.g. *truin* < *truin*) (37 cases). Unlike in English the level category is not an existing strategy in Dutch (cf. Cuskley et al. 2015: 114 fn. 6).

- the preterite showed a suffix *-ste/-sde* (e.g. *huikste* < *huik*) (7 cases). This suffix originally belongs to the preterite-presents (verbs with a present that diachronically goes back to a preterite) and is still used in some dialects to form irregular past tense forms (e.g. *bego(n)st* < *begin* 'begin'). Therefore it is difficult to categorize these forms as regular weak forms.
- the preterite showed a suffix *-t/-d* (e.g. *parpt* < *parp*) (13 cases).
- only a schwa was added but no dental (e.g. *niekke* < *niek*) (25 cases).
- the verb form was an existing preterite (e.g. *spoot* 'squirted' < *spuis*) (57 cases). In these cases we cannot know for sure whether the participant had the existing verb in mind or the non-verb.
- the preterite is unrecognizable as a verb form from the given non-verb (either because the participant had misread the prompt, made a serious typo or for unknown reasons, e.g. *spots* < *spuis, trook* < *truin, kuikte* < *huik*) (173 cases).

In total 567 attestations were excluded. This left us with a total of 16,915 attestations of which 4,259 were strong and 12,656 were weak. This means that overall in 25.18% of the cases a strong inflection was chosen.[5] Figure 2 shows the distribution of the weak and strong inflection over the vowel groups. Vowel groups are ordered by increasing type frequency. Except for /y/ <uu> on the left-hand size of the graph, it can be observed that the vowel groups on the right-hand side (with a higher type frequency) show more strong forms overall. Verbs with /ɛɪ/ <ij> even show more strong than weak forms.

<insert Figure 2 here>

[5] This is less than in Knooihuizen & Strik (2014) who observed a strong inflection in one third of the cases. However, in their experiment they used more nonce verbs with a vowel that occurs among the strong verbs, making their nonce verbs more likely to be used with the strong inflection.

Figure 2. Distribution strong and weak inflection for each vowel group. Vowel groups are ordered by increasing type frequency. Overall, it can be observed that the vowel groups on the right-hand side (with a higher type frequency) show more strong forms.

These results will be analysed in three steps.[6,7] In a first step we will compare the distribution of strong and weak verb forms for native versus non-native speakers. Do non-native speakers prefer weak preterite morphology or do we observe an opposite trend as in Cuskley et al. (2015)? In a second step we will focus on non-native speakers specifically and the effect of proficiency and age of acquisition of Dutch (and other Germanic languages) on the choice for the weak or strong inflection. Finally, in a third step, we will zoom in on the differences in strong strategies used by native versus non-native speakers.

5.2. The use of strong vs. weak preterites in native and non-native speakers

Overall, the data show that native speakers use the strong inflection in 24.23% of the cases, while non-native speakers use the strong inflection in 29.24% of cases. Though this difference is much smaller than the difference reported in Cuskley et al. (2015) (i.e. respectively 21.6% vs. 35.7%), a mixed effects logistic regression does show that this difference is significant, and goes in the same direction. The mixed effects regression was run with INFLECTION as the outcome variable (with 'strong' as default level) and NATIVENESS as a factor of interest, as well as an interaction with TYPE FREQUENCY (centered and standardized, original values can be found in Table 2). As covariates TRIAL NUMBER, PRIMING and TYPE

[6] For our data transformation, analysis and visualisations, we made use of R (R Core Team 2017) and more specifically of the packages dplyr (Wickham et al. 2019), reshape (Wickham 2017), effects (Fox 2013), ggplot2 (Wickham 2016), MuMIn (Barton 2019), vcd (Meyer et al. 2020) and ModelMetrics (Hunt 2018).

[7] Our dataset and the code can be found on the lead author's github page:
https://github.com/IsabeauDeSmet/replication-Cuskleyetal2015

FREQUENCY were included. TRIAL NUMBER is a centered and standardized numerical variable added to account for effects of fatigue or practice. PRIMING is a categorical variable conveying for each participant the answer to the previous trial, in case this primes the participant for a certain answer in the next question. Values could either be 'strong', 'weak' or 'no priming' (for the first trial or when the previous answer of a participant had to be disregarded). As random structure a random intercept for SUBJECT and a random intercept for ITEM were added, as well as a correlated random slope for NATIVENESS by ITEM. As the model with a correlated random slope for NATIVENESS by ITEM did not converge, we downsized the model by switching to an uncorrelated random slope which did not make the model significantly worse and ensured convergence.

Table A in the appendix shows the numerical output of the fixed effects, Table B of the random effects. The model had a C-value of 0.857 (values above 0.8 indicate a good fit), a marginal R^2 of 0.082 (variation explained by fixed effects only) and a conditional R^2 of 0.377 (variation explained by both fixed and random effects). There were no problems with multicollinearity (VIF-scores were all below 2).[8] Figure 3 shows the interaction between TYPE FREQUENCY and NATIVENESS. First of all, it becomes clear once more that native speakers use fewer strong preterites than non-natives speakers. Second, we can see that a higher type frequency leads to a higher probability of strong preterites, but that this effect is more outspoken for non-native speakers. Yet, it seems to be the case that the effect is mostly due to a number of verbs with a very high type frequency, namely the verbs with /ɛɪ/ <ij> in the present (class I). Indeed, when excluding verbs of this class from the model, the interaction fails to reach significance.

