



# The duration of word-final *s* in English: A comparison of regular-plural and pluralia-tantum nouns

Marcel Schlechtweg<sup>1</sup>  · Greville G. Corbett<sup>2</sup>

Received: 10 May 2020 / Accepted: 9 May 2021 / Published online: 2 June 2021  
© The Author(s) 2021

## Abstract

The alveolar fricative occurs in word-final position in English in different grammatical functions. Nominal suffixes may indicate plurality (e.g. *cars*), genitive case (e.g. *car's*) or plurality and genitive case in cumulation (e.g. *cars'*). Further, there are the third person singular verbal suffix (e.g. *she fears*) and the cliticized forms of the third person singular forms of *have* and *be* (e.g. *she's been lucky*; *she's friendly*). There is also non-affixal *s* (e.g. *freeze* (noun)). Against the standard view that all these types are homophonous, several empirical studies have shown that at least some of the fricatives listed can actually be differentiated in their duration. The present article expands this line of research and considers a further case, which has not been included in previous analyses: pluralia-tantum nouns (e.g. *goggles*). We report on a carefully controlled reading study in which native speakers of British English produced pluralia-tantum and comparable regular-plural nouns (e.g. *toggles*). The duration of the word-final fricative was measured, and it was found that the two do not systematically differ in this acoustic parameter. The new data are interpreted in comparison to relevant previous studies, and against the background of the similarities of pluralia-tantum and regular-plural nouns.

**Keywords** Pluralia-tantum nouns · Regular-plural nouns · Word-final *s* · Acoustics · English · Number

## 1 Introduction

The correspondence between form and meaning or function is not always one to one. One meaning or function can be represented by several forms, and several meanings or functions can be expressed by a single form. The second type includes cases of homophony, polysemy, and syncretism. In the present contribution, we are concerned

---

✉ M. Schlechtweg  
[marcelschlechtweg@gmail.com](mailto:marcelschlechtweg@gmail.com)

<sup>1</sup> Carl von Ossietzky University Oldenburg, Oldenburg, Germany

<sup>2</sup> University of Surrey, Guildford, UK

with an example of affixal homophony and we raise the question whether the supposedly identical *s* suffix in English pluralia-tantum nouns (PTN) and regular-plural nouns (RPN) is indeed homophonous. Overall, our project connects to a body of research that has provided evidence in favor of the idea that phonologically identical forms differ in fine acoustic detail. In particular, the influence of various factors on the phonetic realization of words was tested. For instance, words of high frequency (e.g. *time*) seem to be articulated with a shorter duration than homophones of a lower frequency (i.e. *thyme*) (see, e.g., Gahl 2008). A further case in this line of research is the investigation of different types of word-final *s* in English. In the nominal domain, one finds the plural (e.g. *cars*), genitive (e.g. *car's*) and plural-genitive (*cars'*). Further, the English *s* appears in third person singular verbs (e.g. *she fears*) and as cliticized forms (e.g. *she's been lucky*; *she's friendly*). Finally, there is non-affixal *s* (e.g. *freeze* (noun)). Scholars detected acoustic variation in these supposedly homophonous kinds of *s* (see, e.g., Seyfarth et al. 2018); however, one category of word-final English *s* has so far been ignored, namely that of PTN (e.g. *goggles*). Given the linguistic differences between PTN and RPN (e.g. *toggles*) in English, it makes sense to examine how their final segment is uttered. This represents a logical extension of previous research. It has been surprising to see each new apparently homophonous form being shown to be different. PTNs are therefore particularly interesting, since they could be either the next surprise, or they might show that we have finally reached the end point, where the forms are in fact homophonous.

The paper is structured as follows. Section 2 introduces the topic of phonetic variation among supposedly identical forms in general, before specifically focusing on previous research on the acoustic realization of word-final *s* in English. The findings are further examined with respect to models of speech production and the interplay of morphology, phonology, and phonetics. In addition, Sect. 2 introduces the hypothesis that PTN and RPN are possibly a further case of non-homophonous affixation and presents several reasons in support of this idea. Our experiment is described in Sect. 3. We present a reading study in which the data of 40 native speakers of British English is analyzed. The speakers produced PTN (e.g. *goggles*) and RPN (e.g. *toggles*) and we measured the duration of the word-final fricative. To preview our results, no systematic difference between the two groups of items was found. We discuss our findings in connection to previous research and the linguistic properties of PTN and RPN in Sect. 4, and conclude in Sect. 5.

## 2 Theoretical background

### 2.1 Acoustic variation in phonologically identical forms: overview

Research on phonetic variation, specifically duration differences, between phonologically identical forms is growing, and different factors that might be responsible for this variation have been examined. Examples of such factors are frequency, syntactic category, morphosyntactic number and morphological status. Frequency has been shown to affect the articulation of items in that more frequent words are shorter in duration than others (see, e.g., Conwell 2018; Drager 2011; Gahl 2008; Lohmann

2018a, 2018b; Whalen 1991; but see also, for conflicting results, Cohn et al. 2005; Jurafsky et al. 2002). A well-known example is the *time/thyme* homophone pair, in which the more frequent item *time* is shorter than *thyme* (see Gahl 2008). Moreover, there is evidence that morpho-syntactic aspects can have an effect on the duration of words. While one study suggests that noun-verb conversions such as *answer/answers* differ acoustically (Sereno and Jongman 1995), another could not confirm this (Lohmann 2020). Schlechtweg and Heinrichs (2020) and Schlechtweg et al. (2020) show in two studies that German noun syncretisms, which have the same phonological form in their singular and plural (e.g. *Schatten* 'shadow(s)'), are acoustically differentiated in that the singular has a shorter duration than the plural. Finally, there is evidence suggesting that sequences of morphological stems and affixes are articulated differently than comparable sequences without morphological complexity (see, e.g., Kemps et al. 2005a, 2005b; Smith et al. 2012; Sugahara and Turk 2009). That is, for instance, plural stems (e.g. *car* in *cars*) seem to be shorter than the corresponding singular form (i.e. *car*) (see Kemps et al. 2005b). It is the factor morphological status that leads us to the aspect of interest in the current contribution, the *s* suffix in English.

