Morphology (2025) 35:449–487
https://doi.org/10.1007/s11525-025-09445-2

Rhythm-affected plural variation in written Dutch

Tim Zee[1] · **Louis ten Bosch**[1] · **Mirjam Ernestus**[1]

Received: 22 September 2023 / Accepted: 1 June 2025 / Published online: 13 June 2025

Abstract
Previous research has shown that users of certain languages tend to choose and order their words such that they avoid stress clashes. This tendency has been attributed to the Principle of Rhythmic Alternation (PRA). It has also been claimed that language users alternate between morphological structures to abide by the PRA. However, in our estimation, it has not yet been convincingly shown that PRA effects documented for morphology are indeed morphological instead of syntactic or lexical in nature. The present research shows that the alternation between two plural suffixes for certain Dutch nouns is a genuine case of morphological PRA by controlling for alternative mechanisms. We conducted two experiments and one corpus study. In the first experiment, we had participants pluralize nouns which were embedded in sentences such that only one of the plural variants would not lead to stress clash. We did not find a PRA effect in this study—possibly because participants inserted prosodic breaks between the plural and the next word. This explanation was supported by the results of a second follow-up experiment. In the corpus study, we looked at the distribution of plural suffixes for both variable and invariable nouns. As stress clashing plural forms were avoided only for the variable nouns, we can be reasonably confident that this PRA effect was due to a morphological process. Apart from identifying a new PRA phenomenon, this result strengthens previous claims about the relevance of the PRA for models of morphological processing.

Keywords Principle of rhythmic alternation · Morphology-prosody interface · Stress clash · Plural inflection · Dutch

1 Introduction

1.1 Background

Many languages show a general bias towards an alternation of stressed and unstressed syllables (e.g., Jespersen, 1909; Kager, 1989). This observation has led researchers

✉ T. Zee
tim.zee@ru.nl

[1] Radboud University, Nijmegen, The Netherlands

to posit a *Principle of Rhythmic Alternation* (henceforth PRA). One manifestation of the PRA concerns the alternation between word forms expressing the same morphological function. For instance, previous research on English adjectives (e.g., Schlüter, 2015) has shown that, given two alternative morphological forms and a following stressed syllable, the form that does not result in two consecutive stressed syllables is more likely to be produced, e.g. *drunken man* is more likely than *drunk man*. Such findings (see the overview in Arndt-Lappe & Hoffmann, 2022) have been used to argue that prosody and morphology interact during the production process (e.g., Schlüter, 2015). The current paper argues that, although these studies are certainly valuable in illustrating the PRA, they do not provide conclusive evidence that these effects involve a morphological process. In order to make that claim, alternative explanations of the observed data need to be ruled out. By identifying and taking alternative causal mechanisms into account, the current research investigates whether rhythm truly affects the choice between morphological variants during production. Answering this question is relevant in a more general psycholinguistic context as certain psycho-linguistic production models do not allow for a prosodic influence on morphological processing (e.g., Levelt et al., 1999).

The remainder of this section is as follows. Section 1.2 describes the PRA in more detail, and outlines the problems with the interpretation of PRA effects involving morphological alternations. Section 1.3 proposes an experiment to solve the interpretation issues, and explains why Dutch variable plurals are suitable for an experimental investigation of the relation between rhythm and morphological alternations.

1.2 The PRA and the pitfalls of its interpretation

The PRA is a term for the general bias in certain languages towards an alternation of stressed and unstressed syllables. In other words, the PRA is the avoidance of *stress clash*, which refers to the presence of two consecutive stressed syllables. This principle manifests itself in many different phenomena. As proposed by Arndt-Lappe and Hoffmann (2022), these phenomena can be categorized into three types: (1) stress shift, (2) alternation between (a) lexical or (b) morphological variants, and (3) alternation between different word orderings. Although this paper is focussed on morphological (type 2b) phenomena, we briefly introduce all three.

The first type, stress shift, describes the phenomenon whereby speakers change the canonical stress pattern of a word to avoid consecutive stressed syllables. This phenomenon is also known as the *thirteen-men rule*, as the stress in words like *thirteen* shifts from the second syllable to the first if it is followed by a stress-initial word like *men*. Another example is given by Gussenhoven (1983), who shows that the Dutch word *hardóp* 'out loud' has final stress in isolation and initial stress when followed by a stressed syllable, e.g., *hárdop vóórlezen* 'read out loud'.

The second type of PRA phenomena involves language users selecting lexical units—complete words (henceforth lexical PRA, type 2a) or affixes (henceforth morphological PRA, type 2b)—such that a stress clash is avoided. This is illustrated by the distribution of near-synonyms that differ in terms of stress pattern, such as *rich*, which has final stress, and *wealthy*, which does not have final stress. Schlüter and Knappe (2018) show that, before initially stressed nouns, *rich* tends to be avoided

more than *wealthy*. Similar distributional effects have been found for affix (type 2b) alternations, where the morphological variant of a word that is used depends on the position of stress in the next word. For instance, in Dutch, some adjectives double as adverbs, such as *enórm* 'enormous(ly)'. In certain inflections, attributively used adjectives are suffixed with a final *-e*, pronounced [ə], which distinguishes them from their adverbial forms: for example, *de enórm gróte man* 'the enormously big man' vs. *de enórme gróte man* 'the enormous big man'. However, Quené (2006) found that adverbs may also be inflected with *-e*, pronounced [ə], if this resolves a violation of the PRA, e.g., *de enórme gróte man* meaning 'the enormously big man'.

Type 3 or syntactic PRA phenomena involve language users choosing a particular syntactic structure or word ordering such that consecutive stressed syllables are avoided. Such phenomena can be studied using the distributions of alternative word orderings that differ in stress pattern but convey the same meaning. To give an example, in Dutch subordinate clauses, the past participle can either be preceded or followed by an auxiliary verb: 'that he has paid Klaas' can be produced as both *dat hij Kláás heeft betáálд* and *dat hij Kláás betáálд heeft*. Arfs (2007) demonstrates that when the past participle starts with a stressed syllable, the participle-auxiliary ordering is avoided, presumably because it would result in two consecutive stressed syllables: 'that he has called Klaas' is more frequently produced as *dat hij Kláás heeft ópgebeld* than as *dat hij Kláás ópgebeld heeft*.

Although the differences between morphological (type 2b) and other types of PRA phenomena are well-defined on a theoretical level, in practice it appears difficult to prove that a given PRA phenomenon is in fact morphological. This difficulty can be illustrated by considering the causal elements that define a morphological PRA effect. The first causal element of this definition is straightforward: the rhythmic context affects the morphological form that is used rather than the other way around. If this directionality is not taken into account, certain types of PRA effect can no longer be distinguished. For instance, Schlüter (2015) claims that the low frequency of *drunk* relative to *drunken* preceding initially stressed words constitutes evidence of a direct effect of rhythm on the selection of the morphological variant. However, the same distribution of *drunk* and *drunken* can result from a lexical PRA effect, in which the selection of *drunk* favours certain continuations over others, e.g., *drúnk indivídual* over *drúnk pérson* (a type 2a effect), or result from a syntactic PRA effect, in which the selection of *drunk* is associated with certain syntactic structures, e.g., *drúnk and háppy péóple* instead of *háppy and drúnk péóple* (a type 3 effect).

The second causal element crucial to the definition of PRA effects on morphology is that it involves selection of a morphological variant in favour of another by individual language users, rather than an overall preference for a variant in a given context in a population of speakers. Different language users may each use only one morphological variant of a word such that the aggregated data show a distribution that avoids stress clash. For instance, the skewed distribution of uninflected and *-e* suffixed Dutch adverbs (e.g., *enórm*/*enórme* 'enormously') preceding initially stressed adjectives (e.g., *gróte* 'grote') is argued to be the result of language users inserting a weak syllable to avoid stress clash (Quené, 2006). However, this distribution can also be observed if certain language users exclusively use *enorm* and other language users exclusively use *enorme* to modify an adjective. The *enorme*-users would never

run into a stress-clash, and the *enorm*-users could resolve stress-clashes by selecting a near-synonym of *enorm* without final stress such as *gigántisch* 'gigantic'. In other words, if we do not take into account that a PRA effect may occur across a population but not within individuals, we ignore the possibility that the PRA mechanism at work is not morphological but lexical.

Taking the causal elements of morphological PRA effects into consideration is especially difficult in corpus studies. The directionality element is problematic because the context and the morphological variants are observed together, that is, it is not clear which came first in the production process. Likewise, it is often difficult to establish whether the distribution of morphological variants is due to individual language users choosing one variant over another or is due to inter-user differences. Many corpora do not allow researchers to trace data back to individual users, or they are too small to provide a large enough sample of morphological variants of a particular lemma by a particular language user. In spite of the causal pitfalls associated with interpreting PRA effects in corpus studies as morphological, most if not all previous studies on morphological PRA effects are based on corpus data and do not take into account alternative causal mechanisms. Not addressing alternative explanations is not just a problem for the interpretation of specific PRA effects, but also for the theoretical claims about morphological processing that are based on those interpretations.

Importantly, the suggestion that previously described morphological PRA phenomena are all confounded by some other mechanism may not be as unlikely at it first appears. Recent studies (Breiss & Hayes, 2020; Hoffmann et al., 2022) have found that stress clashes were avoided in word bigrams from written English corpora containing millions of words. The bigrams were not selected for any morphological, lexical or syntactic alternations in particular. Given the fact that morphological alternations are rather rare in English (Schlüter, 2015), this overall PRA effect likely is the result of rhythmically induced lexical and structural alternations. Thus, it is a real possibility that these mechanisms are also responsible for the rhythmic effects in the previously reported datasets that were selected for particular morphological alternations. Given such alternative explanations of morphological PRA phenomena in corpus data, a more controlled, experimental study may be more successful in identifying the mechanism behind a given PRA phenomenon.

1.3 An experimental study on PRA-affected plural variation

Many previously described morphological PRA phenomena are not suitable for experimental investigation for a number of reasons. Firstly, one of the morphological variants of the alternation may have completely disappeared from the contemporary standard language, e.g. *worser* as an alternative to *worse*. Secondly, the distribution of the alternation may no longer be driven by a rhythmic preference but rather by semantic or functional differences (e.g., Arndt-Lappe & Hoffmann, 2022). For instance, although it is likely that the *drunk*/*drunken* distribution across attributive and non-attributive positions was originally driven by a rhythmic constraint, over time the use of *drunken* in predicative position completely disappeared. As such, in present day English, there likely is very little active competition between *drunk* and *drunken* during production: some language users may simply reserve *drunken* for attributive

use (with certain nouns), whereas others always use *drunk*, regardless of the syntactic position. Third, the variants of the alternation may be part of very frequent and possibly lexicalized collocations, such as the Dutch adverb *hele* in *hele grote* 'very big'. In sum, it seems that few contemporary PRA phenomena have been described that are likely to involve the active competition between morphological variants that is required for an experimental study.

A contemporary morphological alternation that may be suitable for an experimental investigation of morphological PRA effects can be found in Dutch noun plurals. In Dutch, most plural nouns end in either *-en* or *-s*, for example *dier+en* 'animals' or *bakker+s* 'bakers'. Generally, *-en* is used for nouns with a stressed final syllable, whereas *-s* is used for nouns with an unstressed final syllable. However, these generalizations do not consistently apply to certain final-stress nouns like *directeur* 'director' [ˌdi.rɛkˈtøːr], for which both *directeur+en* and *directeur+s* are acceptable plurals. This instance of morphological overabundance (Thornton, 2019) has been documented in dictionaries (e.g., Van Dale et al., 2015) and textbooks on Dutch morphology (e.g., De Haas & Trommelen, 1993). A number of factors have been suggested to influence the distribution of this variation, such as modality (Kürschner, 2009), register (Baayen et al., 2002), dialect (de Schutter et al., 2005), and time (Theissen, 2003). However, it is still unclear which factors lead a language user, in a given time, modality and style, to choose one variant over the other. One potential candidate might be the rhythm of the phrase. Specifically, the *-s* plural variant of a noun with final stress, e.g., *directéúrs*, might be avoided when it precedes a word with initial stress, e.g., *práten* 'talk', as this would result in two consecutive stressed syllables. In contrast, *-en* plural variants are not expected to be avoided before stress-initial words, because the *-en* suffix always adds an unstressed final syllable. Thus, Dutch variable plurals potentially are an example of a morphological PRA effect.