<insert Figure 3 here>

[8] VIF-scores were calculated by the technique used by Zuur et al. (2009) for mixed effect regression models, values above 5 suggest multicollinearity (Levshina 2015: 160).

Figure 3. Interaction NATIVENESS and TYPE FREQUENCY in model 1. Native speakers use fewer strong preterites than non-native speakers and a higher type frequency leads to a higher probability of strong preterites.

In a second model, to discover other, more fine-grained patterns besides the effect of type frequency in the distribution of the strong and weak inflection for native versus non-native speakers, we included a more detailed interaction between PRESENT STEM VOWEL and NATIVENESS. We use exactly the same fixed and random factors as in the first model, only PRESENT STEM VOWEL instead of TYPE FREQUENCY (again an uncorrelated random slope was selected over a correlated random slope in the model selection). Because PRESENT STEM VOWEL has 15 levels, this variable prevented the model from converging without using an optimizer bobyqa ('better optimalisation by quadratic approximation'). The numerical output of the model can be found in Table C and D in the appendix. The C-value was 0.856, marginal R^2 was 0.175 and conditional R^2 was 0.434. The model did not suffer any problems with multicollinearity (VIF-scores were all below 2).

Figure 4 shows the interaction between PRESENT STEM VOWEL and NATIVENESS. Again the vowel groups are ordered by increasing type frequency. It can be seen that for almost every vowel group, non-native speakers use the strong inflection more often than native speakers. Only /u/ <oe>, /y/ <uu> and /a/ <aa> show the opposite trend. These deviations are all located towards the left-hand side of Figure 4, among the vowel groups with the lower type frequencies. On the righthand side of the Figure, among the vowel groups with higher type frequencies, differences between native and non-native speakers appear to be larger, especially for the /ɛɪ/ <ij> verbs.

<insert Figure 4 here>

Figure 4. Interaction NATIVENESS and PRESENT STEM VOWEL in model 2. Vowel groups are ordered by increasing type frequency. Overall, non-native speakers tend to use the strong inflection more often than native speakers. Deviations from this trend are located among the vowel groups with lower type frequencies (left-hand size).

5.3. Non-native speakers

In a second step we focused the analysis on non-native speakers only. The model built for this analysis contained 3201 attestations. Again INFLECTION (with 'strong' as default level) was the outcome variable and TRIAL NUMBER, PRIMING and TYPE FREQUENCY were included as covariates as well as random intercepts for SUBJECT and ITEM. As factors of interest we looked at SELF-RATED PROFICIENCY (centered and standardized), AGE OF ACQUISITION (centered and standardized), HIGH PROFICIENCY OF GERMAN AS L2[9] and HIGH PROFICIENCY OF ENGLISH AS L2. These factors of interest were also added as correlated random slopes by ITEM. Only the correlated random slope for SELF-RATED PROFICIENCY by ITEM made the model significantly better, as well as the fixed effect SELF-RATED PROFICIENCY. Factors of interest and random slopes that did not make the model significantly better, were removed. The model had a C-value of 0.898, a marginal R^2 of 0.117 and a conditional R^2 of 0.579. The output can be found in Table E and Table F in the appendix. No problems with multicollinearity occurred (VIF-scores were all below 2). As in Cuskley et al. (2015) a higher proficiency seems to lead to a higher probability of weak forms, as can be seen in Figure 5.

<insert Figure 5 here>

[9] Self-rated proficiency was higher than 60 (on a scale from 0-100).

Figure 5. Effect SELF-RATED PROFICIENCY in model 3 (non-native speakers only). A higher proficiency leads to a higher probability of weak forms.

5.4. Strong strategies used by non-native speakers

Finally, we zoom in on the strong strategies used by native versus non-native speakers. First of all, when comparing the number of strong strategies used by native versus non-native speakers, Figure 6 shows that both native and non-native speakers peak at a number of 3 strong strategies per participant, although among non-native speakers there are some outliers using 7, 8 or 9 strategies. Yet, it is important to keep in mind that native speakers use fewer strong forms in total and thus have less opportunity to vary in their strong strategies.

<insert Figure 6 here>

Figure 6. Number of strong strategies used by native versus non-native speakers. Both groups peak at a number of 3 strong strategies per participant, though the non-native speaker group shows several outliers.

To account for this, we calculated an entropy measure (S) for each participant (*j*), following Cuskley et al. (2015: 215): "Given n_j as the total number of irregularizations for a participant, we define p_i^j as the

fraction of irregularizations adhering each sub-rule *i* adopted by the participant. Given this, S_j is defined as:

(1)

$$S_j = \frac{-\sum_{i=1}^{n_j} p_i^j \log_2 p_i^j}{\log_2 n_j}$$

[…]". The maximal value of S_j is acquired when $p_i^j = 1/n_j$ for all the adopted sub-rules. The denominator $\log_2 n_j$ ensures the value is bound between 0 and 1 for each participant *j*. In this way, the normalised quantity S_j provides a value for each participant ranging between 0 and 1. An S_j measure of 0 means only one strategy is used across all strong forms. An S_j measure closer to 1 means more diverse strategies are used. Figure 7 shows the S_j value for each participant on the y-axis and the total number of strong forms on the x-axis. A loess curve is fitted for both native and non-native participants.[10] This Figure shows us that non-native speakers tend to use more different ablaut strategies than native speakers, controlled for their overall use of strong verbs, as their S_j values are usually higher.