## 2.2 Word-final *s* in English

Having looked at four potential reasons for phonetic variation at a general level, we now home in on findings concerning the duration of word-final *s* in English, since this is the dependent variable in the experiment reported in Sect. 3. Seyfarth et al. (2018) conducted a pseudo-conversational speech experiment, in which participants produced prepared and controlled dialogues. The authors concentrated on potential acoustic differences between English words containing an affixal *s* (e.g. third person singular or plural suffix as in *frees* and *laps*) and non-affixal *s* (e.g. *freeze*, *lapse*); they found that affixal *s* had a longer duration than non-affixal *s*. Walsh and Parker (1983) found the same effect in a reading study, in which each of three word pairs (*laps/lapse*, *wrecks/Rex*, *hearts/Hartz*) was integrated in each of three contexts. Schwarzlose and Bradlow (2001) report the same trend for their data gathered from four speakers. Song et al. (2013) considered spontaneous speech from a longitudinal study with three children and six mothers and found that affixal *s* was longer than non-affixal *s* in utterance-final position. Focusing on conversational speech in American English from a sample of the Buckeye corpus (Pitt et al. 2007), Plag et al. (2017) examined acoustic properties of English words with word-final *s*. On the one hand, they contrasted non-affixal to affixal and clitic *s*, and, on the other hand, they compared different kinds of affixal and clitic *s*, namely plural, genitive, plural-genitive, third person singular, cliticized *has* and cliticized *is*. The analysis revealed both differences between non-affixal and affixal/clitic *s* and between various types of affixal and clitic *s*, especially for voiceless *s*. That is, the duration of word-final *s* was shorter if it was used as a clitic in comparison to its use as a suffix. The latter, in turn, showed a shorter duration than non-affixal *s*. Plag et al.'s (2017) results were replicated by Tomaschek et al. (2021) relying on the complete Buckeye corpus, by Zimmermann (2016) for New Zealand English and by Schmitz et al. (2021) with pseudowords for Southern British English. Overall, the aforementioned studies revealed that affixal

and non-affixal word-final *s* in English differ in duration. The direction of the effect differs, however: while Walsh and Parker (1983), Schwarzlose and Bradlow (2001), Song et al. (2013) and Seyfarth et al. (2018) present evidence in favor of longer affixal *s*, Plag et al. (2017), Tomaschek et al. (2021), Zimmermann (2016) and Schmitz et al. (2021) point to the opposite.<sup>1</sup>

So far, we have primarily taken into account research that looked at the duration of affixal and non-affixal word-final *s* in English. Three other contributions are dedicated to possible duration differences of various types of affixal *s*. Hsieh et al. (1999) analyzed stories read by mothers to their children and demonstrated that third person singular *s* is shorter than the plural *s*; however, the authors also emphasize that plural forms occur more frequently in sentence-final position than the verbal suffix, which, in turn, contributes to the effect. Song et al. (2013) did not find a difference between the two. In a reading study, Plag et al. (2020) contrast the duration of the plural *s* (e.g. *colleagues*) with that of the plural-genitive *s* (e.g. *colleagues'*) and show that the duration of the former is shorter than that of the latter. They attribute the effect to the higher frequency of usage of the plural in comparison to the plural-genitive. This completes our overview of studies that examined the acoustics of word-final *s* in English. Our goal is to expand this line of research and investigate a so far unexplored type of word-final *s* in English, the *s* of PTN, and to compare its duration to that of RPN.

### 2.3 Implications for linguistic modelling

The detection of fine acoustic differences between phonologically identical words or word parts calls central assumptions of well-established linguistic models into question.<sup>2</sup> Broadly speaking, these models are of two types. First, in the light of psycholinguistic feed-forward models of speech production (see, e.g., Fromkin 1971; Harley 1984; Levelt 1989, 1995; Levelt et al. 1999; Roelofs 1997), one should not expect this kind of acoustic variation. The reason is that there is no direct connection between morphology and phonetics. Once we have left behind semantic, syntactic and morphological aspects here, and have arrived at the phonological level containing discrete symbolic representations, if two forms (say, the affixal and non-affixal *s*) are phonologically identical, we should no longer expect variation in the pronunciation of the two (provided that they are spoken in identical contexts). Second, well-known theories dealing with the interaction of morphology and phonology (e.g. Bermúdez-Otero 2018; Chomsky and Halle 1968; Kiparsky 1982) face similar problems. If the

<sup>1</sup>When considering these conflicting effects, we should take the following into consideration. First, as outlined in Tomaschek et al. (2021:128), the findings from Plag et al. (2017) and Seyfarth et al. (2018) seem to be more similar than one might assume at first glance. Seyfarth et al. (2018) analyzed more voiced than voiceless variants of word-final *s*. Concentrating on the voiced *s* only in Plag et al.'s (2017) data, one realizes that the affixal [z] is longer than the non-affixal one here as well. Second, some of the findings have to be treated with caution. As stated in Plag et al. (2017:185), Schwarzlose and Bradlow's (2001) study is described only in a short abstract and is difficult to discuss at all. We further agree with Plag et al. (2017:185) that Walsh and Parker's (1983) project is problematic in that they looked at a rather small data set (168 sound files) and do not provide inferential statistics. Also, more details on the experiment they conducted are needed in order to evaluate the findings.