Dutch plural variation has a number of characteristics that make it suitable for experimental investigation. Firstly, individual Dutch speakers know and use both variants of many plurals. This is evidenced by speakers using both variants in the same sentence (Zee et al., 2021). The claim that Dutch speakers know both variants is also evidenced by the finding that the token frequency of a variant relative to its alternative affects the pronunciation of the chosen plural variant (Zee et al., 2021). This suggests that both variants and their relative frequencies are somehow represented in the psychological processes involved in spoken word production. Secondly, although the two plurals for some nouns differ in their meanings (e.g., *wortelen* is primarily used for 'carrots' whereas *wortels* is used for the different senses of 'roots', Haeseryn et al., 1997), the two plurals have the same meaning for most nouns. For instance, *saxofonen* and *saxofoons* both refer to more than one *saxofoon* 'saxophone'. Third, Dutch noun plurals can be followed by many different word types, which makes it likelier that many noun plurals will not be part of lexicalized collocations. Given these characteristics, it is a distinct possibility that a rhythmic effect on the production of Dutch variable plurals involves competition between morphological alternatives, which should be measurable in an experiment.

Two aspects of the production process need to be controlled in an experiment that investigates whether the choice of a plural variant is affected by the presence of following stress. Firstly, we need to make sure that the prosodic context affects the

choice of plural form rather than the other way around. This can be controlled by providing the context of the plural. Secondly, we need to be sure that the participants choose between plural variants of the same lemma (e.g. *directéúrs* and *directéúren*) rather than choosing between plurals of two different lemmas (e.g. *directéúrs* and *leidinggévenden*). This can be achieved by also providing the lemma.

The present study implements both controls by presenting participants with written sentences containing the singular form of variable plurals and asking them to type a plural version of the sentence. We manipulated the context directly following the target noun to have initial stress or not, see (1) for an example in which *sergeant* is the target noun and *dankte* / *bedankte* reflects the stress manipulation.

(1) De sergeant *dankte* / *bedankte* zijn troepen voor hun inzet.
'The sergeant *thanked* his troops for their efforts.'

If the Dutch plural alternation represents a case of morphological PRA, we predict fewer *-s* variant productions when the plural is followed by a stressed syllable (e.g., *sergéánts dánkte*) compared to an unstressed syllable (e.g., sergéánts *bedánkte*). If the results do not show an effect of stress pattern on the number of -s variant realisations, further investigations are necessary to determine the reason for the null effect.

2 Experimental study 1

2.1 Methodology

2.1.1 Participants

Participants were recruited and remunerated using the online *Prolific* platform. In total, 82 participants signed up, two of which revoked their participation during the experiment. No further exclusions were made, resulting in a sample of 80 participants. Participants were selected to be adult native speakers of Dutch without language impairments. The participants' age ranged from 19 to 75 years with a mean of 32.7. In terms of gender, the participants were relatively evenly distributed between 37 identifying as women, 41 identifying as men and 2 identifying as other.

2.1.2 Materials

To maximize the probability that participants in an experiment adapt their choice of plural to the prosodic context, they should be presented with those nouns that show the most variability. It is difficult to determine these nouns on the basis of existing corpora because these corpora do not contain sufficient data from different speakers to show within speaker variability.

As such we conducted a pilot study in which 21 native Dutch speakers rated 65 nouns on how often the *-s* and *-en* variants occurred. The mean age of the participants was 28.76 years old. The ages ranged from 19 to 70 years old. Most of the participants also spoke English (18). About half of the participants also spoke German (10), and seven participants spoke French. The participants indicated their estimated proportion

of the two variants on a slider scale that ranged from 0 to 1, where 0 reflected always *-s* and 1 reflected always *-en*. We assume that participants' answers would reflect their own preference to some degree. The resulting ratings (see Appendix A) were used to select the nouns for the main experiment as follows. Only those nouns were selected for which at least 20% of participants indicated a proportion between 0.1 and 0.9. Furthermore, only nouns with mean proportions between 0.25 and 0.75 were selected. This resulted in a final selection of 28 nouns for the main experiment.

Each noun was embedded in a sentence as the subject and it was immediately followed by a finite verb. The finite verb and the rest of the sentence were chosen to make sense in combination with the subject noun. Importantly, each verb was selected such that it could be prefixed without large modifications to the structure or meaning of the sentence, see (2) for an example.

(2) a. De generaal kondigde zijn nieuwe plannen aan.
'The general announced his new plans.'
b. De generaal verkondigde zijn nieuwe plannen.
'The general proclaimed his new plans.'

In (2a) we used the particle verb *aankondigen* which has the same stem and a similar meaning as the prefixed verb *verkondigen* in (2b). As a result, the only difference in the immediate context of the target noun *generaal* is the unstressed prefix *ver-* in (2b), which contrasts with the stressed syllable *kon-* in (2a). When participants are asked to produce a plural version of (2b), the plural variant *generaals* is not expected to be avoided because the following verb *verkondigden* does not have initial stress. However, in (2a), the initial stress of the verb *kondigden* is expected to make the *generaals* plural variant less likely to occur. To prevent priming effects, each participant was presented with either the *unstressed* (2b) or the *stressed* (2a) condition of a given target noun.

We included 12 filler sentences which contained nouns with invariable plurals to make it less obvious that the experiment was about variable plurals. In order to further mask the purpose of the experiment, we also asked the participants to produce singular versions of 40 sentences with invariable plural noun subjects. Some of these sentences contained gender-stereotyped nouns and (gender-neutral) plural possessive pronouns, requiring the participant to make a gender assignment when constructing the singular sentence as Dutch does not have a gender-neutral singular possessive pronoun, see (3) for examples.

(3) a. De agenten trokken hun dienstwapen.
'The police officers drew their service weapons.'
b. De receptionisten hielden hun pauze om 12 uur.
'The receptionists held their break at 12 o'clock.'

A singular version of the sentences in (3) would require the participant to choose between *zijn* 'his' and *haar* 'her' for the possessive pronoun corresponding to the respective subjects *agent* 'police officer' and *receptionist* 'receptionist'. By including jobs like police officer and receptionist, which are male- and female-dominated, respectively (CBS, 2022), we expected that participants would become aware of the

gender choice they had to make and would be distracted from the true purpose of the experiment.

In total, each participant was presented with 80 sentences (28 target sentences + 12 filler sentences with singulars of invariable plurals + 40 filler sentences with invariable plurals). Four different orderings of the sentences were made, each containing 14 target sentences in the stressed and 14 in the unstressed condition. In each ordering, four filler sentences were used at the start, the number of consecutive target sentences was limited to four, and phonologically similar target nouns (e.g., *directeur* and *redacteur*) were prevented from following each other. For each of these four orderings, two versions were made, resulting in eight stimulus lists in total. For each pair of lists, the 14 target nouns that were presented in the stressed condition in version A were presented in the unstressed condition in version B, and *vice versa*. List 1A and 1B can be found in Appendix B.

2.1.3 Procedure

The experiment was conducted online in Qualtrics. The procedure of the experiment was as follows. Participants were presented visually with sentences such as those in (2) and (3). They were then asked to indicate how surprising the sentence was (not surprising, a little surprising, very surprising) to ensure that they fully processed the sentence. As some of the filler sentences were manipulated to be semantically surprising, these responses could also be used to check whether participants were still paying attention,[1] see (4) for an example of a surprising sentence.

(4) De baviaan liet zijn dagboek zien.
'The baboon showed his diary.'

After indicating their surprisal, participants were asked to change the number of the sentence and type it out. That is, if they had been presented with a singular sentence, they should produce a plural version and *vice versa*. They could then click on a button that started the next trial. After all trials had been presented, participants answered a few questions about their linguistic background, a few demographic questions, and they indicated what they thought the experiment was about.[2]

2.1.4 Analysis

We automatically extracted the subject noun of each target response, and coded the dependent variable *Plural* as 1 if the noun was pluralized using *-s* and as 0 if it was pluralized using *-en*. We excluded responses in which participants had left the target noun in singular form, responses that were misspelled to such a degree that the intended plural could not be established, and responses in which the plural form belonged to a different lexeme than the singular. This resulted in a final data set of

[1] All participants had more *a little* / *very surprising* responses to surprising sentences and more *not surprising* responses to unsurprising sentences.

[2] Most of the participants thought the experiment was about gender. Only two participants thought the experiment was about the type of plural that was used.

Table 1 Distribution of the variables in the data from the first experimental study

Dependent variable	Levels	Token numbers
Plural	1 (*-s*)	1136
	0 (*-en*)	1055
Fixed predictor		
Next Stress	Stressed	1089
	Unstressed	1102
Random variables	Number of levels	Examples
Lemma	28	admiraal, alarm, balkon
Participant	80	PP1, PP2, PP3

2191 observations. The predictor of interest *Next Stress* had two levels: *Stressed* if the verb following the plural had initial stress; *Unstressed* if the verb did not have initial stress. In addition, we took into account the random variables *Lemma* and *Participant*. Table 1 summarizes the variables in the data set.

We used Bayesian multi-level logistic regression, as implemented in the R packages *rstan* (Stan Development Team, 2021) and *rethinking* (McElreath, 2020), to model the *Plural* variable. The predictor (*Next Stress*) and the random variables (*Lemma* and *Participant*) were entered into the models as index variables rather than dummy variables, and we used mildly regularizing priors for all model parameters. To find out whether *Next Stress* was a relevant predictor of the plural variant distribution, we created two models: a *baseline model*, which only contained the random variables and a fixed intercept, and a second, *PRA model*, which also included *Next Stress*. We then compared the models using the Watanabe-Akaike information criterion (**WAIC**) to see whether adding the fixed predictor improved the model. Subsequently, we interpreted the effect of *Next Stress* by subtracting the posterior distribution of the parameter for the *Unstressed* level from the posterior distribution of the parameter for the *Stressed* level. The resulting *Stressed – Unstressed* contrast indicates to what extent the log-odds of an *-s* variant decreases or increases before a verb with initial stress. Finally, we constructed a third model that included by-participant and by-lemma random contrasts for the *Next Stress* parameter to investigate whether the plural distribution of some participants or nouns showed more of an effect of the PRA than others.

2.2 Results

The *baseline model* of the experimental data had a WAIC of 2279.2, whereas the *PRA model* had a WAIC of 2280.7, resulting in a difference of 1.4 with a standard error of 0.26. This means that, in terms of WAIC, including the *Next Stress* predictor resulted in a worse fitting model. The *Stressed – Unstressed* contrast for the *Next Stress* predictor was calculated from the posterior distributions summarized in Table 6 in Appendix C. As Fig. 1 illustrates, there is no clear difference in log-odds of an *-s* variant between the *Stressed* and *Unstressed* conditions.

The null result from the PRA model leaves open the possibility that some participants or some lemmas showed the PRA effect and that other participants or lemmas

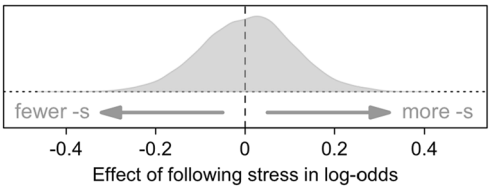

Fig. 1 Posterior distribution of the *Stressed – Unstressed* contrast in the PRA model of the first experimental study. See running text for details

showed an effect in the opposite direction. As such, a model was fitted with random slopes for both participants and lemmas. The random slopes revealed that none of the participants or lemmas showed clear differences across the stressed and unstressed conditions. The unnecessary complexity of this model resulted in a WAIC of 2292.8, making it much worse than the baseline model.

3 Interim discussion 1

The results from Experiment 1 show no evidence of a rhythmic effect on the selection of plural variants. This null result can be explained in a number of ways. Firstly, it could be that the production of Dutch variable plurals is not affected by the PRA at all. Secondly, the null result could mean that Dutch variable plurals are only involved in a lexical or structural PRA effect. Third, the absence of an effect could be due to some issue with the experiment.

We considered the third option to be the most likely explanation for two reasons. Firstly, participants might have resolved the stress clashes using prosodic boundaries. Boundary insertion may have been encouraged by the fact that participants first read a sentence with the noun in singular form before producing a plural version. This meant that participants were presented with a stress clash for all stimuli in which the singular noun was followed by an initially-stressed verb. Participants may have inserted a prosodic boundary in those stimuli to resolve the stress clash when they processed those sentences (see Kentner, 2015 for a similar effect on accentuation). When they subsequently pluralized the sentence, they may have just kept the prosodic boundary in place, which prevented the stress of the word following the boundary from playing any role in the selection of the plural variant. Unfortunately, prosodic boundaries in this sentence position are not represented in orthography. As such, we could not investigate this possible explanation using the participants' responses.