<insert Figure 7 here>

Figure 7. S_j values against total number of strong verbs for each participant. Loess curves are fitted for both native and non-native speakers. Non-native speakers use more different ablaut strategies than native speakers.

So far we have seen that non-native speakers use more strong forms than native speakers and use a more diverse range of strategies. A question that remains unanswered is whether non-native speakers are as acquainted with the existing Dutch strong verb patterns or whether they improvise and invent new rules, diverging from existing patterns. In a new model, we therefore took a look at whether or

[10] A logistic regression model with NATIVENESS as outcome variable and the S_j MEASURE and the NUMBER OF IRREGULAR VERBS as fixed effects confirms that S_j values are significantly higher for non-native speakers.

not native and non-native speakers differ in their use of 'unexpected strategies'. Only strong attestations (*N* = 4259) are retained. A strategy was coded as 'expected' when the vowel pattern occurs among existing strong verbs and as 'unexpected' when it does not.[11] This served as an outcome variable (with 'expected' as default level). NATIVENESS was the factor of interest. Again TRIAL NUMBER, PRIMING and TYPE FREQUENCY were added as control factors and a correlated random slope for NATIVENESS by ITEM and an intercept for SUBJECT were added as random structure. The C-value was 0.896, marginal R^2 was 0.244 and conditional R^2 0.678. Table G and H in the appendix show the output. No problems with multicollinearity were attested (VIF-scores were all below 2). NATIVENESS did not turn out to have a significant effect on the expectedness of the vowel used in the Wug task. Native and non-native speakers seem to follow existing vowel patterns to more or less the same degree.

Still, in more than one third of cases both native and non-native speakers deviate from existing vowel patterns. The next question then is what happens when native and non-native speakers do diverge. Is it in these deviations that non-native speakers bring simplification to the Dutch strong verb system? Figure 8 shows the vowels used by respondents when we only take into account answers that diverge from the existing vowel patterns (*N* = 1736). Two large differences between native and non-native speakers stand out in this figure. The first one can be found in the 'other' category. This category contains all vowels that do not occur in the preterite singular of existing strong verbs (/a/ <aa>, /ø/ <eu>, /ɛɪ/ <ij>, /ɑʊ/ <ou>, /ʏ/ <u>, /œy/ <ui> and /y/ <uu>) or only in a very limited number of verbs (i.e. /ɛ/ <e> in *word-werd* 'become-became' and /ɪ/ <i> in *hang-hing* 'hang-hung' and *vang-ving* 'catch-caught'). Non-native speakers make use of one of these vowels in 12.7% of their answers, while native speakers only make use of them in 3.3% of their answers. This suggests that, although native and non-

[11] Strong preterites of verbs with a present stem vowel that does not occur among the existing strong verbs (i.e. /ø/ <eu> and /ʏ/ <u>) were always counted as 'unexpected'.

native speakers do not differ in the number of times they diverge from existing verb patterns, when they do diverge, non-native speakers more often 'invent' new strategies.[12]

A second notable difference can be found in the use of the /o/-preterite. Figure 7 suggests that /o/ <oo> is the most popular vowel among both native and non-native speakers, which confirms the results of earlier studies on the 'default' status of /o/ <oo> (Knooihuizen & Strik 2014; Nowak 2010; De Smet 2021). However, there is quite a large discrepancy between both groups: native speakers use /o/ <oo> in 33.2% of the cases (almost twice as much as the next most used vowel, /e/ <ee>), while non-native speakers use this vowel only in 20.5% of the cases.

<insert Figure 8 here>

Figure 8. Unexpected vowels used in the preterite by native versus non-native speakers. /o/ <oo> is the most popular vowel among both native and non-native speakers, though the difference with the other vowels is much larger among native speakers.

To confirm this higher proportion of the use of the /o/ <oo>-preterite among native speakers, we ran a final mixed effects logistic regression model with /o/ <oo> as the outcome variable (default level is 'no /o/ <oo>'). All strong attestations were included in this model (*N* = 4259). The same control factors and random intercepts were added, as well as an uncorrelated random slope for NATIVENESS by ITEM (again a correlated slope did not significantly improve the model). The factor of interest is NATIVENESS once more. In order for the model to converge, an optimizer 'bobyqa' had to be added. The C-value is 0.958. Marginal R^2 is 0.026 and conditional R^2 is 0.460. The output can be found in Table I and J in the

[12] Related to this, we observed that the cases we dismissed as not making use of either the strong or the weak variant (see 5.1. Overview data) came significantly more often from non-native speakers than from native speakers.

appendix. No problems with multicollinearity occurred (VIF-scores were all below 2). Results show that non-native speakers indeed tend to use /o/ <oo> less than native-speakers.