<sup>2</sup>See also, e.g., Plag et al. (2017), Tomaschek et al. (2021), Vijver and Tomaschek (2021).

abstract and underlying outcome of lexical phonology is the same for affixal and non-affixal *s*, post-lexical phonology and phonetics do not have the means to create acoustic distinctions. A way out of the dilemma might be exemplar-based accounts, which are less rigid in the organization of speech production and in which the activation of semantic, syntactic or morphological information can have a direct influence on the articulation of items (see, e.g., Dell 1986; Pierrehumbert 2001, 2002).

## 2.4 PTN versus RPN: why the *s* might differ in duration

PTN are understood in the current contribution as nouns that are morphologically and syntactically, but often not semantically, plural; they lack a singular counterpart, and thus show a defective paradigm (see, e.g., Anderson 1992:108; Corbett 2019:54; Karlsson 2000:648–649; see also, e.g., Acquaviva 2008:12–16; Matthews 1997:284; Payne and Huddleston 2002: 340–345; Quirk et al. 1985:297–304; Wisniewski 2010:181). That is, the word *goggles* in *The goggles are new* carries the regular English plural suffix *s*, controls agreement of the verb, but does not necessarily refer to more than one item; one can interpret the noun in the sense of *a pair of goggles* or *more than one pair of goggles*.<sup>3</sup>

Although PTN and RPN are similar in that both control plural agreement, the two also differ substantially. In the present analysis, we rely on three aspects – the informative value of the *s* in the two noun types, the paradigmatic properties of PTN and RPN and the potential psycholinguistic characteristics of the two – to hypothesize that the duration of the word-final *s* is different between PTN and RPN. The first argument is that less informative, or less functionally relevant, information is reduced in speech production (see, e.g., Demuth 2011; Krasheninnikova 1979:75). An example of this is Engelhardt and Ferreira's (2014) study. The authors conducted an experiment in which subjects saw a panel with several sub-sections. For instance, on one panel participants saw a white heart, a yellow star, a blue triangle and a blue heart. One of these was marked with an arrow, in this case the blue triangle, and subjects were requested to describe this particular object (by saying, e.g., *the triangle* or *the blue triangle*). In the given example, the modifier *blue* was not necessary to identify the target object uniquely, since there was only one triangle. In another scenario, the same panel with the same objects was presented. This time, however, the arrow pointed at the blue heart and unique identification required the modifier *blue*, as there was also a white heart. Engelhardt and Ferreira (2014) showed that necessary modifiers (e.g. *blue* in the second scenario) were produced with a longer duration than the same modifiers if they were unnecessary (e.g. *blue* in the first scenario). Since PTN do not have a singular counterpart, the *s* does not play a crucial role in distinguishing the plural from the singular version of a noun. In contrast, in RPN, the suffix expresses the number distinction and differentiates the singular and plural form. This supports the idea that the *s* in RPN could be longer than in PTN.

Second, the paradigmatic properties of PTN and RPN might cause different *s* durations. The defective paradigm of PTN affects how language users integrate these

---

<sup>3</sup>An anonymous reviewer pointed out that PTN – but not RPN – occasionally occur with a singular determiner in spoken English (e.g. *a scissors*). This seems to derive from the semantic difference of the two, since only PTN but not RPN can refer to a single item.

nouns in a sentence. Using a sentence-completion task, Bock et al. (2001) provided subjects with the beginning of a sentence (e.g. *The advertisement for the scissors*) and asked them to repeat the phrase given and then finish the sentence. They were interested in the verb form produced and its agreement with either the head of the noun phrase (e.g. *advertisement*) or the noun in the complement (e.g. *scissors*). For the present purpose, the results from Experiment 1 on English are of particular interest. All of the given noun phrases were, overall, singular in meaning but they had different types of complements; this could include a singular noun (e.g. *The advertisement for the razor*), a RPN (e.g. *The advertisement for the razors*) or a PTN (e.g. *The advertisement for the scissors*). If a singular noun (e.g. *razor*) appeared, almost always a singular verb followed (proportion of plural verbs = 0.01). If a RPN (e.g. *razors*) occurred, a plural verb became more likely (proportion of plural verbs = 0.34). Finally, in the PTN condition, the proportion of plural verbs was 0.22. One might speculate that the fact that the PTN are more likely to refer to a single item than the RPN (which are always semantically plural) is responsible for the effect. However, in Experiment 2, Bock et al. (2001) discovered that PTN such as *suds*, which “tend to be conceived of as multiple” (Bock et al. 2001:101), showed a similar pattern to items like *scissors*, which are by default interpreted as single items (proportion of plural values: singular nouns (e.g. *bubble*): 0.01; RPN (e.g. *bubbles*): 0.21; PTN (e.g. *suds*): 0.12). Bock et al. (2001) suggest that the lack of a singular counterpart and/or the lack of a regular inflectional process in PTN, rather than semantic peculiarities, explain the results. This might possibly trigger an acoustic difference between PTN and RPN.