Secondly, the experiment involved online participants that were doing the experiment in an uncontrolled environment. It may be that PRA effects only surface when language users are sufficiently focused on a given task.

In order to find out whether the null effect in Experiment 1 can be explained by one of these shortcomings, we conducted a second experiment. To test whether participants used prosodic boundaries to prevent stress clashes, Experiment 2 elicited spoken sentences rather than written sentences, so that we could observe whether participants had inserted a prosodic boundary after a noun. Moreover, Experiment 2 took place in a controlled lab environment, to prevent participants from becoming distracted. Section 4 describes Experiment 2 in more detail.

4 Experimental study 2

4.1 Methodology

4.1.1 Participants

Participants were recruited from the participant pool of Radboud University. In total, 40 participants were recruited none of which were excluded. Participants were selected to be native speakers of Dutch without language impairments who were at least 16 years old. The participants' age ranged from 16 to 62 years with a mean of 26.2. In terms of self-identified gender, the participants were strongly biased with 33 women, 6 men and 1 person who would rather not say.

4.1.2 Materials

We used the sentences and orderings of these sentences from Experiment 1, see Sect. 2.1.2 for details. The main difference in the materials was the language modality. In Experiment 2, the sentences were presented to the participants both in text and in speech. The recordings were made by the first author using the internal microphone of a 14-inch 2021 MacBook Pro.

The auditory target stimuli were designed to have no prosodic boundary between target noun and verb, as we did not want to bias participants towards using a prosodic boundary between the plural and the following verb in their responses. In order to make prosodic boundaries in the target stimuli less likely, the prosodic focus was always placed on a word after the finite verb, see the word in uppercase in example (5).

(5) De generaal kondigde zijn NIEUWE plannen aan.
'The general announced his NEW plans.'

As it would be rather salient if all stimuli contained such a prosodic focus, some of the fillers showed a more conventional focus on the noun, see example (6).

(6) De PRINSEN zwaaiden naar hun onderdanen.
'The PRINCES waved to their underlings.'

Of all the auditory stimuli that the participants heard, exactly 50 percent had a prosodic focus on the first noun, as in (6), and the other 50 percent, which included all of the target stimuli, had a prosodic focus after the finite verb, as in (5).

4.1.3 Procedure

Participants sat at a desk in a sound-attenuated booth in front of computer screen, mouse and keyboard and a Audio-Technica ATR20 microphone. After the participants had read the instructions, they put on AKG K371 headphones. The experiment was built and conducted in the browser-based Gorilla platform. Each trial started with a simultaneous textual and auditory presentation of the stimulus. After the auditory

stimulus finished playing, a (visual) 5 second countdown started, during which the participants had to repeat the sentence out loud while changing the number of the subject and conjugating the finite verb to match the subject. While the countdown was running, participants could still see the textual stimulus. This was done because we found that participants in a pilot version of the experiment had trouble fluently speaking the entire sentence if the textual stimulus disappeared after the auditory stimulus had finished. After the countdown ran out, participants were asked how surprising the sentence was using a multiple-choice question (not surprising, a little surprising, very surprising). As in Experiment 1, some of the sentences were designed to be semantically surprising, see the example in (4). After they had answered the multiple-choice question, the next trial started. After all trials had been completed, participants answered the same questionnaire as in Experiment 1.[3]

Apart from the change to auditory stimuli and responses, the main procedural change compared to Experiment 1 is the order in which participants produced their modified sentence and indicated their surprisal. We chose to have participants say their modified sentence out loud before indicating their surprisal, because we thought this might increase the probability that participants mirrored the prosody from the stimulus, thereby avoiding prosodic boundaries between target noun and verb.

4.1.4 Analysis

We listened to each target response and annotated whether the *Plural* was produced with the *-en* or *-s* suffix. Certain responses were excluded from further analysis for a variety of reasons. As was the case in Experiment 1, some participants pluralized *doorn* 'thorn' as *dorens*. Similarly, in one response, *redacteur* 'editor' was pluralized as *redactrices* 'female editors'. In addition, sometimes participants simply repeated the singular sentence, forgetting to pluralize the noun. Finally, in a few cases participants responded too late for their productions to be recorded. In total, 50 of the 1120 data points were excluded due to a divergent response.

Two annotators, TZ and KK, indicated whether a response was produced with a *Prosodic Boundary* between plural noun and verb. For these annotations, the annotators had three options: *boundary*, *no boundary*, and *unclear*. Responses were annotated as *unclear* if it could not be decided whether there was a very slight boundary or not, or if the participant misspoke in such a way that it could not be determined whether a boundary was present. By excluding all responses that were annotated as *unclear*, we hoped to prevent any bias from the annotators, who knew about the goal of the experiment, from influencing the results. These exclusions resulted in a final dataset of 863 data points. Excluding *unclear* annotations, the annotators agreed on 80% of the data. This relatively low agreement percentage probably stems from the fact that annotator KK had very few *unclear* annotations. For the analysis, we used the annotations by TZ, as he had more *unclear* annotations.

[3] All participants had more *a little/very surprising* responses to surprising sentences and more *not surprising* responses to unsurprising sentences. None of the participants suspected the experiment was about variable plurals.

Table 2 Distribution of the variables in the data from Experiment 2. Interactions refers to predictors that are entered into our statistical models as an interaction with Next Stress

Interactions		Stressed	Unstressed
Plural	*-s*	221	242
	-en	201	199
Prosodic Boundary	Boundary	174	132
	No Boundary	248	309
Fixed predictor			
Previous plural	*-s*	513	
	-en	350	
Random variables		Number of levels	Examples
Lemma		28	admiraal, alarm, balkon
Participant		40	PP1, PP2, PP3

The *Next Stress*, *Lemma*, and *Participant* variables were defined as they were in Experiment 1. For each target stimulus, we also wanted to account for a potential priming effect from the previous stimulus. For instance, if two subsequent trials require the participant to form a plural, and the participant uses an *-s* suffix in the first trial, they might be primed to use the *-s* suffix again in the following trial. Alternatively, if the participant hears a filler stimulus that contains an *-en* plural as the subject, they might also be primed to use an *-en* plural in a following target trial. To account for such a potential priming effect, we recorded the *Previous Plural* that was either produced or heard by the participant in the subject position. Table 2 summarizes the variables in the data set.

We analyzed the data according to three different hypotheses. The first hypothesis extends the hypothesis behind Experiment 1: Participants are less likely to choose the *-s* plural if it results in two consecutive stressed syllables, but only if there is no prosodic boundary between the plural and the following word. The second hypothesis reverses the causality: Participants are more likely to insert a prosodic boundary if the use of an *-s* plural results in two consecutive stressed syllables. The third hypothesis assumes that more prosodic boundaries are inserted if the stimuli contain a stress clash between noun and verb, i.e., regardless of which plural suffix is eventually used. It should be noted that these hypotheses are not mutually exclusive. Both directions of causality may play a role and the degrees to which the stimulus and the participant's plural selection affect the insertion of prosodic boundaries may vary between participants and stimuli. Our analysis will attempt to establish which hypothesis is most compatible with the data.

We tested these hypotheses in two statistical analyses: the *Plural Model*, in which *Plural* was the dependent variable, and the *Boundary Model*, in which *Prosodic Boundary* was the dependent variable. It is, of course, not possible to determine the causal relationship between variables by simply reversing the roles of predictor and dependent variable in different models.

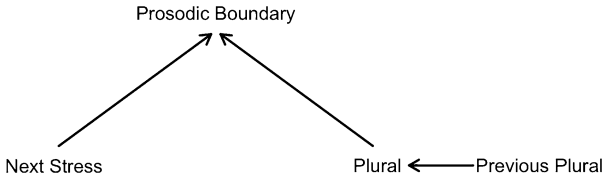

Fig. 2 Simplified causal diagram of the hypothesis that the presence of Prosodic Boundaries depends on which Plural is used and whether the next syllable is stressed or not. Which plural is produced depends on the Previous Plural that was used. See Appendix D for causal diagrams of the alternative hypotheses

However, given a causal model it is possible to generate certain predictions for both analyses. To illustrate this, consider the simplified (it omits the random variables) causal diagram in Fig. 2.

If the causality indicated by Fig. 2 is true, then a model with *Prosodic Boundary* as a dependent variable should show clear effects of (an interaction between) the *Next Stress* and *Plural* variables. On the other hand, if the model has *Plural* as a dependent variable, we would expect effects of (an interaction between) the *Next Stress* and *Prosodic Boundary* variables, due to something called *collider bias* (see McElreath, 2020). Applied to the causal relations between variables in Fig. 2, the collider works as follows: if we know that prosodic boundaries are inserted when the next syllable is stressed and the plural has an *-s* suffix, then given a prosodic boundary and following stress, we can guess that an *-s* plural was used. However, if the same *Plural Model* also includes other predictors that actually have a causal effect on which plural is used, such as *Previous Plural*, the collider effect of *Next Stress* and *Prosodic Boundary* has to compete with these predictors to explain the same variance. As a result, if the causality in Fig. 2 represents reality, the apparent effect of *Next Stress* and *Prosodic Boundary* in the *Plural Model* might be diminished, compared to the true causal effect of *Next Stress* and *Plural* in the *Boundary Model*.

If it turns out that the insertion of prosodic boundaries mostly depends on whether the stimulus contained a stress clash, then we should expect the *Plural* × *Next Stress* interaction in the *Boundary Model* to look more like a main effect of *Next Stress* as the produced plural suffix should not matter.

To measure the effects of the *Plural* × *Next Stress* and *Prosodic Boundary* × *Next Stress* interactions in their respective models, we compared each of these models to baseline models which did not include the respective interactions.

4.2 Results

4.2.1 Plural model

The baseline model of the *Plural* variable had a WAIC of 871.3, whereas the full *Plural Model* had a WAIC of 874.8, resulting in a difference of 3.5 with a standard error of 3.52. This means that, in terms of WAIC, including the *Prosodic Boundary* × *Next Stress* interaction resulted in a worse fitting model. The *Stressed – Unstressed* contrasts for the Next Stress predictor at the *Boundary* and *No Boundary* levels were calculated from the posterior distributions summarized in Table 7 in Appendix C. As

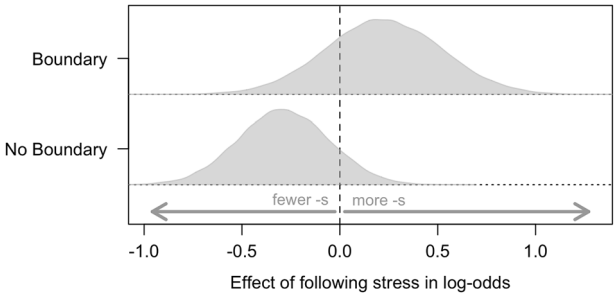

Fig. 3 Posterior distribution of the *Stressed – Unstressed* contrasts in the *Plural Model*. See running text for details

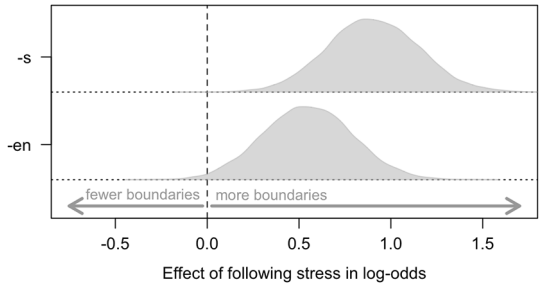

Fig. 4 Posterior distribution of the *Stressed – Unstressed* contrasts in the *Boundary Model from Experiment 2*. See running text for details

Fig. 3 illustrates, there is no clear difference in log-odds of an -s variant between the *Stressed* and *Unstressed* conditions, regardless of the presence of a *Prosodic Boundary*.

4.2.2 Boundary model

The baseline model of the plural variable had a WAIC of 889.2, whereas the full *Boundary Model* had a WAIC of 872.8, resulting in a difference of 16.5 with a standard error of 8.77. This means that, in terms of WAIC, including the *Plural* × *Next Stress* interaction resulted in a better fitting model. The *Stressed – Unstressed* contrasts for *-s* plural and *-en* plurals respectively were calculated from the posterior distributions summarized in Table 8 in Appendix C. As Fig. 4 illustrates, for both the *-s* plurals and the *-en* plurals, following stress resulted in more prosodic boundaries, although the effect was clearer for *-s* plurals.