Ideally, we would also add an interaction with PRESENT STEM VOWEL to the model to map where exactly native and non-native speakers differ in their use of /o/ <oo>, but this model could not converge (mainly because for some present stem vowels, e.g. /o/ <oo>, /o/ <oo> never shows up in the preterite). However, we can show plain proportions of the use of /o/ <oo> for each vowel for both native and non-native speakers in Figure 9. This figure confirms that the higher proportion of /o/ <oo> among native speakers can be observed for almost every vowel. Only for the present stem vowels which always occur with /o/ <oo> in the preterite (i.e. /i/ <ie> and /œy/ <ui>), we barely see a difference for native and non-native speakers.

<insert Figure 9 here>

Figure 9. Proportion /o/ preterites for each present stem vowel for native and non-native speakers. Present stem vowels are ordered by increasing type frequency. For almost every vowel a higher proportion of /o/ can be observed among native speakers.

6. Discussion results

To sum up, our experiment shows first of all that non-native speakers use more strong preterites in Dutch than native speakers and thus confirms Cuskley et al.'s (2015) findings on English. Not only do we see a difference between native and non-native speakers, but among non-native speakers we also observed that a higher proficiency leads to a decrease in the use of strong forms, which is again in line with the results reported by Cuskley and colleagues. The fact that the difference in proportion of weak forms is not as large as in Cuskley et al. (2015) may be due to the overall higher proficiency of our non-

native speakers as well.[13] Cuskley et al.'s (2015) explanation with regards to the frequency of strong verbs in the input of non-native speakers holds for Dutch too. Non-native speakers encounter a lower proportion of the less frequent weak verbs and therefore over-estimate the proportion of strong verbs in the language. The interaction with proficiency can be explained in the same way. Less proficient non-native speakers will come across even fewer of the less frequent weak verbs than more proficient non-native speakers and therefore over-estimate the proportion of strong verbs even more.

The difference between native and non-native speakers is largest for the /ɛɪ/ <ij> present stems. It would appear that this is a salient strong verb pattern, and consequently easier for the non-native speakers to memorize. Not only does this pattern show the highest type frequency, it also shows a high cue reliability: /ɛɪ/ <ij> verbs only ever show up with an /e/ <ee> preterite and at the same time the /e/ <ee> preterite only ever shows up for /ɛɪ/ <ij> verbs among the existing strong verbs. These advantages of course also hold for native speakers. Yet, while for most native speakers these advantages only play a role subconsciously, non-native speakers are explicitly made aware of them in L2-education by teachers who try to provide their students with practical guidance and mnemonic aids to bring some order in what at first sight may seem an irregular mess to the learner. As a result, some kind of 'Matthew effect'[14] emerges: the cognitive advantages of the /ɛɪ-e/ <ij-ee> pattern leads to a high emphasis on this pattern in L2-education, which makes this pattern even more salient for L2-learners. Though this effect of education probably most clearly plays a role for the /ɛɪ-e/ <ij-ee> pattern, it might also be part of the explanation for the overall higher rate of strong forms of non-native speakers, as was also suggested in Cuskley et al. (2015). In general a greater emphasis is put on

[13] A second factor that might add to the smaller difference between non-native and native speakers in Dutch is the more 'shattered' strong (or irregular) verb system in English. The higher degree of regularity and therefore the higher type frequency of many of the vowel patterns in Dutch probably leads to a higher productivity of the strong verb patterns in Dutch, making it a more plausible strategy to apply to nonce verbs.

[14] An ubiquitous effect in natural sciences and human culture describing the phenomenon that the rich get richer and the potent more potent or more in general that advantage creates advantage (Perc 2014).

the strong verbs in L2-education because of their greater difficulty, and they are often drilled by rote learning, which again could lead to an over-estimation of strong verb forms by non-native speakers. Furthermore, the school-like nature of the Wug task may also have led to some form of the observer's paradox, and to a social desirability bias: L2-speakers might feel eager to display their acquaintance with the Dutch strong verb system causing them to use more strong forms.

L2-speakers do indeed seem to know the Dutch strong verb system quite well. When we compare the strong strategies used by native and non-native speakers, it turns out that non-native speakers diverge no more from existing verb patterns than native speakers. Furthermore, an entropy measure shows that non-native speakers use a more diverse range of ablaut strategies: it is thus not the case that non-native speakers tend to simplify the verb system more than native speakers by extrapolating a limited number of strategies to all different patterns. Even though non-native speakers do not diverge from existing verb patterns more often, when they do diverge, they tend to use very rare strategies more frequently than native speakers.