Stronger support for the idea that the paradigmatic properties of PTN and RPN might be reflected in the duration of the *s* comes from contributions showing that higher predictability has an impact on length. Generally speaking, more predictable elements – such as segments, syllables, words or syntactic structures – are often subject to reduction (see, e.g., Bell et al. 2003,2009; Gahl and Garnsey 2004; Jurafsky et al. 2001; Moore-Cantwell 2013; for an overview, see also Rose 2017:3–4). Reduction contributes to less effortful communication and high predictability ensures that communication remains successful in presence of reduced linguistic material (see, e.g., Frank and Jaeger 2008; Kurumada and Grimm 2017; Kurumada and Jaeger 2015; Norcliffe and Jaeger 2016; for a related issue, see Clopper and Turnbull 2018; Lindblom 1990). Focusing on morphological predictability, one can differentiate between syntagmatic and paradigmatic predictability (see, e.g., Cohen 2014; Rose 2017). While the former refers to the probability of occurrence of a form in a specific context, the latter represents the probability of occurrence of a form in comparison to the other members of the same word paradigm. For our purpose, the notion of paradigmatic predictability is relevant. Due to the defective paradigm of PTN and the absence of a singular counterpart, the paradigmatic predictability is higher for PTN than for RPN. As outlined and discussed in Cohen (2014), and in contrast to other variants of predictability, such as syntagmatic predictability, the effects of paradigmatic predictability are less homogenous and both acoustic reduction and enhancement were detected in previous studies (see also, e.g., Hay 2001, 2003; Hay and Baayen 2001; Kuperman et al. 2007; Schuppler et al. 2012). In any case, the two noun types under investigation in the present article clearly differ in predictability: the *s* in

PTN is more predictable due to the lack of a singular form and we hypothesize that this paradigmatic deviation between PTN and RPN leads to a distinct duration of the suffix.

The third argument for hypothesizing acoustic differences between the *s* of PTN and RPN is their possible psycholinguistic characteristics. Before considering these two groups, however, some general remarks are in order. While simplex words are retrieved – and processed and represented more generally – as a whole, two routes are in principle available for morphologically complex words, the whole-word and / or the (de)composition route (see, e.g., Baayen 1992; Baayen and Schreuder 1999; Butterworth 1983; Caramazza et al. 1985, 1988; Clahsen and Almazan 2001; Frauenfelder and Schreuder 1992; Giraudo and Grainger 2000; Manelis and Tharp 1977; Pinker 1991; Taft and Forster 1975; for an overview, see, e.g., Schlechtweg 2018). Cohen (2014:299) connects the route question to the acoustic realization of words and states that “forms which are more likely to be retrieved as whole words, rather than morphologically decomposed forms, would be pronounced more like whole words – namely, with phonetic reduction.” One of her examples to illustrate this idea is the finding that the English plural suffix *s* (as in *laps*) was found to be longer than the non-affixal equivalent (as in *lapse*) (see also Sect. 2.2). That is, the acoustic difference between the two types of *s* might reflect the use of different routes, the composition (*laps*) and the whole-word route (*lapse*), respectively. Plag et al. (2020) refer to a similar point, namely the idea that (higher) morphological complexity (as in *laps*) causes higher processing complexity and a slower – and therefore longer – production.

Now, hypothesizing that the *s* in PTN is shorter than the *s* in RPN derives from the idea that the former might be more prone to whole-word retrieval than RPN. Beyersmann et al. (2015:871) refer to the possibility that whole-word retrieval in production might be more likely for plural-dominant plurals while singular-dominant plurals might be more prone to be composed (see also, e.g., Biedermann et al. 2013). Transferring this idea to PTN and RPN, we hypothesize that PTN are more likely to be retrieved as a whole and that RPN are more likely to be composed of stem and suffix since the frequency of the former, which are plural-dominant, is by definition higher than the frequency of their (non-existent) stem. For RPN, the frequency of the whole word is lower than that of the stem. Note that all of our RPN in the experiment were singular-dominant and had a higher frequency in the singular than in the plural. This possible psycholinguistic difference could then lead to a difference in the duration of the *s* in that the *s* of PTN is shorter than the suffix of RPN because PTN are retrieved as wholes while RPN are composed (see also Cohen 2014).<sup>4</sup>

---

<sup>4</sup>Cohen (2014) refers to work from Hay and colleagues (Hay 2001, 2003; Hay and Baayen 2001), which further supports the idea that reduction is connected to the frequency ratio of a complex word and its stem. The alveolar plosive /t/ is less likely to be deleted in derivatives (e.g. *softly*) whose frequency is lower than the frequency of their stem (*soft*). In derivatives (e.g. *swiftly*) whose frequency is higher than the frequency of their stem (*swift*), /t/ deletion is more probable.

**Table 1** Test items used

PTN	RPN
<i>shears</i>	<i>beers</i>
<i>trousers</i>	<i>browsers</i>
<i>earnings</i>	<i>yearnings</i>
<i>pliers</i>	<i>fires</i>
<i>tweezers</i>	<i>freezers</i>
<i>goggles</i>	<i>toggles</i>
<i>tongs</i>	<i>gongs</i>
<i>jeans</i>	<i>screens</i>
<i>odds</i>	<i>pod</i>

### 3 Methodology

Using Praat (Boersma and Weenink 2019), we conducted an experiment in which subjects read the same English sentences containing either PTN (e.g. *goggles*) or RPN (e.g. *toggles*). The aim was to investigate whether the duration of the word-final fricative is distinct in the two types of words.

#### 3.1 Subjects

We tested 40 monolingual native speakers of British English (29 females, 11 males).<sup>5</sup> Their average age was 20.9 years (Standard deviation (SD): 2.3), and they were students at the University of Surrey (UK). None of these subjects reported a speech disorder and all of them had normal or corrected-to-normal vision. They received modest payment for participating.