5 Interim discussion 2

The results from the *Boundary Model* in Experiment 2 clearly show that participants inserted more prosodic boundaries if the plural was followed by verbs with initial stress, especially if the plural had an *-s* suffix. This suggests that participants may have inserted prosodic boundaries when their selection of the *-s* plural resulted in a stress clash.

Interestingly, following stress was also associated with more boundaries for the *-en* plurals, though to a lesser extent than the *-s* plurals. This finding is in line with the

hypothesis that, when the singular stimulus sentence contained a stress clash, sometimes participants planned to include a prosodic boundary before deciding which plural to use. These results support our suspicion that the design of our experiments stimulated participants to use prosodic boundary insertion as a strategy to prevent stress clashes, explaining why we did not find a morphological PRA effect in either experiment.

The question that remains after Experiment 2 is whether following stress does influence the choice between plural variants outside of the context of our experimental design. In order to investigate this possibility, we conducted a corpus study on the relationship between the PRA and variable plurals.

As discussed in Sect. 1.2, investigating the PRA using corpus studies has its difficulties, namely, determining directionality and determining the relationship between population-level effects and mechanisms in individual language users. Given a noun stem with final stress, it may be the case that one group of language users always use the *-s* plural, whereas another group always uses the *-en* plural. If the lexical or syntactic PRA mechanisms affect individuals in those groups, the distribution of *-s* and *-en* variants in the population-level data avoids consecutive stressed syllables. Likewise, such a stress clash avoiding distribution can be achieved if one group always uses the *-s* variant for a given noun and solves stress clashes using prosodic boundary insertion, and another group uses the *-en* variant and therefore never encounters stress clashes with following words. It follows that a population-level distribution of plural variants which shows rhythmic alternation is not sufficient evidence for a morphological PRA effect.

Importantly, for variable plurals, we can check for this potential confound by also considering invariable plurals. There is no reason why any of the alternative mechanisms should exclusively affect variable plurals. So, if a potential PRA effect on variable plurals is actually driven by the alternative mechanisms, we should observe a similar PRA effect on invariable plurals. Alternatively, if we observe no PRA effect on invariable plurals, but we do observe a PRA effect on variable plurals, then it is very likely that the morphological mechanism caused the effect.

The hypotheses and predictions for the corpus study are summarized in (7).

(7)
a. If there is no link between plural variant choice and following stress (the *Null Hypothesis*), there should be no association between following stress and the distribution of plural forms in either variable or invariable plurals.
b. If following stress affects the choice between plural variants (the *Morphological Hypothesis*), the proportion of *-s* forms in variable plurals should be lower before stressed syllables than before unstressed syllables (if there is no prosodic boundary between the plural and the next syllable), whereas there should be no association between following stress and the distribution of plural forms in invariable plurals.
c. If an alternative mechanism prevents stress clashes between variable plurals and following words (the *Alternative Hypothesis*), the proportion of *-s* forms in both variable and invariable plurals should be lower before stressed syllables than before unstressed syllables (if there is no prosodic boundary between the plural and the next syllable).

6 Corpus study

6.1 Materials and methods

6.1.1 Data

In order to reliably estimate the distribution of variable plurals we need a very large corpus. This constraint limited us to corpora that are primarily based on written language, as just using the available spoken language corpora would result in only a handful of observations for most variable plurals. Furthermore, to determine the stress pattern of the following word, we needed a corpus which contained the context of each word rather than a corpus consisting of just a word frequency list. The SoNaR corpus (Oostdijk et al., 2013), which contains 500 million words, meets these criteria and it is also easily accessible through the OpenSONAR web-interface. SoNaR comprises texts from various sources, such as internet pages, books, and the spoken language corpus CGN (Oostdijk, 2000).

We excluded some of the sources for the SoNaR corpus because we judged them unsuitable for our study. Firstly, we excluded the *Subtitles* sub-corpus because manual inspection raised suspicions that some of the subtitles had been machine-translated from other languages, resulting in very unnatural Dutch sentences. Secondly, *Chats, Discussion lists, SMS*, and *Tweets* were excluded because these sources often contained spelling mistakes, non-standard spellings *etc*., which would have made further processing very difficult.

In order to find variable plurals in the resulting subset of SoNaR, we used the list of noun lemmas from Zee and colleagues (2021) and selected those 69 lemmas that have final stress in the singular. For none of these lemmas, the two plural variants differ in their meanings. Searching SoNaR for sentences containing these variable plurals, while automatically excluding sentences that were written in a different language, resulted in 106 828 tokens. This dataset showed that many of the 69 noun lemmas were predominantly pluralized with either *-s* or *-en*. As variation in the plural is a prerequisite for finding an effect of the following stress, we further restricted our list of nouns to those with a proportion of *-s* variants above 0.05 and below 0.95. Additionally, we excluded any duplicate sentences, which resulted in a final dataset of 15683 tokens and 21 lemmas.

The nouns with invariable plurals were selected to match the nouns with variable plurals on a number of features that may affect their sensitivity to rhythm. We started by only selecting nouns that have final stress in their singular form. Subsequently, we selected nouns that approximately matched the nouns with variable plurals in terms of number of *Syllables*, *Plural Frequency*, and *Plural Proportion* (plural frequency relative to combined plural and singular frequency). With regard to *-s Proportion* (*-s* variant frequency relative to combined *-s* and *-en* variant frequency), we matched variable lemmas with an *-s Proportion* above 0.5 with invariable *-s* plurals, and we matched variable lemmas with a majority of *-en* variants with invariable *-en* plurals. All frequency measures were derived from the SUBTLEX corpus (Keuleers et al., 2010). Figure 5 visualizes these measures for the variable and invariable plurals.

We extracted all tokens of the 21 selected invariable plurals from the SoNaR corpus (excluding the *Chats, Discussion lists, SMS*, and *Tweets* subcorpora), excluded

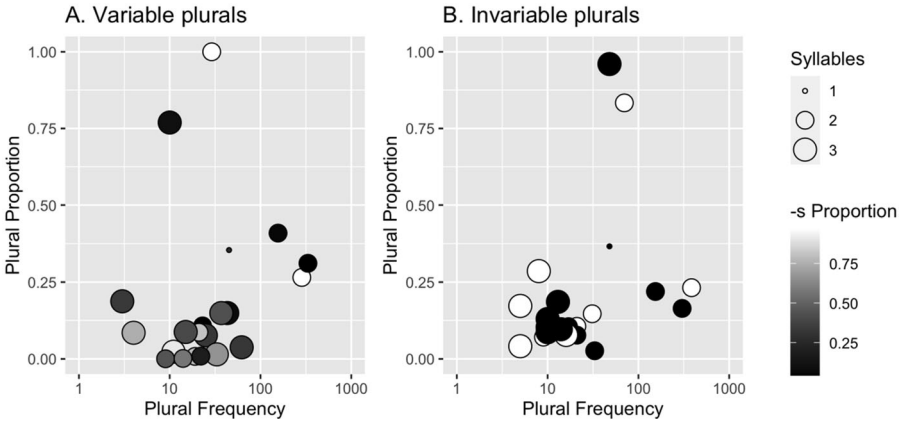

Fig. 5 Visualization of the measures on which the invariable plurals (in panel B) were matched to the variable plurals (in panel A). Each circle represents a lemma. The size of the circle represents the number of syllables in the lemma. The circle's position along the *x*-axis represents the plural frequency (the combined frequency of the *-s* and *-en* variants). The shade of the circle represents the *-s* variant frequency relative to the combined plural frequency. The circle's position along the *y*-axis represents the plural frequency relative to combined plural and singular frequency

tokens in languages other than Dutch, and removed duplicate tokens. This procedure resulted in a final dataset of 26700 observations.

6.1.2 Variables

Tables 3 and 4 summarize the overall distribution of the dependent and independent variables in the variable and invariable data respectively. The dependent variable of this study is *Plural*, which has the value of 1 for *-s* variant tokens and the value of 0 for *-en* variant tokens. The first predictor of interest is *Next Stress* which has two levels: *Stressed* for plurals preceding a stressed syllable, and *Unstressed* for plurals preceding an unstressed syllable. We automatically coded this variable by looking up the canonical pronunciation of the word following the plural in the CELEX database (Baayen et al., 1996). The second predictor of interest is the *Prosodic Boundary* variable, which indicated whether a prosodic boundary separated the plural noun and the following word. As punctuation marks often cooccur with strong prosodic boundaries (Ingulfsen, 2004), nouns preceding a punctuation mark were coded *Boundary*, and all other nouns were coded *No Boundary*. Although punctuation marks, and the syntactic boundaries they represent, do not perfectly match prosodic boundaries (e.g., Bennett & Elfner, 2019), previous research on Dutch has shown that orthography-derived boundary annotations outperform those based on more sophisticated features and algorithms (Marsi et al., 2003). For variable plurals, the *Morphological Hypothesis* expects *Next Stress* and *Prosodic Boundary* to interact such that *-s* plurals occur less frequently before a *Stressed* syllable than before an *Unstressed* syllable but only if there is *No Boundary* between the plural and the following syllable. In contrast, for invariable plurals, the *Morphological Hypothesis* predicts that there is no effect of Next Stress, neither as a simple effect nor in the *Prosodic Boundary* × *Next Stress*

Table 3 Distribution of the variables in the variable plurals corpus dataset

Dependent variable	Levels	Token numbers
Plural	1 (*-s*)	9154
	0 (*-en*)	6529
Fixed variables		
Next Sound	[s] / [z]	1385
	Vowels	4525
	Consonants	9773
Next Stress	Stressed	13031
	Unstressed	2652
Prosodic Boundary	Boundary	2581
	No Boundary	13102
Random variables	Number of levels	Examples
Lemma	21	admiraal, ballon, baron, etc.
Style	19	auto cues, blogs, books, etc.

Table 4 Distribution of the variables in the invariable plurals corpus dataset

Dependent variable	Levels	Token numbers
Plural	1 (*-s*)	14368
	0 (*-en*)	12089
Fixed variables		
Next Sound	[s] / [z]	1301
	Vowels	8637
	Consonants	15519
Next Stress	Stressed	21668
	Unstressed	4789
Prosodic Boundary	Boundary	5211
	No Boundary	21246
Random variables	Number of levels	Examples
Style	19	auto cues, blogs, books, etc.

interaction. In other words, the *Morphological Hypothesis* predicts a three-way interaction of *Prosodic Boundary* × *Next Stress* × *Variability*, where *Variability* has the levels *Variable* and *Invariable*.

In addition to the predictors of interest, a number of covariates were included to adjust for potential confounds. The first potential confound involves the *Next Sound* variable. Previous research (e.g., Breiss & Hayes, 2020) has reported that speakers tend to avoid consecutive sibilants. As a consequence, a language user may be more likely to choose *-en* rather than *-s* if the following word starts with a sibilant. Likewise, a following vowel might reduce the probability of an *-en* variant as this suffix

is usually pronounced as [ə], which could lead to a vowel hiatus. Since the stress position and the initial sound of the following word are likely to cooccur to some extent (for example, many words starting with the /s/ sound may have initial stress), the effects of these two features may be confounded. We based the variable *Next Sound* on CELEX transcriptions, distinguishing three sound categories: *Sibilant, Vowel*, or *Consonant*. As we did not know whether a potential effect of *Next Sound* would be mediated by the presence of a prosodic break, we included an interaction between *Next Sound*, *Prosodic Boundary*, and *Variability* in our regression model.

A second potential confound involves the *Style* variable, which groups the data into 19 levels that were derived from the subcorpus identifiers in SoNaR: *e-magazines, periodicals/magazines, newspapers, policy documents, press releases, web sites, wikipedia, proceedings, reports, teletext pages, books, auto cues, texts for the visually impaired, brochures, newsletters, guides/manuals, legal texts, blogs*, and *written assignments*. Previous research has claimed that *-en* plural variants are preferred in more formal speech registers (Baayen et al., 2002). Given that claim, it is not unthinkable that more formal writing *Styles* also elicit the *-en* plural for nouns with variable plurals. Furthermore, writing style certainly affects the type of words that are used. As such, the lexical stress patterns of the words following the plurals may differ across different writing styles. We therefore included *Style* as a random variable in the regression analysis with random slopes for *Variability*.

The final variable we included in some of the regression analyses is *Lemma*, which encodes the singular form of each plural. Including *Lemma* as random variable allows for the investigation of lemma-specific effects and for better predictions, which results in more accurate parameter estimates for the predictors of interest (*Next Stress* and *Prosodic Boundary*).