A morphological simplification of the Dutch strong verb system by non-native speakers cannot be found in the use of the /o/ <oo> preterite either, as was hypothesized. Though the popularity of /o/ <oo> is again confirmed, non-native speakers do not tend to make use of this vowel quite as often as native speakers, even though the adoption and extrapolation of /o/ <oo> to other verb patterns could be argued to entail a simplification of the Dutch verb system. This could mean that the rise of /o/ <oo> as a default strategy to form the preterite in Dutch is a very subtle change that is much more difficult for non-native speakers to pick up on. Though the rise of /o/ <oo> has, throughout the history of Dutch, permeated in some cases into the standard language (e.g. *weeg-woog* 'weigh-weighed'*, zweer-zwoor* 'fester-festered'), in most cases the extrapolation of /o/ <oo> still clearly belongs to the informal language (e.g. *breek-brook* 'break-broke'*, steel-stool* 'steal-stole' etc.). Being largely confined to informal registers, the /o/ <oo> preterite might be hiding in the thicket of sociolinguistic contexts that remain elusive to non-native speakers: non-native speakers will not easily come across these preterites in average language education settings, and the advantages of the /o/ <oo> preterite will remain out

of reach, as opposed to the low-hanging fruit of the /ɛɪ-e/ <ij-ee> pattern, which is fully standard, highly salient, and is explicitly drawn attention to in language education. The lack of difference in the use of /o/ <oo> in salient, high-type frequency patterns where /o/ <oo> is expected in standard language as well (i.e. /i/ <ie> and /œy/ <ui>) supports this hypothesis.

7. What does this mean for the language contact hypothesis?

The converging evidence this paper brings forward for a higher rate of strong forms among non-native speakers suggests it is worth to take another look at earlier literature that suggests an impact of language contact on the weakening of the strong verbs in the Germanic language and morphological simplification in general. Where Cuskley et al. (2015) are still able to broadly confirm theories about the effect of language contact on morphological simplification in general by showing that even though non-native speakers display a higher rate of irregularization, they do tend to simplify their irregular verb morphology in other ways (e.g. by extrapolating the level category), this does not hold up for Dutch. We did not find any evidence of simplification among the non-native speakers. Furthermore, even if this would have been the case, neither Cuskley et al.'s (2015) results nor ours offer support for theories ascribing a higher rate of weakening (or regularization) to the influx of non-native speakers, such as the studies by Carroll et al. (2012), De Smet et al. (2017) and Dale & Lupyan (2012) (see §1). When non-native speakers do not tend to use more weak forms, can migration and language contact really drive this change?

Of course, we need to be very careful in treating both Cuskley et al.'s (2015) results and our results as definite negative evidence for the language contact hypothesis. Combined, only two (closely related) Germanic languages (English and Dutch) have been scrutinized in this way. Furthermore, we do not know to what extent our results from a Wug task can be transferred to a real life situation. As Cuskley et al. (2015) already noted, an experiment with existing verbs might lead to very different results. At the same time, the high rate of strong forms might also be due to the relatively simple nature of the task, which stands in high contrast to actual language use that is much more complex.

When the rest of the utterance is more complex, non-native speakers may have to devote more cognitive resources to other aspects of morphological and syntactic planning, and may resort more to the simpler weak past tense variant. Another caveat lies in the proficiency of our participants. SLA research shows that especially early L2-learners tend to simplify verb morphology, while our participants were specifically more advanced L2-speakers. At the same time, however, our results show that proficiency is inversely correlated with the use of strong forms, which suggests less advanced L2-speakers would have used even more strong forms. Furthermore, we also noted that the higher rate of strong forms among non-native speakers might be partly due to formal L2-education, which was probably not always available to the same extent in the course of history or for all language groups. This could explain why, on a global scale, an inverse correlation between morphological complexity and language contact can be observed. Finally, even some form of the observer's paradox or social desirability bias may have skewed the results.

Yet, let us, for the sake of the argument, assume none of these factors play a role and our Wug test is an adequate representation of real language use. How can we then align our results with previous research about the impact of language contact on language change? A few possibilities come to mind. First, it might be the case that we do not need to look at previous studies about *morphological* complexity, but rather at *lexical* complexity. Reali et al. (2018), for example, show that languages with larger speaker populations (and thus also larger population sizes) feature larger vocabularies. It could be hypothesized that strong verb forms are processed and saved as lexical entries (unlike the weak verbs which are formed using a rule-based mechanism as has been famously claimed by i.a. Pinker 1998; Clahsen 1999; Pinker & Ullman 2002 as advocates of the *dual mechanism model*). Our observation that L2-speakers use more strong forms would thus be in line with the effects of language contact on the lexicon. However, we do have some important reservations about this theory. First of all, our experiment uses non-words, which means the participants have never seen these verbs before and thus they cannot conjure them up from their lexicon. Second, the *dual mechanism model* has been nothing but controversial. A decade-long debate has been going on between advocates of the *dual*

mechanism model on the one hand and advocates of a *single mechanism model* or *connectionist model* (such as Bybee & Moder 1983; Rumelhart & McLelland 1986; MacWhinney & Leinbach 1990; Hare & Elman 1995), who do not believe strong verbs are stored in the lexicon, but that both strong and weak verbs are processed by one single associative process, on the other hand. This is also reflected in our experiment: both L1- and L2-speakers form past tense forms of these non-words which they cannot have stored in their memory in analogy with existing patterns. Finally, even if we could align our results with the effects of language contact on the lexicon, this would still not explain why other studies showed or suggested a connection between weak forms and a growing population.