#### 3.2 Materials

Nine test pairs, each including a PTN and a RPN, were used in the study (see Table 1). The PTN were taken from Payne and Huddleston (2002:341–343) and Quirk et al. (1985:301–303).

The test nouns were embedded in the test sentences presented in the Appendix; an example is given in (1).

- (1) a. *The goggles appear to be broken and they're useless.*
- b. *The toggles appear to be broken and they're useless.*

When creating the test sentences, we took the following potentially confounding variables into account, and controlled for them across the two conditions “PTN” and “RPN”. First, the conditions were matched with respect to their phonetic and phonological properties. All of the 18 test nouns contained the final fricative [z], which appears after vowels and consonants that are neither sibilants (after sibilants [iz] appears) nor voiceless non-sibilants (after voiceless non-sibilants [s] appears) (see, e.g.,

<sup>5</sup>We did not consider the data of bilinguals and a person with a speech problem in the analysis.



**Table 2** Frequencies of the test items used. Values per million words

PTN	Frequency	RPN	Frequency
<i>shears</i>	0.4	<i>beers</i>	4.6
<i>trousers</i>	5.6	<i>browsers</i>	9.4
<i>earnings</i>	15.4	<i>yearnings</i>	0.2
<i>pliers</i>	0.5	<i>fires</i>	9.9
<i>tweezers</i>	0.4	<i>freezers</i>	0.7
<i>goggles</i>	1.0	<i>toggles</i>	0.2
<i>tongs</i>	0.4	<i>gongs</i>	0.3
<i>jeans</i>	3.5	<i>screens</i>	11.3
<i>odds</i>	10.3	<i>Pods</i>	1.2
<b>Mean</b>	<b>4.2</b>	<b>Mean</b>	<b>4.2</b>
Median	1.0	Median	1.2
SD	5.4	SD	4.7

Quirk et al. 1985:304; and Ahmed et al. 2021 for complications). In the two words of each pair, all the following were identical: at least two segments before the final fricative, the rhyme of the ultimate syllable and the stress pattern. Apart from the target nouns, the two sentence versions of each pair were exactly the same on the segmental and suprasegmental level (intonation, accents, stress). Second, both sentences of a pair were identical with respect to their syntax and the position of the target word within the sentence. Third, all target nouns were inanimate. Fourth, the two target words as well as the entire sentences of each pair had the same number of syllables. Fifth, no frequency differences were detected between the two conditions. The individual values, the means, medians and standard deviations are given in Table 2. An independent t-test confirmed that the two conditions did not significantly differ ( $t = 0.02$ ,  $p = .988$ ). The frequency values are given in “occurrences per million words” and are based on the ukWaC corpus,<sup>6</sup> which contains about two billion words collected from UK-based websites.

All PTN had an entry only in their plural form (e.g. *goggles*) in the Oxford Advanced Learner’s Dictionary (Hornby 2005) and no PTN had a singular entry (e.g. *goggle*). In contrast, all RPN had only an entry in their unmarked singular form (e.g. *toggle*) but no entry in the plural (e.g. *toggles*) (see also, e.g., Quirk et al. 1985:304). Hence, while the PTN lack a singular form, the RPN have both a singular and a (regularly formed) plural. Crucially, all RPN were singular-dominant: their frequency in the ukWaC corpus was higher in the singular than in the plural (see Table 3).<sup>7</sup>

Sixth, the sequence “target noun + following word” was novel, in the sense that it had a frequency of 0 per million words for all of the 18 test items. That is, our items were comparable in terms of syntagmatic predictability and were equally likely to occur given the following word. This is crucial since the reduction of linguistic elements increases with raising syntagmatic predictability (see, e.g., Cohen 2014). We

<sup>6</sup><http://corpus.leeds.ac.uk/itweb/htdocs/Query.html#> (Hartley et al. 2011).

<sup>7</sup>A dependent t-test indicated marginal significance ( $t = 2.14$ ,  $p = .065$ ).

**Table 3** Frequencies of the RPN in the singular and plural. Values per million words

Target noun	Singular	Plural
<i>beer(s)</i>	20.4	4.6
<i>browser(s)</i>	35.8	9.4
<i>yearning(s)</i>	1.2	0.2
<i>fire(s)</i>	86.6	9.9
<i>freezer(s)</i>	3.9	0.7
<i>toggle(s)</i>	0.7	0.2
<i>gong(s)</i>	0.7	0.3
<i>screen(s)</i>	74.7	11.3
<i>pod(s)</i>	2.1	1.2
<b>Mean</b>	<b>25.1</b>	<b>4.2</b>
Median	3.9	1.2
SD	33.8	4.7

further controlled for the potential confounds speaker, order and way of presentation (see Sect. 3.3).