6.1.3 Modelling

As in Experiments 1 and 2, we used Bayesian multi-level logistic regression to model the *Plural* variable. Again, predictors were entered into the models as index variables rather than dummy variables, as this results in more sensible assumptions about the priors (McElreath, 2020), and we used mildly regularizing priors for all model parameters. We modelled the combined variable and invariable plural data using three models corresponding to the three hypotheses. The *Null model* contained the *Next Sound* × *Prosodic Boundary* × *Variability* interaction and by-*Style* random intercepts for *Variability*. The *Alternative model* contained all parameters of the *Null model* and added the *Next Stress* × *Prosodic Boundary* interaction. To find out whether distinguishing between variable and invariable plurals with regard to the *Next Stress* × *Prosodic Boundary* interaction improved model predictions, we also fitted the *Morphological model*. This model expanded on the *Alternative model* by adding the *Next Stress* × *Prosodic Boundary* × *Variability* interaction. We again compared the models using WAIC to see which model best fits the data.

Subsequently, we interpreted the *Next Stress* × *Prosodic Boundary* × *Variability* interaction by using the posterior distribution to compute contrasts between the relevant conditions. For the variable and invariable plurals, respectively, *Stress*/*Unstressed* contrasts in the *No Boundary* condition were calculated by subtracting the posterior distribution of the *No Boundary-Unstressed* parameter estimate

from the posterior distribution of the *No Boundary-Stressed* parameter estimate. Likewise, the *Stress/Unstressed* contrasts were calculated for the *Boundary* condition by performing the same operation for the *Boundary-Unstressed* and *Boundary-Stressed* conditions.

Finally, we fitted two more models to just the variable plural dataset in order to investigate whether any potential effects of the PRA on plural variant distribution were shared by all lemmas or limited to only a few lemmas. Including the *Lemma* variable only makes sense for the variable dataset because, for the invariable dataset, including *Lemma* would be the same as incorporating the correct plural morpheme as a predictor, which would minimize any effect of the other predictors. A *Baseline Model* was fitted which contained the *Next Stress* × *Prosodic Boundary* interaction, the *Next Sound* × *Prosodic Boundary* interaction and random intercepts for both *Style* and *Lemma*. The *Lemma-specific Model* expanded on the *Baseline Model* by adding by-*Lemma* random intercepts for the *Next Stress* × *Prosodic Boundary* interaction. We then compared the models using WAIC to see whether estimating lemma-specific effects of the prosodic context resulted in a better model.

6.2 Results

6.2.1 Combined dataset

Model comparison of the three hypothesis-based models showed that the *Morphological model* performed best with a WAIC of 54539.8, followed by the *Alternative model* (WAIC of 54562.8) and the *Null model* (WAIC of 54567.2). The WAIC differences between the *Morphological* model and the other models (23.0 and 27.4, respectively) were relatively robust, as evidenced by the estimated Standard Errors of those differences at 10.22 and 11.80 respectively. This means that adding the *Next Stress* × *Prosodic Boundary* × *Variability* interaction resulted in a better fitting model. The *Stressed/Unstressed* contrasts in the *Boundary* and *No Boundary* conditions for the variable and invariable plurals were calculated from the posterior distributions summarized in Table 9 in Appendix C. As Fig. 6 illustrates, for the *Invariable* plurals, there is no difference between the *Stressed* and *Unstressed* conditions regardless of whether a prosodic boundary is present. For the *Variable* plurals, in contrast, if there is no prosodic boundary separating the plural and the following word, the model estimates that following stress does reduce the log-odds of an *-s* variant.

Posterior distributions of the *Null* and *Alternative model* are summarized in Table 10 in Appendix C.

6.2.2 Variable plurals

The *baseline model* of the variable plurals had a WAIC of 11680.3, whereas the *Lemma-specific Model* had a WAIC of 11679.2, resulting in a difference of 1.1 with a standard error of 15.6. This means that, in terms of WAIC, allowing the *Next Stress* × *Prosodic Boundary* interaction to vary among lemmas hardly improves predictive accuracy. In the *Lemma-specific Model*, the fixed Stresses/Unstressed contrast in the *No Boundary* condition no longer shows a clear effect, which suggests that the rhythmic effect is not shared by all lemmas. Inspecting the by-Lemma contrasts in Fig. 7

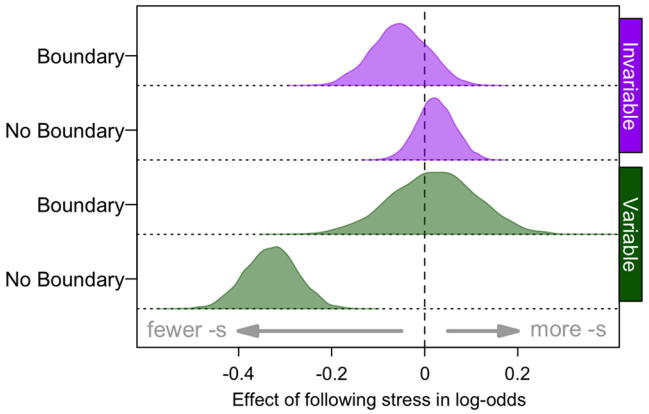

Fig. 6 Posterior distributions of the *Stressed/Unstressed* contrasts in the *Morphology Model* of the combined corpus data. Separate contrasts are defined for the *Boundary* and *No Boundary* conditions in the *Variable* and *Invariable* plurals respectively

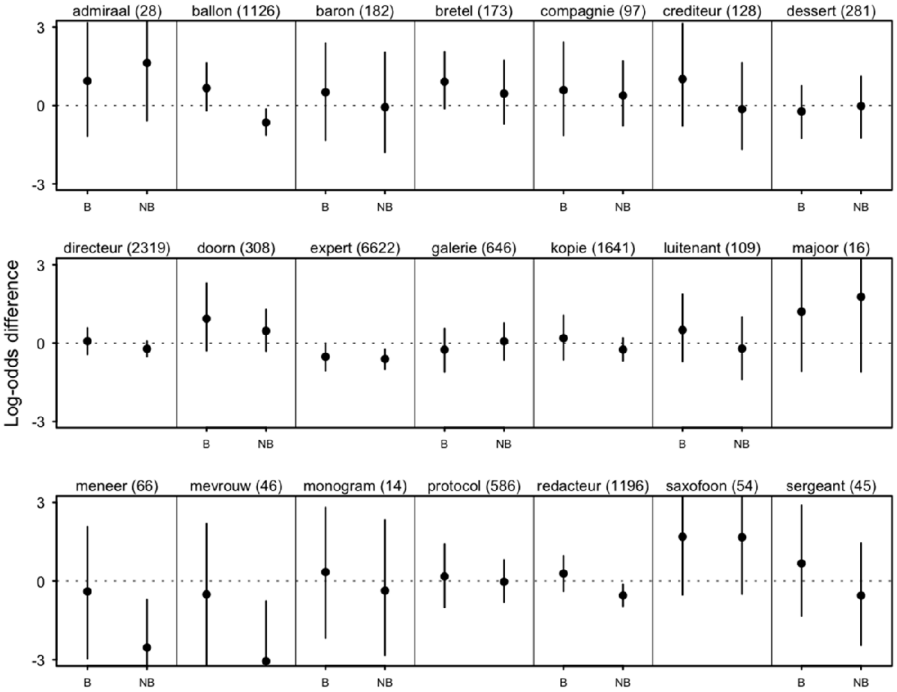

Fig. 7 By-lemma random contrasts between stressed and unstressed contexts in both the boundary (B) and no boundary (NB) conditions of the *Lemma-specific Model*. The y-axis reflects the effect of following stress on the log-odds of *-s* variant productions. The number of plural observations of each lemma is included in parentheses. The lines represent 95% credible intervals of the contrasts

confirms that the negative effect of following stress in the *No Boundary* condition affects some lemmas (e.g., *ballon*, *redacteur*) more than others (e.g., *dessert*, *galerie*). One reason why the WAIC improvement may be so small, even though there is clear variation in the effect of following stress, is that almost all high frequency plurals do show the effect, see Fig. 7.

7 Discussion

In this research, we investigated whether the presence of following stress affected the choice between Dutch plural variants that differ with respect to the stress position, e.g., *vampíérs* (final stress) vs. *vampíéren* (penultimate stress). In doing so, we studied the effect of the *Principle of Rhythmic Alternation* (PRA) on the distribution of morphological alternatives. Dutch variable plurals were chosen because they had not been investigated before in light of the PRA and because they enabled us to distinguish between different mechanisms behind such an effect. Specifically, this research investigated whether a PRA effect on Dutch variable plurals is best explained by a prosodic effect on the competition between morphological variants or by alternative mechanisms that do not involve morpho-prosodic processing. One alternative explanation involves a structural PRA effect: the selection of a plural variant affects the structure of its context (i.e., word order) such that the PRA is observed. Another alternative explanation entails a lexical PRA effect, in which a plural variant competes with rhythmically different near-synonyms for selection. To establish which mechanism most likely explains a potential PRA effect on Dutch variable plurals, we conducted two experiments and a corpus study.

In Experiment 1, we investigated whether the stress pattern of the following word affected the choice of plural variants in a writing task. We excluded alternative lexical and structural PRA mechanisms by allowing participants to choose only between plural variants, not between structures of the sentence or between lemmas. Participants were asked to come up with the plural version of sentences containing singular noun subjects. We hypothesized that participants would avoid stress clashes by using *-s* plural variants less often preceding words with initial stress. However, the data showed no difference in the distribution of plural variants before stressed and unstressed syllables, which suggests that rhythm did not play a role in the participants' choice between plural variants. We suspected that the lack of an effect may have been the result of participants inserting prosodic boundaries between the plural and the following word, thereby removing the need to prevent stress clash through plural selection. This suspicion could not be confirmed because the boundaries would have occurred in a sentence position that does not allow for punctuation.

Experiment 2 was almost identical to Experiment 1, but it required participants to respond through speech. The change in modality allowed us to check whether participants prevented stress clashes by inserting prosodic boundaries instead of choosing an *-en* plural. Generally, this hypothesis was supported by the results, which explains the lack of a PRA effect in Experiment 1. Interestingly, participants also inserted boundaries between *-en* plurals and subsequent stressed syllables, although to a lesser extent

than between *-s* plurals and subsequent stressed syllables. This suggests that the presence of stress clashes between the nouns and the following words in the stimuli also triggered participants to insert prosodic boundaries.

Taken together, the experimental results suggest that the design of our two experiments probably prevented a prosodic effect on morphological variation from occurring regardless of whether such an effect exists. To investigate whether the PRA affects plural variation outside of an experimental context, we conducted a corpus study. Specifically, we analysed the distribution of plural variants of nouns with final stress in a large text corpus consisting of websites, books, and other types of text. We hypothesized that *-s* variants would be less likely to occur preceding a stressed syllable but only if no punctuation marks separated the plural and the stressed syllable, as a punctuation mark would indicate a prosodic boundary between the variable plural and the following word. This prediction was supported by the analysis of the corpus data, which contained 15683 variable plural tokens of 21 noun types. As such, this is the first study to show that Dutch variable plurals are involved in a PRA effect. Furthermore, it shows that the interaction between the PRA and the presence of prosodic boundaries found for English (Schlüter, 2005; Azzabou-Kacem, 2018; Breiss & Hayes, 2020) also applies to Dutch, thereby underlining the prosodic nature of the effect.

By analyzing variable plurals together with invariable plurals, we investigated whether the distribution of variable plurals was likely to be the result of a rhythmic effect on morphological alternation rather than of lexical or structural variation or of prosodic boundary insertion. We hypothesized that, if the PRA effect on variable plurals isdue to morphological alternation, the distribution of *-s* and *-en* invariable plurals should not be associated with the presence of following stress. By contrast, if the avoidance of stress clash was achieved through prosodic boundary insertion or a choice between different noun lemmas or sentence structures, invariable plurals should show a PRA effect as well. The invariable plural data, which consisted of 26700 plural tokens of 21 noun types, showed no effect of following stress, suggesting that the PRA effect for the variable plurals arises through competition between the plural variants of a lemma. This result constitutes the first evidence for a morphological PRA effect that takes alternative mechanisms into account. As such, it supports previous claims about the relevance of PRA effects for models of morphological processing (e.g., Schlüter, 2015).