It could also be the case that it is not non-native speakers who drive change in situations of language contact, but rather native speakers. This could work in two ways. First, the instigation of morphological simplification could lie in situations of dialect contact between native speakers only, with dialect levelling and koineisation as a result (i.a. Kerswill 2002; Carroll et al. 2012). Secondly morphological simplification could also be a consequence of native speakers trying to accommodate their own language to the influx of immigrants by using the simpler variant. This suggestion is made by Cuskley et al. (2015: 218). They refer to the study by Dale & Lupyan (2012) which showed an effect of language contact on morphological simplification, in native speakers and mention studies by Smith (2007) and Bell (2007) who show that native speech can indeed be altered as a result of accommodation. Little (2011) also offers evidence that native speakers reduce morphosyntactic complexity more when talking to non-native speakers than when talking to native speakers and Atkinson et al. (2018) show with an artificial language learning experiment that speakers of a more complex language simplify their language use in interaction with speakers of a simpler language. However, it is not clear whether this accommodation would also take place when the non-native speakers do not actually tend to use the simpler forms themselves, as is shown in this Wug-task. Perhaps the subconscious idea that non-native speakers would use more simpler forms, is enough for native speakers to accommodate their speech, though this explanation may be rather far-fetched.

Finally, the introduction of new, more diverse past tense forms by non-native speakers may also be the instigation of morphological simplification at a later stage[15]. These more diverse past tense forms can complicate the verbal system for L1-speakers (and new L2-speakers) which may then in turn lead to learning problems and simplification of the system by the adoption of weak forms. This was also observed by Berdicevskis & Semenuks (2020). In their artificial learning experiment, irregularity of an abundant marker for agreement tended to increase first in the imperfect learning condition, which resulted in a lower learnability of the language and later in a complete loss of the marker.

8. Conclusion

In this experiment, we have replicated the study by Cuskley et al. (2015) for Dutch and shown that non-native speakers indeed tend to use more strong forms than native speakers. This effect tends to decrease as proficiency of Dutch increases and seems to be most outspoken for the most salient strong patterns, such as /ɛɪ-e/ <ij-ee>. No other types of simplification were found in the strong strategies used by non-native speakers, but we did find evidence that patterns that are part of informal language of native speakers are less easily picked up on. We could not straightforwardly confirm earlier theories about the effect of language contact on the weakening of the Germanic strong verbs or on morphological simplification in general. Though there are clear limitations to the use of a Wug task to study this, our paper suggests that morphological simplification through language contact might work in a more complex way than previously assumed.

[15] We thank an anonymous reviewer for this suggestion.

References

Atkinson, Mark, Kenny Smith & Simon Kirby. 2018. Adult learning and language simplification. *Cognitive Science* 42(8). 2818-2854.

Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed effects models using lme4. *Journal of Statistical Software* 67(1). 1-48.

Barton, Kamil. 2019. MuMIn: Multi-Model Inference. R package version 1.43.6. https://CRAN.R-project.org/package=MuMIn

Bell, N. 2007. How native and non-native english speakers adapt to humor in intercultural interaction. *International Journal of Humor Research* 20. 27-48.

Bentz, Christian & Bodo Winter. 2013. Languages with more second language learners tend to lose nominal case. *Language Dynamics and Change* 3. 1-27.

Berdicevskis, Aleksandrs & Arturs Semenuks. 2020. Different trajectories of morphological overspecification and irregularity under imperfect language learning. In Peter Arkadiev & Francesco Gardani (eds.), *The Complexities of Morphology*, 283-305. Oxford: Oxford University Press.

Berko, Jean. 1958. The child's learning of English morphology. *Word* 14. 150–177.

Brezina, Vaclav & Gabriele Pallotti. 2019 Morphological complexity in written L2 texts. *Second Language Research* 35(1). 99-119.

Bybee, Joan & Carol Lynn Moder. 1983. Morphological classes as natural categories. *Language* 59(2). 251-270.

Bybee, Joan. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10(5). 425-455.

Carroll, Ryan & Ragnar Svare & Joseph Salmons. 2012. Quantifying the evolutionary dynamics of German verbs. *Journal of Historical Linguistics* 2. 153-172.

Clahsen, Harald. 1999. Lexical entries and rules of language: a multidisciplinary study of German inflection. *Behavioral and brain sciences* 22. 991-1060.

Cuskley, Christine, Francesca Colaiori, Claudio Castellano, Vittorio Loreto, Martina Pugliese & Francesca Tria. 2015. The adoption of linguistic rules in native and non-native speakers: Evidence from a Wug task. *Journal of Memory and Language* 84. 205-223.

Dale, Rick & Gary Lupyan. 2012. Understanding the origins of morphological diversity: The linguistic niche hypothesis. *Advances in Complex Systems* 15. 1-16.

De Clercq, Bastien & Alex Housen. 2019. The development of morphological complexity: A cross-linguistic study of L2 French and English. *Second Language Research* 35(1). 71-97.

De Smet, Isabeau, Katrien Beuls, Dirk Pijpops & Freek Van de Velde. 2017. Language-specific differences in regularization rates of the Germanic preterite. Paper presented at the 23rd *International Conference on Historical Linguistics* (ICHL). San Antonio, July 31-August 4.

De Smet, Isabeau & Freek Van de Velde. 2019. Reassessing the evolution of West-Germanic preterite inflection. *Diachronica* 36(2). 139-179.

De Smet, Isabeau. 2021. *De sterke werkwoorden in het Nederlands*. PhD dissertation, University of Leuven.