### 3.3 Procedure

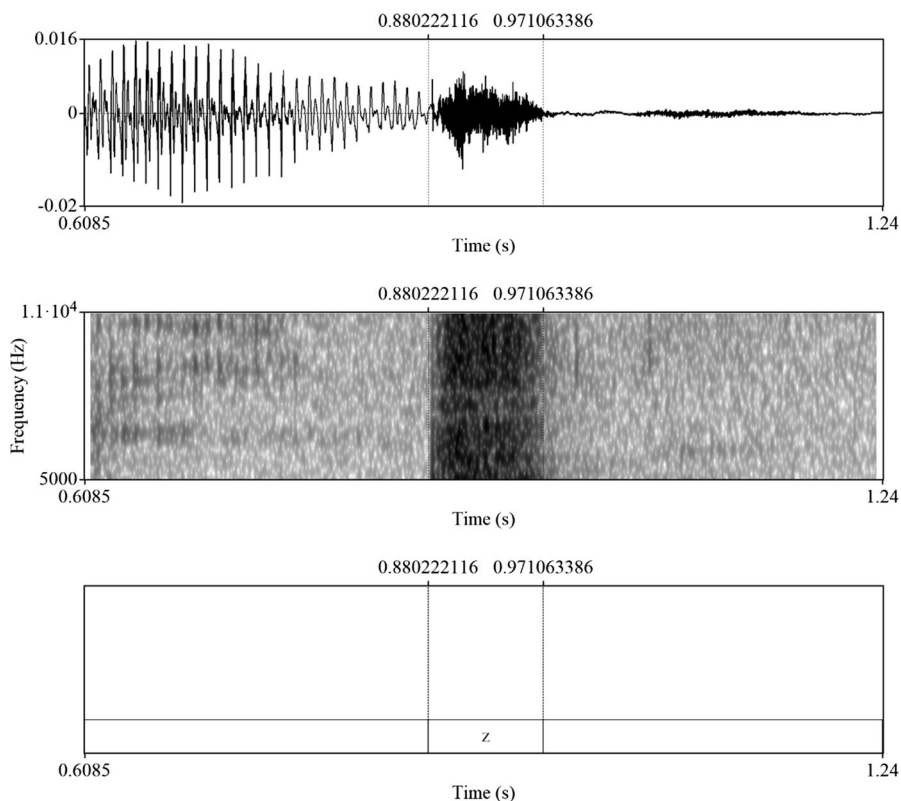
The study was conducted in a silent room at the University of Surrey. Subjects saw each sentence on a computer screen, which was placed approximately 60 centimeters (24 inches) in front of them. A large-diaphragm condenser microphone was placed centrally between the subject and the screen. Participants read all sentences silently first and then aloud while being recorded with Praat. On the computer screen, all sentences were left-aligned and appeared in the same font and font size, in a single line and in the middle of the screen.

Each person produced both sentences of a pair, that is, all of the 18 test sentences. The critical aspect of inter-subject variation was thus controlled for, since all subjects served as their own control. Moreover, 36 filler sentences were read out as well. Between each test sentence (e.g. the sentence containing the PTN *goggles*) and the respective sentence in the other condition (i.e. the sentence containing the RPN *toggles*), we inserted 26 other sentences in order first to avoid an influence of repetitions, and second to increase the distance between the two sentence versions of each pair. We counterbalanced the order of the two experimental conditions “PTN” and “RPN” not only within but also across participants. Further, the item order varied across subjects.

### 3.4 Data analysis

#### 3.4.1 Data preparation and segmentation

A total of 720 sound files (40 subjects  $\times$  18 sound files per subject) were collected, out of which 56 (8%) were discarded due to slips of the tongue or technical problems. Hence, 664 sound files remained, which were then phonetically segmented in Praat. In order to keep the segmentation consistency at the maximum level, a PTN (e.g.



**Fig. 1** Segmentation of [z] using waveform (top), spectrogram (middle) and Praat TextGrid (bottom)

*goggles*) and the respective RPN (e.g. *toggles*) spoken by the same person were segmented together. Relying on the information from the spectrogram and waveform, we identified the beginning and end of the word-final [z]. The spectrum settings were set to 5,000 to 11,000 Hertz (Hz) to improve the visibility of the target fricatives. On the basis of the acoustic properties of these fricatives and segmentation procedures specified in the literature (see, e.g., Ladefoged 2003; Ladefoged and Maddieson 1996; Machač and Skarnitzl 2009; Schlechtweg and Härtl 2020; Turk et al. 2006), we developed the following segmentation strategy. Our primary criterion in the segmentation process for detecting the beginning and end of the word-final fricative was the visibility of increased energy in the higher frequencies in the spectrogram. The second criterion was the visible fricative noise in the waveform. Priority was given to the primary criterion if the two did not coincide. The [z] was marked on a Praat TextGrid. Figure 1 shows an example file.

Note that we focused on the duration of the word-final [z] only and did not consider stem or word durations in our analysis. The reason for this decision was that different words were used in the two conditions PTN and RPN and, hence, it is simply not possible to entirely control the two for potential confounds. It is, for instance, impossible to compare the duration of the word-initial segment of *shears*

(voiceless postalveolar fricative) to that of *beers* (voiced bilabial plosive). Carefully controlling for potential confounds is always important and even more so when we investigate cases for which differences of no more than some milliseconds can be expected.