The evidence for a morphological PRA effect in our corpus study reinforces the idea that the lack of PRA effects on plural variation in the experimental studies is likely due to some aspect of the experimental design promoting prosodic boundary insertions. In our experiments, boundary insertion may have been encouraged by the fact that participants first read or heard a sentence with the noun in singular form before producing a plural version. This meant that participants were presented with a stress clash for all stimuli in which the singular noun was followed by an initially-stressed verb. Participants may have inserted a prosodic boundary in those stimuli to resolve the stress clash when they processed those sentences (see Kentner, 2015 for a similar effect on accentuation). When they subsequently pluralized the sentence, they may have kept the prosodic boundary in place, which prevented the stress of the word following the boundary from playing any role in the selection of the plural variant.

A second explanation of the experimental results assumes that the PRA only affects a certain type of psychological process. Certain psycho-linguistic models (e.g., Baayen et al., 1997) suggest that morphologically complex words can be produced by means of two distinct processing routes. In the retrieval route, mental representations of complex words compete with each other for selection, e.g., *vampiers* vs. *vampieren*. In the compositional route, the complex word is composed from the base representation according to competing morphological generalisations, e.g., *vampier* + *-en* vs. *vampier* + *-s*. The current experimental paradigm may not allow us to measure an effect on the retrieval route because it encourages a compositional process by explicitly asking the participants to turn a singular noun into a plural noun. As such, the experimental null result can also be explained if we assume that the PRA only affects variable plurals processed by means of the retrieval route.

More research is needed to find out whether morphological PRA effects mostly result from word form representations with similar meanings competing for selection (the retrieval route) or also from some compositional mechanism. Interactions of morphological PRA effects with predictors that reflect the representational strengths of the morphological doublets could be inspected to find out which of these options is more likely. For instance, a high token frequency has been associated with a stronger representation of a word form (Bybee, 1995) and a similar association has been found for Dutch variable plurals (Zee et al., 2021). Hence, if rhythm only affects alternation between morphological variants with high frequencies, this would suggest that strong word form representations are a requirement for a morphological PRA effect to emerge. We looked at such an interaction effect *post hoc* in the lemma-specific analysis of our corpus study. This analysis suggested that it was mostly plurals with low token frequency that did not exhibit the PRA effect. These results are in line with the hypothesis that the PRA effect mainly affects competition between representations of the respective plural variants. However, more research is needed to provide robust evidence for such a claim, especially because the results of our experimental study can also be explained by participants using prosodic boundaries to avoid stress clashes.

Regardless of the exact psychological mechanism that is at play in the morphological PRA effect on Dutch variable plurals, our results have implications for psycholinguistic models of morphological processing. Whether a model describes the production of a variable plural form as the result of a selection process involving the lexical representations of complete plural forms or as the result of a composition process involving a stem and plural suffix, in either case this process must be open to influences from the rhythmic properties of the plural form and the following word. This means that production models cannot be strictly serial in the sense that inflected word forms are selected or composed before they are prosodically integrated (e.g., Levelt et al., 1999; see overview in Wheeldon & Konopka, 2018). Moreover, production models that are not strictly serial (e.g., Dell et al., 2014) need to be expanded to implement the influence of rhythm on selection or composition of an inflected word.

In sum, our research illustrated that a more substantive and well-supported psycholinguistic account of morphological PRA effects can be established by investigating PRA effects on known morphological alternations in contemporary language use. By showing that the production of Dutch variable plurals is affected by the presence

of following stress, we identified a new morphological alternation that is affected by the PRA. Furthermore, this research argued for the importance of considering alternative causal mechanisms behind a given PRA effect and provided a strategy to do so in corpus studies. Using this strategy, we showed that the PRA effect on Dutch variable plurals likely involves a prosodic effect on morphological rather than lexical or syntactic processing, thereby providing evidence for previous claims about the psycholinguistic relevance of morphological PRA effects. We hope that future research builds on our findings to further uncover the exact mechanisms that underly morphological PRA effects.

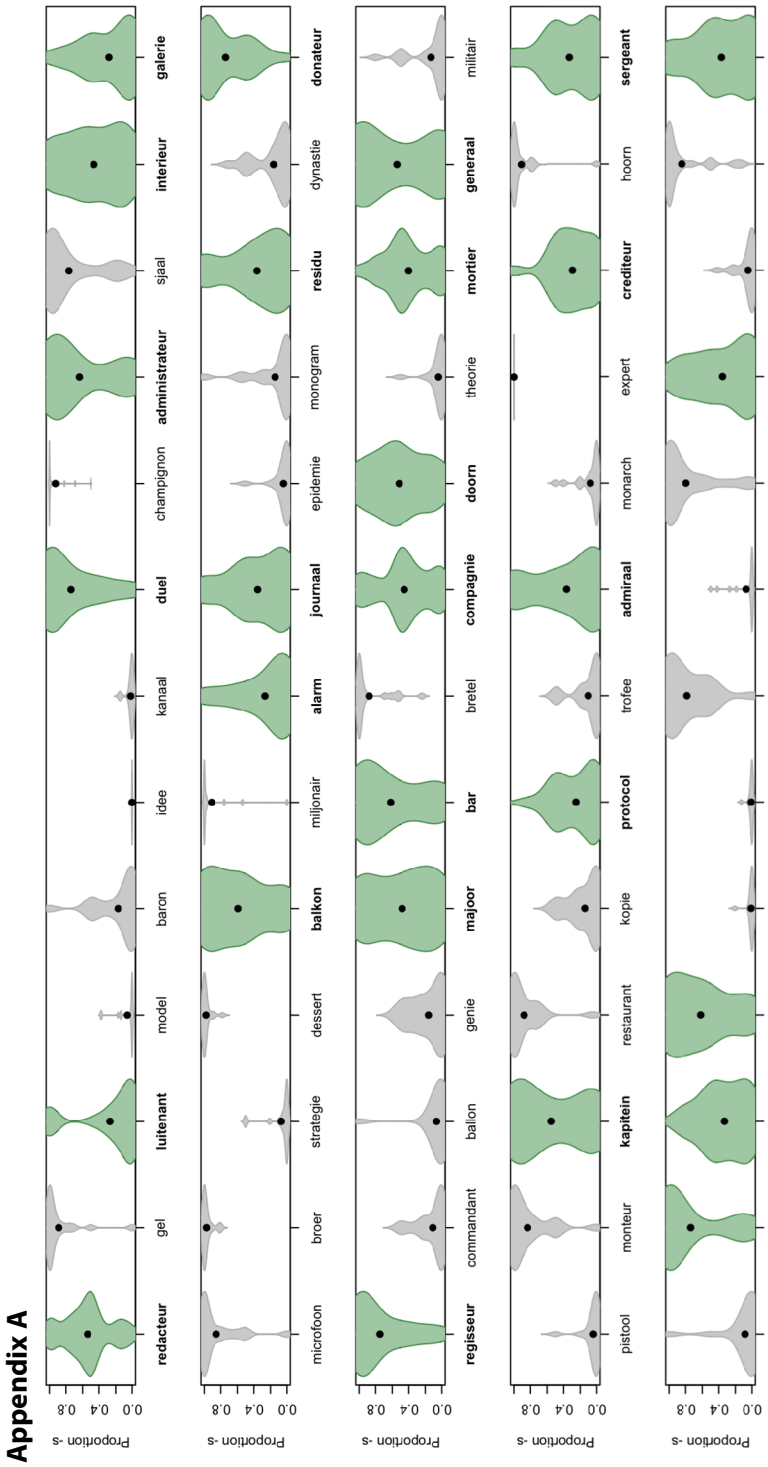

Fig. 8 Ratings of variable plurals in pilot study. Filled areas represent the density of ratings, and black dots represent the average ratings. Nouns with green-shaded areas and labels in bold font were selected for use in the experimental study. (Color figure online)

Appendix B

Table 5 Experimental Lists 1a and 1b

Noun	Type	Change to	List 1a	List 1b
priester	filler	singular	De katholieke priesters wijden hun leven aan God.	De katholieke priesters wijden hun leven aan God.
receptionist	filler	singular	De receptionisten hielden hun pauze om 12 uur.	De receptionisten hielden hun pauze om 12 uur.
soldaat	filler	singular	De soldaten pakten hun waterpistool.	De soldaten pakten hun waterpistool.
varken	filler	plural	Het varken spitste zijn oren.	Het varken spitste zijn oren.
crediteur	target	plural	De crediteur verwachtte de betaling.	De crediteur wachtte de betaling af.
agent	filler	singular	De agenten trokken hun dienstwapen.	De agenten trokken hun dienstwapen.
kassamedewerker	filler	singular	De kassamedewerkers openden hun kassa's.	De kassamedewerkers openden hun kassa's.
hengst	filler	singular	De hengsten kregen hun voer in de stal.	De hengsten kregen hun voer in de stal.
duel	target	plural	Het duel liep teleurstellend af.	Het duel verliep teleurstellend.
protocol	target	plural	Het protocol deelde de taken in 4 stappen op.	Het protocol verdeelde de taken in 4 stappen.
interieur	target	plural	Het interieur mengde een moderne stijl met traditionele meubels.	Het interieur vermengde een moderne stijl en traditionele meubels.
doktersassistent	filler	singular	De doktersassistenten haten hun leidinggevende.	De doktersassistenten haten hun leidinggevende.
crimineel	filler	singular	De criminelen hadden geen spijt van hun misdaden.	De criminelen hadden geen spijt van hun misdaden.
diva	filler	singular	De filmdiva's begonnen hun loopbaan als danseres.	De filmdiva's begonnen hun loopbaan als danseres.
sergeant	target	plural	De sergeant bedankte zijn troepen voor hun inzet.	De sergeant dankte zijn troepen voor hun inzet.
schoonheidsspecialist	filler	singular	De schoonheidsspecialisten kochten hun producten bij de dierenwinkel.	De schoonheidsspecialisten kochten hun producten bij de dierenwinkel.
dochter	filler	singular	De dochters vertelden alles aan hun ouders.	De dochters vertelden alles aan hun ouders.
officier	target	plural	De officier mijdt de blik van zijn leidinggevende.	De officier vermijdt de blik van zijn leidinggevende.
tante	filler	singular	De tantes geven hun nichtje 5 euro zakgeld.	De tantes geven hun nichtje 5 euro zakgeld.
redacteur	target	plural	De redacteur twijfelt sinds kort aan de betrouwbaarheid van haar bronnen.	De redacteur betwijfelt sinds kort de betrouwbaarheid van haar bronnen.

Table 5 (*Continued*)

Noun	Type	Change to	List 1a	List 1b
mondhygiënist	filler	singular	De mondhygiënisten geven hun patiënten advies over gordijnen.	De mondhygiënisten geven hun patiënten advies over gordijnen.
piloot	filler	plural	De piloot stelde zijn collega's gerust.	De piloot stelde zijn collega's gerust.
journaal	target	plural	Het journaal lichtte de gebeurtenissen toe.	Het journaal belichtte de gebeurtenissen.
luitenant	target	plural	De luitenant berekende de uitgaven van zijn eenheid.	De luitenant rekende de uitgaven van zijn eenheid uit.
bouwvakker	filler	singular	De bouwvakkers haalden hun speelgoed uit de bus.	De bouwvakkers haalden hun speelgoed uit de bus.
loodgieter	filler	singular	De loodgieters gebruikten hun toverstaf.	De loodgieters gebruikten hun toverstaf.
schrijver	filler	plural	De schrijver vertelde over haar boeken.	De schrijver vertelde over haar boeken.
directeur	target	plural	De directeur begroette de investeerders.	De directeur groette de investeerders.
kater	filler	singular	De katers markeren hun territorium.	De katers markeren hun territorium.
compagnie	target	plural	De compagnie bewaakte sinds 2005 de grensovergang.	De compagnie waakte sinds 2005 over de grensovergang.
baviaan	filler	plural	De baviaan liet zijn dagboek zien.	De baviaan liet zijn dagboek zien.
bewaker	filler	plural	De bewaker liep haar laatste rondje.	De bewaker liep haar laatste rondje.
non	filler	singular	De nonnen zeiden hun gebed voor het avondmaal.	De nonnen zeiden hun gebed voor het avondmaal.
zeemeermin	filler	singular	De zeemeerminnen lagen met hun staart op de rots.	De zeemeerminnen lagen met hun staart op de rots.
administrateur	target	plural	De administrateur werkte alle bonnetjes weg.	De administrateur verwerkte alle bonnetjes.
model	filler	singular	De modellen deden hun werk op de modeshow.	De modellen deden hun werk op de modeshow.
personage	filler	plural	Het personage veroorzaakt haar eigen dood.	Het personage veroorzaakt haar eigen dood.
piraat	filler	plural	De piraat begroef zijn laptop.	De piraat begroef zijn laptop.
jongen	filler	singular	De jongens maakten hun huiswerk.	De jongens maakten hun huiswerk.
ballerina	filler	singular	De ballerina's oefenden hun danspasjes.	De ballerina's oefenden hun danspasjes.
perron	target	plural	Het perron grensde zowel aan trein- als metrosporen.	Het perron begrensde zowel trein- als metrosporen.