Fox, John. 2003. Effect displays in R for generalised linear models. *Journal of Statistical Software* 8(15). 1-27.

Haeseryn, Walter, Kristin Romijn, Guido Geerts, Jaap de Rooij & Maarten C. van den Toorn. 1997. *Algemene Nederlandse Spraakkunst.* 2de editie. Groningen/Deurne: Martinus Nijhoff uitgevers/Wolters Plantyn.

Hare, Mary & Jeffrey Elman. 1995. Learning and morphological change. *Cognition* 56. 61-98.

Haspelmath, Martin. 2018. Revisiting the anasynthetic spiral. In Heiko Narrog & Bernd Heine (eds.), *Grammaticalization from a typological perspective*. Oxford: Oxford University Press.

Haspelmath, Martin & Susanne Michaelis. 2017. Analytic and synthetic: typological change in varieties of European languages. In Isabelle Buchstaller & Beat Siebenhaar (eds.*), Language variation* –

European perspectives VI: Selected papers from the Eighth International Conference on Language Variation in Europe (ICLaVE 8), Leipzig, May 2015, 3-22. Amsterdam: John Benjamins.

Hunt, Tyler. 2018. ModelMetrics: Rapid Calculation of Model Metrics. R package version 1.2.2. https://cran.r-project.org/web/package=ModelMetrics

Kerswill, Paul. 2002. Koineization and accommodation. In J.K. Chambers, Peter Trudgill & Natalie Schilling-Estes (eds.), *The handbook of Language Variation and Change,* 669-702. Oxford: Blackwell.

Keuleers, Emmanuel & Marc Brysbaert. 2010. Wuggy: A multilingual pseudoword generator. *Behavior Research Methods* 42(3). 627-633.

Keuleers, Emmanuel, Marc Brysbaert & Boris New. 2010. SUBTLEX-NL: A new frequency measure for Dutch words based on film subtitles. *Behavior Research Methods* 42(3). 643-650.

Knooihuizen, Remco & Oscar Strik. 2014. Relative productivity potentials of Dutch verbal inflection patterns. *Folia Linguistica Historica* 35. 173-200.

Koplenig, Alexander. 2019. Language structure is influenced by the number of speakers but seemingly not by the proportion of non-native speakers. *Royal Society Open Science* 6. 181274.

Kusters, Wouter. 2003. *Linguistic complexity: the influence of social change on verbal inflection.* Utrecht: LOT.

Levshina, Natalia. 2015. *How to do linguistics with R: data exploration and statistical analysis.* Amsterdam: John Benjamins.

Little, Hannah. 2011. *The role of foreigner directed speech in the cultural transmission of language and the resulting effects on language typology.* PhD dissertation. University of Edinburgh.

Lupyan, Gary & Dale, Rick. 2010. Language structure is partly determined by social structure. *PloS ONE* 5. 1-10.

MacWhinney Brian & Jared Leinbach. 1990. Implementations are not conceptualizations: revising the verb learning model. *Cognition* 40. 121-157.

McWhorter, John. 2005. *Defining creole*. Oxford: Oxford University Press.

McWhorter, John. 2007. *Language interrupted: Signs of non-native acquisition in standard language grammars.* Oxford: Oxford University Press.

Meyer David, Achim Zeileis & Kurt Hornik. 2020. *vcd: Visualizing Categorical Data.* R package version 1.4-8.

Miestamo, Matti, Kaius Sinnemäki & Fred Karlsson (eds.). 2008. *Language complexity: typology, contact, change.* Amsterdam: John Benjamins.

Navarro, Danielle. 2015. *Learning statistics with R: A tutorial for psychology students and other beginners.* R package version 0.5.1.

Nichols, Joanna. 2009. Linguistic complexity: A comprehensive definition and survey. In G. Sampson, D. Gil, & P. Trudgill (eds.), *Language complexity as an evolving variable*, 110-125. Oxford: Oxford University Press.

Nowak, Jessica. 2010. On the emergence of an eighth ablaut class in German and Dutch. *Journal of Germanic Linguistics* 22(4). 351-380.

Oldenkamp, Loes. 2013. *The trouble with inflection for adult learners of Dutch. A study of the L1-L2 interplay of morphosyntactic and phonetic-phonological factors.* PhD-dissertation. LOT.

Perc, Matjaž. 2014. The Matthew effect in empirical data. *Journal of The Royal Society Interface*, 11(98). 20140378.

Pinker, Steven. 1998. Words and rules. *Lingua* 106. 219-242.

Pinker, Steven & Michael Ullman. 2002. The past-tense debate: the past and future of the past tense. *TRENDS in Cognitive Sciences* 6(11). 456-463.

Reali, Florencia, Nick Chater & Morten H. Christiansen. 2018. Simpler grammar, larger vocabulary: How population size affects language. *Proceedings of the Royal Society B: Biological Sciences*, 285(1871). 20172586.

Raviv Limor, Antje Meyer & Shiri Lev-Ari. 2019. Larger communities create more systematic languages. *Proceedings of the Royal Society B: Biological Science* 286(1907). doi:10.1098/rspb.2019.1262

R Core Team. 2017. *R: A language and environment for statistical computing. R Foundation for Statistical Computing*. Vienna. http://www.R-project.org/

Ritchie, Stuart. 2020. *Science fictions: how fraud, bias, negligence, and hype undermine the search for truth*. New York: Henry Holt.