### 3.4.2 Statistical analysis and modelling

After the segmentation of the sound files, we used the program R (R Core Team 2021) for the following statistical analyses. First, we analyzed the descriptive statistics of the data. Since we carefully controlled the experiment prior to the study and were able to exclude the influence of many potential confounds, these simple values provide a first picture of what is going on in our data. Next, the data was examined with linear mixed effects models and the lme4 package (Bates et al. 2015) in R.<sup>8</sup> A reduced dataset (648 values), without statistical outliers, was used, and the log transformed (to the base 10) suffix durations represented the response variable. Statistical outliers were defined as values plus and minus 2.5 standard deviations from the mean (see, e.g., Loewen and Plonsky 2016:134). NounType (P (= PTN), R (= RPN)) was entered as a fixed effect. Further, although our experiment had been carefully controlled for many aspects, there are two factors that could not be controlled for in advance and these were entered as fixed-effect control variables. One was Log10SpeechRate\_z, the log transformed (to the base 10), centered and standardized speech rate (on log transformation, centering and standardizing, see, e.g., Winter 2020:86–98). The speech rate was defined as the quotient of the number of syllables of the entire sentence and the duration of this sentence in seconds. The other variable was GoogleBigrams\_z. It refers to the centered and standardized counts of the sequence “target noun + following word” in the Google Books Ngram Viewer (<https://books.google.com/ngrams>). Remember from Sect. 3.2 that the combination “target noun + following word” did not appear in the ukWaC corpus for all of the 18 test nouns and we therefore had evidence that the syntagmatic predictability across the two conditions of interest was controlled for. Nevertheless, to further reduce the probability that syntagmatic predictability might play a role, we included another measure, namely GoogleBigrams\_z (British English) in the model. For 16 of our 18 target words, the sequence “target noun + following word” had zero counts as well, similar to the ukWaC analysis. Apart from these two fixed-effect control variables, which could not be controlled for in advance, we added further fixed-effect control variables. Although our two groups of nouns, PTN and RPN, did not differ with respect to these variables, we aimed at investigating their role in the dataset. These were Log10Frequency\_z (log transformed, centered and standardized frequency of the target noun), VerbTense (present, past), DeterminerAgreement (Yes, No), PrecedingAdjective (Yes, No), FollowingWord (Verb, Adverb) and Log10LengthNoun\_z (log transformed, centered and standardized length of the target noun in number of syllables). VerbTense and DeterminerAgreement were included to see whether the absence and presence of overt morpho-syntactic agreement has an impact. If the verb

<sup>8</sup>Further packages involved in the data analysis are the readxl (Wickham and Bryan 2019) and the tidyverse package (Wickham et al. 2019).

has a present form (e.g. *appear*), it overtly agrees with the target plural noun with respect to the number feature since the singular noun would take a different verb form (*appears*). If the verb is in the past (e.g. *dropped*), no overt agreement is present. Considering the determiner and the plural noun, overt agreement is found with *these*, but not with *the*. Moreover, random intercepts for Subject and Item as well as random slopes for NounType by both Subject and Item were entered in the initial maximal model. Model fitting proceeded as follows (fit by maximum likelihood, see, e.g., Field et al. 2012:879). Even though a random effects structure with both intercepts and slopes can be generally justified and is in principle desirable (see, e.g., Winter 2020:235), complex random effects structures can be problematic (see, e.g., Barr et al. 2013; Cohen and Kang 2018; Matuschek et al. 2017; Martin Schweinberger p.c.). Since the initial maximal model was indeed inappropriate (“Singular fit” issue), the random slopes were removed from the model, which was then adequate.

We continued with the model containing the nine fixed effects and the random intercepts as specified above. The next objective was to eliminate non-significant factors following the strategy from Plag et al. (2017:194). That is, a factor had to pass three tests in order to remain in the model. First, when the factor was part of the model, the *t*-statistics needed to be greater than 2 or smaller than  $-2$ . Second, when the factor was part of the model, the Akaike Information Criterion (AIC) had to be smaller in comparison to the model without the factor (see also, e.g., Pinheiro and Bates 2000:10; Wu 2010:90). Third, including the factor should significantly improve the fit of the model, evidenced by a *p* value smaller than .05 in an ANOVA examining the model with and the model without the respective factor. The elimination order was determined on the basis of the factors’ values in the R output in the “Pr(>|t|)” column, starting with the highest one.<sup>9</sup> Apart from this stepwise, backward and manual model fitting, we additionally conducted a stepwise, backward and automatic model fitting using the step function of the lmerTest package (Kuznetsova et al. 2020; see also, e.g., Lohmann 2020:436).<sup>10</sup>

---

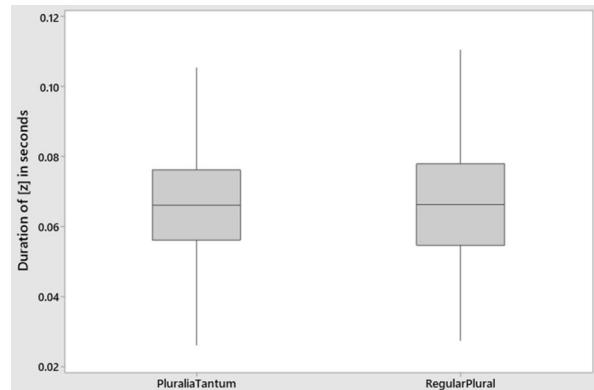
<sup>9</sup>The exclusion order was as follows: Log10LengthNoun\_z, NounType, GoogleBigrams\_z, Log10Frequency\_z, DeterminerAgreement, FollowingWord.

<sup>10</sup>To improve the transparency and documentation of the analyses conducted for the present experiment, we provide Supplementary Material for this article, which includes the following items. First, the whole dataset (664 values) is available (file called “Analysis”). Second, the R script for the analyses presented so far and in Sect. 3.5 is provided (file called “R script 1”). Third, we include a document with additional analyses that shows that the results introduced in Sect. 3.5 hold if specific modifications are implemented during the statistical procedure (file called “Further analyses”). The first analysis given in this file used the raw and untransformed *s* durations (648 values, statistical outliers excluded) as the response variable. The continuous variables (fixed effects) were centered and standardized (but not log transformed). The second analysis in this document relied on the log transformed (to the base 10) *s* durations, but no statistical outliers were excluded prior to the transformation (664 values). The continuous variables (fixed effects) were log transformed (to the base 10), centered, and standardized, with the exception of GoogleBigrams, which were centered and standardized only, due to the many zeroes in the dataset. Moreover, the descriptive values of the whole dataset (664 values) are given, and can be compared to those presented in Table 4. Fourth, we provide the R scripts for the two additional analyses just referred to (files “R script 2” and “R script 3”).