Table 5 (*Continued*)

Noun	Type	Change to	List 1a	List 1b
brandweerman	filler	singular	De brandweermannen trokken hun vuurbestendige pakken aan.	De brandweermannen trokken hun vuurbestendige pakken aan.
regisseur	target	plural	De regisseur bekeek altijd eerst de belichting van de scène.	De regisseur keek altijd eerst naar de belichting van de scène.
fotograaf	filler	plural	De fotograaf belichtte haar onderwerp.	De fotograaf belichtte haar onderwerp.
astronaut	filler	singular	De astronauten legden hun onderzoek uit.	De astronauten legden hun onderzoek uit.
haarstilist	filler	singular	De haarstilisten moesten hun feesthoedje ophouden.	De haarstilisten moesten hun feesthoedje ophouden.
haan	filler	singular	De hanen beschermden hun kippen.	De hanen beschermden hun kippen.
majoor	target	plural	De majoor kende die verslagen houding van zijn mannen.	De majoor herkende die verslagen houding van zijn mannen.
kuiken	filler	plural	Het kuiken opende zijn cadeau.	Het kuiken opende zijn cadeau.
houthakker	filler	plural	De houthakker gebruikte zijn knuffel elke dag.	De houthakker gebruikte zijn knuffel elke dag.
glazenwasser	filler	singular	De glazenwassers zaten om hun middel vast aan een tomaat.	De glazenwassers zaten om hun middel vast aan een tomaat.
wolvin	filler	plural	De wolvin bleef dicht bij haar vriendinnen.	De wolvin bleef dicht bij haar vriendinnen.
bokser	filler	singular	De boksers strikten hun veters.	De boksers strikten hun veters.
oma	filler	singular	De oma's verwennen hun kleinkinderen.	De oma's verwennen hun kleinkinderen.
mortier	target	plural	De mortier plofte met een enorme knal uit elkaar.	De mortier ontplofte met een enorme knal.
stier	filler	singular	De stieren richtten hun horens op de matador.	De stieren richtten hun horens op de matador.
verloskundige	filler	singular	De verloskundigen moeten tijdens hun nachtdienst altijd dronken zijn.	De verloskundigen moeten tijdens hun nachtdienst altijd dronken zijn.
vader	filler	singular	De vaders kregen verlof voor de geboorte van hun kind.	De vaders kregen verlof voor de geboorte van hun kind.
kapitein	target	plural	De kapitein belandde vanwege de storm in Engeland.	De kapitein landde vanwege de storm in Engeland.
prinses	filler	singular	De prinsessen lieten hun jurk op maat maken.	De prinsessen lieten hun jurk op maat maken.
ridder	filler	singular	De ridders trokken hun zwaard.	De ridders trokken hun zwaard.

Table 5 (*Continued*)

Noun	Type	Change to	List 1a	List 1b
balkon	target	plural	Het balkon brokkelde al een tijdje af.	Het balkon verbrokkelde al een tijdje.
galerie	target	plural	De galerie vertoonde de beste werken van Rembrandt.	De galerie toonde de beste werken van Rembrandt.
voetballer	filler	singular	De voetballers oefenden hun zangtechniek.	De voetballers oefenden hun zangtechniek.
moeder	filler	singular	De moeders werden gek van hun drukke kinderen.	De moeders werden gek van hun drukke kinderen.
admiraal	target	plural	De admiraal noemde de onderbevelhebber als zijn opvolger.	De admiraal benoemde de onderbevelhebber tot zijn opvolger.
donateur	target	plural	De donateur verkoos het WNF boven Greenpeace.	De donateur koos het WNF in plaats van Greenpeace.
doorn	target	plural	De doorn bekraste bij het inparkeren de nieuwe auto.	De doorn kraste bij het inparkeren langs de nieuwe auto.
prins	filler	singular	De prinsen zwaaiden naar hun onderdanen.	De prinsen zwaaiden naar hun onderdanen.
bar	target	plural	De bar eindigde het deurbeleid.	De bar beëindigde het deurbeleid.
alarm	target	plural	Het alarm verjaagde de onervaren inbrekers.	Het alarm jaagde de onervaren inbrekers weg.
meisje	filler	singular	De meisjes bestelden hun cadeaus online.	De meisjes bestelden hun cadeaus online.
gazon	target	plural	Het gazon beperkte de uitbreiding van de parkeergarage.	Het gazon perkte de uitbreiding van de parkeergarage in.
generaal	target	plural	De generaal kondigde zijn nieuwe plannen aan.	De generaal verkondigde zijn nieuwe plannen.
residu	target	plural	Het residu verbindt zich niet met water.	Het residu bindt zich niet aan water.
vampier	target	plural	De vampier verkrijgt zijn energie door bloed te drinken.	De vampier krijgt zijn energie door bloed te drinken.
hacker	filler	singular	De hackers testen hun code.	De hackers testen hun code.
babysitter	filler	singular	De babysitters vonden hun werk erg makkelijk.	De babysitters vonden hun werk erg makkelijk.
verpleegkundige	filler	singular	De verpleegkundigen hielpen hun patiënten bij het darten.	De verpleegkundigen hielpen hun patiënten bij het darten.
gravin	filler	plural	De gravin verloor haar dinosaurus.	De gravin verloor haar dinosaurus.

Appendix C

Table 6 Posterior distribution of the PRA model in the first experimental study. All estimates are in log-odds of an *-s* variant. Estimates for individual levels of the Lemma and Participant variables are omitted for the sake of brevity

Fixed Parameters	Mean	Standard deviation	95% interval lower bound	95% interval upper bound
Stressed	0.09	0.26	−0.42	0.59
Unstressed	0.08	0.26	−0.42	0.59
Contrast				
Stressed – Unstressed	0.01	0.10	−0.20	0.21
Random Parameters				
Lemma Standard Deviation	1.17	0.18	0.88	1.57
Participant Standard Deviation	1.31	0.13	1.07	1.59

Table 7 Posterior distribution of the model of the plural data from Experiment 2. All estimates are in log-odds of an *-s* variant. Estimates for individual levels of the Lemma and Participant variables are omitted for the sake of brevity

Fixed Parameters	Mean	Standard deviation	95% interval lower bound	95% interval upper bound
Boundary: Stressed	0.04	0.46	−0.85	0.94
Boundary: Unstressed	−0.18	0.47	−1.12	0.76
No Boundary: Stressed	−0.02	0.45	−0.92	0.86
No Boundary: Unstressed	0.27	0.45	−0.61	1.15
Previous Plural: -s	0.26	0.47	−0.68	1.19
Previous Plural: -en	−0.14	0.48	−1.10	0.79
Contrasts				
Boundary: Stressed – Unstressed	0.22	0.30	−0.36	0.80
No Boundary: Stressed – Unstressed	−0.29	0.23	−0.74	0.17
Random Parameters				
Lemma Standard Deviation	1.22	0.21	0.87	1.68
Participant Standard Deviation	1.56	0.23	1.16	2.08

Table 8 Posterior distribution of the model of the boundary data from Experiment 2. All estimates are in log-odds of a prosodic boundary. Estimates for individual levels of the Lemma and Participant variables are omitted for the sake of brevity

Fixed Parameters	Mean	Standard deviation	95% interval lower bound	95% interval upper bound
-s: Stressed	−0.28	0.30	−0.86	0.33
-s: Unstressed	−1.18	0.30	−1.78	−0.58
-en: Stressed	−0.23	0.30	−0.82	0.37
-en: Unstressed	−0.77	0.31	−1.38	−0.16
Contrasts				
-s: Stressed – Unstressed	0.90	0.25	0.42	1.40
-en: Stressed – Unstressed	0.55	0.24	0.08	1.03
Random Parameters				
Lemma Standard Deviation	0.76	0.15	0.50	1.09
Participant Standard Deviation	1.52	0.23	1.13	2.02

Table 9 Posterior distribution of the Morphological model of the combined corpus data. All estimates are in log-odds of an *-s* variant. Estimates for individual levels of the Style variable are omitted for the sake of brevity

Fixed Parameters	Mean	Standard deviation	95% interval lower bound	95% interval upper bound
Invariable – Boundary – Sibilant	−0.08	0.47	−0.84	0.67
Invariable – Boundary – Vowel	0.17	0.47	−0.58	0.92
Invariable – Boundary – Consonant	0.04	0.47	−0.73	0.78
Invariable – No Boundary – Sibilant	0.21	0.46	−0.54	0.92
Invariable – No Boundary – Vowel	−0.16	0.45	−0.91	0.55
Invariable – No Boundary – Consonant	0.12	0.45	−0.62	0.83
Variable – Boundary – Sibilant	0.37	0.49	−0.40	1.15
Variable – Boundary – Vowel	−0.14	0.48	−0.90	0.62
Variable – Boundary – Consonant	−0.20	0.48	−0.95	0.59
Variable – No Boundary – Sibilant	0.44	0.47	−0.32	1.18
Variable – No Boundary – Vowel	−0.06	0.46	−0.82	0.67
Variable – No Boundary – Consonant	−0.19	0.46	−0.94	0.55
Invariable – Boundary – Stressed	0.01	0.48	−0.77	0.78
Invariable – Boundary – Unstressed	0.06	0.48	−0.71	0.83
Invariable – No Boundary – Stressed	0.11	0.47	−0.62	0.86
Invariable – No Boundary – Unstressed	0.09	0.47	−0.64	0.84
Variable – Boundary – Stressed	−0.05	0.51	−0.86	0.77
Variable – Boundary – Unstressed	−0.07	0.51	−0.90	0.75
Variable – No Boundary – Stressed	−0.06	0.51	−0.88	0.75
Variable – No Boundary – Unstressed	0.27	0.51	−0.55	1.09

Table 9 (*Continued*)

Contrasts	Mean	Standard deviation	95% interval lower bound	95% interval upper bound
Invariable – Boundary – Stressed-Unstressed	−0.05	0.06	−0.17	0.06
Invariable – No Boundary – Stressed-Unstressed	0.02	0.04	−0.05	0.10
Variable – Boundary – Stressed-Unstressed	0.02	0.09	−0.16	0.19
Variable – No Boundary – Stressed-Unstressed	−0.33	0.06	−0.44	−0.22
Random Parameters				
Invariable – Style Standard Deviation	1.27	0.29	0.89	1.79
Variable – Style Standard Deviation	1.96	0.40	1.41	2.67

The contrasts in Table 9 show a clear difference between the *Stressed* and *Unstressed* estimates in the *No Boundary* condition for the *Variable* plurals, and model comparison shows that adding the *Prosodic Boundary* × *Next Stress* × *Variability* results in a strong improvement. However, the individual parameter estimates for this interaction show that all conditions largely overlap. This uncertainty in the parameter estimates is a result of the model containing multiple interactions that incorporate the *Boundary* variable. Specifically, parameters representing the same level of this variable in different interactions., e.g., *Variable – No Boundary – Consonant* and *Variable – No Boundary – Stressed*, are colinear and have negatively correlated posterior distributions. However, parameters representing the same level of the *Boundary* variable within the same interaction, e.g., *Variable – No Boundary – Stressed* and *Variable – No Boundary – Unstressed*, are not colinear with each other because they are estimated on different subsets of the data. In fact, if we subtract the posterior distributions of one of these parameters from the other, we get rid of their shared uncertainty due to collinearity with the parameters of other interaction terms, leaving us with uncorrelated estimates of the relevant contrasts, e.g., *Variable – No Boundary – Stressed-Unstressed*.