Rumelhart, David Everett & James Lloyd McClelland. 1986. On learning the past tense of English verbs. In James Lloyd McClelland & David Everett Rumelhart (eds.), *Parallel distributed processing: Explorations in the microstructure of cognition.* Cambridge: MIT Press.

Schepens, Job, van der Slik Frans & Roeland van Hout. 2013. Learning complex features: A morphological account of L2 learnability. *Language Dynamics and Change* 3(2). 218-244.

Sinnemäki, Kaius. 2009. Complexity in core argument marking and population size. In G. Sampson, D. Gil, & P. Trudgill (eds.), *Language complexity as an evolving variable*, 126-140). Oxford: Oxford University Press.

Sinnemäki Kaius & Francesca Di Garbo. 2018. Language Structures May Adapt to the Sociolinguistic Environment, but It Matters What and How You Count: A Typological Study of Verbal and Nominal Complexity. *Frontiers in Psychology* 9: 1141. https://www.frontiersin.org/article/10.3389/fpsyg.2018.01141. DOI: 10.3389/fpsyg.2018.01141

Smith, C. 2007. Prosodic accommodation by French speakers to a nonnative interlocutor. In Jürgen Trouvain, *Proceedings of the XVIth international congress of phonetic sciences 6-10 August 2007, Saarbrücken, Germany*, 313-348. Saarbrücken: Saarbrücken Universität des Saarlandes.

Sönning, Lukas & Werner, Valentin. 2021. The replication crisis, scientific revolutions, and linguistics. *Linguistics* 59(5). 1179-1206.

Trudgill, Peter. 2002. Linguistic and Social Typology. In Jack K. Chambers, Peter Trudgill & Natalie Schilling-Estes (eds.), *The handbook of language variation and change*, 707-728. Oxford: Blackwell.

Trudgill, Peter. 2004. Linguistic and social typology: The Austronesian migrations and phoneme inventories. *Linguistic Typology* 8. 305-320.

Trudgill, Peter. 2011. *Sociolinguistic typology: social determinants of linguistic complexity*. Oxford: Oxford University Press.

Van der Slik, Frans, Roeland van Hout & Job Schepens. 2015. The gender gap in second language acquisition: gender differences in the acquisition of Dutch among immigrants from 88 countries with 49 mother tongues. *Plos One* 10(11). e0142056.

Van der Slik, Frans, Roeland van Hout & Job Schepens. 2019. The role of morphological complexity in predicting the learnability of an additional language: The case La (additional language) Dutch. *Second Language Research* 35(1). 47-70.

Van Haeringen, Coenraad B. 1940. De taaie levenskracht van het sterke werkwoord. *De Nieuwe taalgids* 31. 241-255.

Wickham, Hadley. 2007. Reshaping Data with the reshape Package. *Journal of Statistical Software*, 21(12). 1-20. http://www.jstatsoft.org/v21/i12/

Wickham, Hadley. 2016. *ggplot2: Elegant Graphics for Data Analysis.* New York: Springer-Verlag.

Wickham, Hadley, Romain François, Lionel Henry & Kirill Müller. 2019. *dplyr: A Grammar of Data Manipulation*. R package version 0.8.3. https://CRAN.Rproject.org/package=dplyr

Zuur, Alain F., Elena N. Ieno, Neil J. Walker, Anatoly A. Saveliev & Graham M. Smith (eds.). 2009. *Mixed effects models and extensions in ecology with R: Statistics for Biology and Health.* New York: Springer.

Acknowledgments

We want to sincerely thank our two anonymous reviewers and Christine Cuskley whose comments have been incredibly useful for this paper. We also want to thank all participants and everyone who shared this experiment and helped us in the search for non-native participants.

Appendix

Table A. Numerical output model 1 – strong vs. weak inflection for native vs. non-native speakers: fixed effects

<insert Table A here>

Table B. Numerical output model 1 – strong vs. weak inflection for native vs. non-native speakers: random effects

<insert Table B here>

Table C. Numerical output model 2 – strong vs. weak inflection for native vs. non-native speakers with interaction with present stem vowel: fixed effects

<insert Table C here>

Table D. Numerical output model 2 – strong vs. weak inflection for native vs. non-native speakers with interaction with present stem vowel: random effects

<insert Table D here>

Table E. Numerical output model 3 – non-native speakers only: fixed effects

<insert Table E here>

Table F. Numerical output model 3 – non-native speakers only: random effects

<insert Table F here>

Table G. Numerical output model 4 – use of unexpected strong strategies: fixed effects

< insert Table G here>

Table H. Numerical output model 4 – use of unexpected strong strategies: random effects

<insert Table H here>

Table I. Numerical output model 5 – use of /o/ versus other vowels for native versus non-native speakers: fixed effects

<insert Table I here>

Table J. Numerical output model 5 – use of /o/ versus other vowels for native versus non-native speakers: random effects

<insert Table J here>

Data availability statement

The data underlying this article will be shared on the lead author's github page.