Table 10 Posterior distribution of the Null and Alternative models of the combined data. All estimates are in log-odds of an *-s* variant. Estimates for individual levels of the Style variable are omitted for the sake of brevity

Fixed Parameters	Null Model				Alternative Model			
	Mean	Standard deviation	95% interval lower bound	95% interval upper bound	Mean	Standard deviation	95% interval lower bound	95% interval upper bound
Invariable – Boundary – Sibilant	−0.06	0.26	−0.58	0.46	−0.08	0.42	−0.88	0.75
Invariable – Boundary – Vowel	0.18	0.26	−0.33	0.69	0.17	0.41	−0.62	0.98
Invariable – Boundary – Consonant	0.05	0.25	−0.44	0.55	0.04	0.41	−0.75	0.85
Invariable – No Boundary – Sibilant	0.30	0.26	−0.20	0.81	0.22	0.41	−0.60	1.00
Invariable – No Boundary – Vowel	−0.06	0.25	−0.56	0.44	−0.15	0.41	−0.96	0.64
Invariable – No Boundary – Consonant	0.21	0.25	−0.29	0.71	0.12	0.41	−0.70	0.90
Variable – Boundary – Sibilant	0.35	0.33	−0.32	0.97	0.36	0.43	−0.47	1.19
Variable – Boundary – Vowel	−0.17	0.31	−0.79	0.43	−0.15	0.41	−1.00	0.64
Variable – Boundary – Consonant	−0.22	0.31	−0.84	0.37	−0.21	0.41	−1.04	0.60
Variable – No Boundary – Sibilant	0.42	0.31	−0.19	1.02	0.36	0.42	−0.43	1.21
Variable – No Boundary – Vowel	−0.06	0.31	−0.69	0.52	−0.13	0.42	−0.93	0.71
Variable – No Boundary – Consonant	−0.17	0.31	−0.78	0.41	−0.24	0.42	−1.04	0.59
Boundary – Stressed					0.02	0.39	−0.75	0.82
Boundary – Unstressed					0.04	0.40	−0.72	0.85
No Boundary – Stressed					0.09	0.39	−0.68	0.84
No Boundary – Unstressed					0.19	0.39	−0.60	0.95

Table 10 (*Continued*)

Contrasts	Null Model				Alternative Model			
	Mean	Standard deviation	95% interval lower bound	95% interval upper bound	Mean	Standard deviation	95% interval lower bound	95% interval upper bound
Boundary – Stressed-Unstressed					−0.03	0.05	−0.13	0.07
No Boundary – Stressed-Unstressed					−0.10	0.03	−0.16	−0.04
Random Parameters								
Invariable – Style Standard Deviation	1.27	0.29	0.83	1.94	1.27	0.30	0.83	1.99
Variable – Style Standard Deviation	1.98	0.42	1.34	2.94	1.96	0.39	1.36	2.85

Appendix D

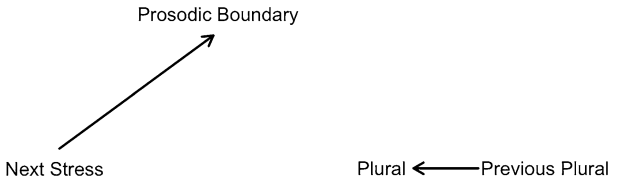

Fig. 9 Simplified causal diagram of the hypothesis that the presence of Prosodic Boundaries depends on whether the next syllable is stressed or not. Which Plural is produced depends on the Previous Plural that was used

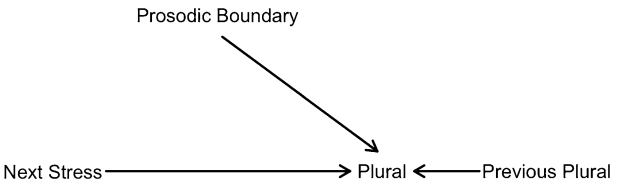

Fig. 10 Simplified causal diagram of the hypothesis that the type of Plural depends on the Previous Plural that was used, on the presence of Prosodic Boundaries and whether the next syllable is stressed or not

Acknowledgements The authors thank the editor and reviewers for their useful feedback, Ingo Plag for the discussions that helped shape this research, Vasudha Raju for contributing to the pilot study, Kyra Kashigin for helping with data collection and annotation, and Yung Han Khoe for a helpful suggestion.

Author contribution All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Tim Zee. The first draft of the manuscript was written by Tim Zee and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding information This research was funded by the Deutsche Forschungsgemeinschaft (Research Unit FOR2373 'Spoken Morphology', Project ER 547/1-1 'Dutch morphologically complex words: The role of morphology in speech production and comprehension', which we gratefully acknowledge.

Data availability Data, materials and code generated in this research can be accessed here: https://doi.org/10.34973/hqyq-7q43

Declarations

Ethics approval The studies involving human participants were reviewed and approved by Ethics Assessment Committee Humanities Radboud University. The participants provided their written informed consent to participate in this study.

Competing interests The authors have no relevant financial or non-financial interests to disclose.

References

Arfs, M. (2007). *Rood of groen? De interne woordvolgorde in tweeledige werkwoordelijke eindgroepen met een voltooid deelwoord en een hulpwerkwoord in bijzinnen in het hedendaags Nederlands*. Göteborg: Acta Universitatis Gothoburgensis. https://hdl.handle.net/2077/17039.

Arndt-Lappe, S., & Hoffmann, S. (2022). Comparing approaches to phonological and orthographic corpus formats. In O. Schützler & J. Schlüter (Eds.), *Data and methods in corpus linguistics: Comparative approaches* (pp. 46–72). Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108589314.

Azzabou-Kacem, S. (2018). *Stress shift in English rhythm rule environments: Effects of prosodic boundary strength and stress clash types*. Doctoral thesis, University of Edinburgh.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1996). The CELEX lexical database. https://doi.org/10.35111/gs6s-gm48.

Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language*, *37*, 94–117. https://doi.org/10.1006/jmla.1997.2509.

Baayen, R. H., Schreuder, R., De Jong, N., & Krott, A. (2002). Dutch inflection: The rules that prove the exception. In S. G. Nooteboom, F. Weerman, & F. N. K. Wijnen (Eds.), *Storage and computation in the language faculty* (pp. 61–92). Dordrecht: Kluwer. https://doi.org/10.1007/978-94-010-0355-1_3.

Bennett, R., & Elfner, E. (2019). The syntax–prosody interface. *Annual Review of Linguistics*, *5*, 151–171. https://doi.org/10.1146/annurev-linguistics-011718-012503.

CBS (2022). Beroepen van werkenden. https://www.cbs.nl/nl-nl/visualisaties/dashboard-arbeidsmarkt/werkenden/beroepen-van-werkenden.

Breiss, C., & Hayes, B. (2020). Phonological markedness effects in sentence formation. *Language*, *96*(2), 338–370. https://doi.org/10.1353/lan.2020.0023.

Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, *10*(5), 425–455. https://doi.org/10.1080/01690969508407111.

De Haas, W., & Trommelen, M. (1993). *Morfologisch handboek van het Nederlands: Een overzicht van de woordvorming*. Hague: The SDU Uitgeverij.

de Schutter, G., van den Berg, B., Goeman, T., & de Jong, T. (2005). *Morfologische Atlas van de Nederlandse Dialecten (vol. 1)*. Amsterdam: Meertens Instituut/KANTL/Amsterdam UP.

Dell, G. S., Nozari, N., & Oppenheim, G. M. (2014). Word production: Behavioral and computational considerations. In M. Goldrick, V. S. Ferreira, & M. Miozzo (Eds.), *The Oxford handbook of language production* (pp. 88–104). Oxford: OUP. https://doi.org/10.1093/oxfordhb/9780199735471.013.014.

Gussenhoven, C. (1983). Stress shift in Dutch as a rhetorical device. *Linguistics*, *21*(4), 603–619. https://doi.org/10.1515/ling.1983.21.4.603.

Haeseryn, W., Romijn, K., Geerts, G., de Rooij, J., & Van den Toorn, M. C. (1997). *Algemene Nederlandse Spraakkunst. Groningen/Deurne: Martinus Nijhoff uitgevers/Wolters Plantyn*.

Hoffmann, S., Arndt-Lappe, S., & Uhrig, P. (2022). Rhythm in World Englishes: Evidence from a quantitative analysis of co-occurrence patterns in a corpus of L1 and L2 varieties of English. In P. Rautionaho, H. Parviainen, M. Kaunisto, & A. Nurmi (Eds.), *Social and regional variation in world englishes (191-213)*, New York: Routledge. https://doi.org/10.4324/9781003227342.

Ingulfsen, T. (2004). *Influence of syntax on prosodic boundary prediction*. (Report No. UCAM-CL-TR-610), Uniiversity of Cambridge Computer Laboratory. https://doi.org/10.48456/tr-610.

Jespersen, O. (1909). *A modern English grammar on historical principles* (Vol. 1). London: George Allen & Unwin.

Kager, R. (1989). A Metrical Theory of Stress and Destressing in English and Dutch [Doctoral dissertation, Utrecht University]. https://hdl.handle.net/1874/8680.

Kentner, G. (2015). Stress clash hampers processing of noncanonical structures in reading. In R. Vogel & R. van de Vijver (Eds.), *Rhythm in cognition and grammar: A Germanic perspective* (pp. 111–135). Berlin: de Gruyter. https://doi.org/10.1515/9783110378092.

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods*, *42*(3), 643–650. https://doi.org/10.3758/BRM.42.3.643.

Kürschner, S. (2009). Morphological non-blocking in Dutch plural allomorphy: A contrastive approach. *STUF – Language Typology and Universals*, *62*(4), 285–306. https://doi.org/10.1524/stuf.2009.0022.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–38. https://doi.org/10.1017/s0140525x99001776.

Marsi, E., Reynaert, M., Van Den Bosch, A., Daelemans, W., & Hoste, V. (2003). Learning to predict pitch accents and prosodic boundaries in Dutch. In *41st annual meeting of the association for computational linguistics* (pp. 489–496). Somerset: Association for Computational Linguistics. https://hdl.handle.net/1854/LU-598073.

McElreath, R. (2020). *Statistical rethinking: A Bayesian course with examples in R and Stan*. New York: Chapman and Hall/CRC. https://doi.org/10.1201/9780429029608.

Oostdijk, N. (2000). The spoken Dutch corpus: Overview and first evaluation. In *LREC-2000 proceedings* (pp. 887–894).

Oostdijk, N., Reynaert, M., Hoste, V., & Schuurman, I. (2013). The construction of a 500-million-word reference corpus of contemporary written Dutch. In P. Spyns & J. Odijk (Eds.), *Essential speech and language technology for Dutch* (pp. 219–247). Berlin: Springer. https://doi.org/10.1007/978-3-642-30910-6_13.

Quené, H. (2006). Rhythmic factors in weak-syllable insertion: An Internet corpus study. In *Proc. Speech Prosody 2006*.

Schlüter, J. (2005). *Rhythmic grammar: The influence of rhythm on grammatical variation and change in English*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110219265.

Schlüter, J. (2015). Rhythmic influence on grammar: Scope and limitations. In R. Vogel & R. van de Vijver (Eds.), *Rhythm in cognition and grammar: A Germanic perspective* (pp. 179–206). Berlin: de Gruyter. https://doi.org/10.1515/9783110378092.

Schlüter, J., & Knappe, G. (2018). Synonym selection as a strategy of stress clash avoidance. In S. Hoffmann, A. Sand, S. Arndt-Lappe, & L. M. Dillmann (Eds.), *Corpora and lexis* (pp. 69–105). Leiden: Brill. https://doi.org/10.1163/9789004361133_005.

Stan Development Team (2021). RStan: The R interface to Stan (Version 2.21.3) [Computer software]. https://mc-stan.org/.

Theissen, S. (2003). Dubbel meervoud -(e)n / -s. In *Handelingen - Koninklijke Zuid-Nederlandse maatschappij voor taal- en letterkunde en geschiedenis 57* (pp. 89–116). https://doi.org/10.21825/kzm.v57i0.17305.

Thornton, A. M. (2019). Overabundance in morphology. In *Oxford research encyclopedia of linguistics*, Oxford: Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.554.

Van Dale, J. H., Den Boon, T., Hendrickx, R., & Van der Sijs, N. (2015). *Van Dale Groot woordenboek van de Nederlandse taal*. Utrecht/Antwerpen: Van Dale Lexicografie.

Wheeldon, L. R., & Konopka, A. (2018). Spoken word production: Representation, retrieval, and integration. In S. Rueschemeyer & M. G. Gaskell (Eds.), *The Oxford handbook of psycholinguistics* (2nd ed., pp. 335–371). Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198786825.013.15.

Zee, T., Ten Bosch, L., Plag, I., & Ernestus, M. (2021). Paradigmatic relations interact during the production of complex words: Evidence from variable plurals in Dutch. *Frontiers in Psychology*, 3629. https://doi.org/10.3389/fpsyg.2021.720017.