**Table 4** Reduced dataset without statistical outliers. Duration of *s* in seconds

Type of value	PTN	RPN	Total
N	327	321	648
Mean	0.067	0.067	0.067
Median	0.066	0.066	0.066
SD	0.017	0.017	0.017
Min	0.022	0.028	0.022
Max	0.114	0.118	0.118

**Fig. 2** Reduced dataset<sup>12</sup>



### 3.5 Results

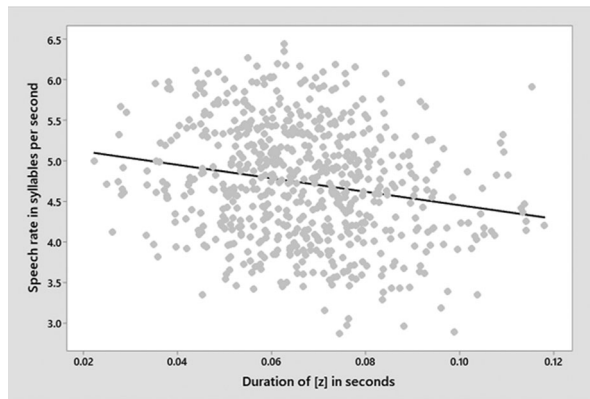
We present the descriptive statistics of the dataset without statistical outliers (648 data points) in Table 4 and Fig. 2.<sup>11</sup>

The two groups of interest, PTN and RPN, do not differ with respect to the duration of the suffix. Mixed effects modelling confirmed the result: the final model did not contain NounType but only, apart from the two random intercepts, Log10SpeechRate\_z, PrecedingAdjective and VerbTense. That is, (a) the duration of the *s* suffix increased if speech rate decreased; (b) the duration of the *s* suffix was longer if the verb in the sentence was in the past tense and there was therefore no overt agreement between the noun and the verb; and (c) the duration of the *s* suffix was longer if no adjective preceded the noun. We illustrate the respective descriptive statistics in Figs. 3, 4, 5, relying on, for ease of interpretation, the raw duration and

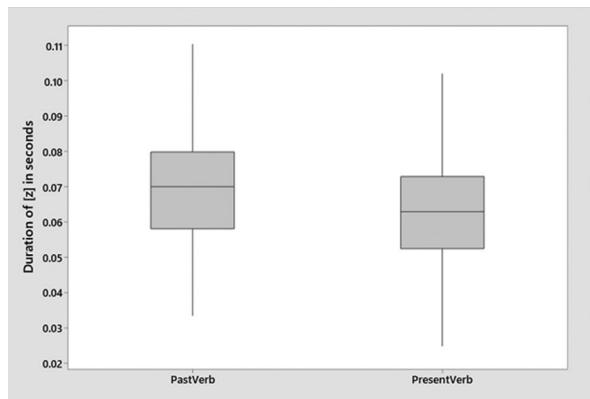
<sup>11</sup>Figs. 2 to 6 of this paper were created in Minitab (Minitab 2019).

<sup>12</sup>We would like to add a few words on the boxplot (see also, e.g., Larson-Hall 2010:245–246; Minitab 2019). The solid line in the middle of the box represents the median. The lower and the higher end of the box are the first (25th percentile) and the third quartile (75th percentile), the box includes the “central” 50 percent of the data (interquartile range). The vertical lines (whiskers) reach to the lowest and highest value in the dataset with the exception of values that are 1.5 times the length of the box away. These extreme values can be considered outliers (Larson-Hall 2010:245–246). However, these values are not visualized here since we relied on a different and common technique to identify outliers: outliers were defined as data points plus and minus 2.5 standard deviations from the mean in our analysis.

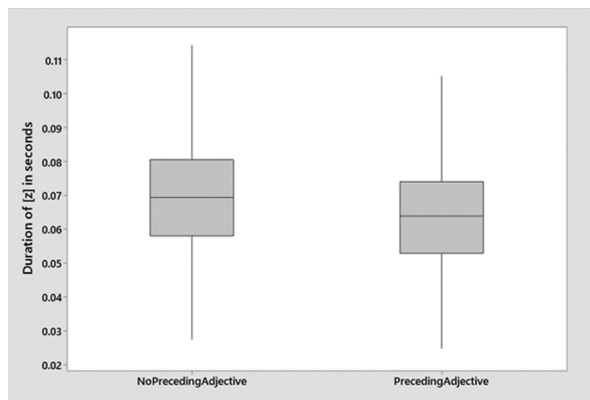
**Fig. 3** Effect of SpeechRate on the duration of *s*



**Fig. 4** Effect of VerbTense on the duration of *s*



**Fig. 5** Effect of PrecedingAdjective on the duration of *s*



speech rate values. The final model’s details are given in Tables 5 and 6. Note that the automatic procedure found the same model as the manual one.

We observe that our factor of interest, NounType, is not even part of the final model. To examine this factor in some more detail, we conducted a second analysis,

**Table 5** Random effects statistics of the mixed-effects model

	Variance	SD
Subject (Intercept)	0.0024160	0.04915
Item (Intercept)	0.0006934	0.02633
Residual	0.0075445	0.08686

**Table 6** Fixed effects statistics of the mixed-effects model

	Estimate	SE	df	t value	Pr(> t )
(Intercept)	-1.109210	0.015490	33.134832	-71.607	<2e-16***
Log10SpeechRate_z	-0.042009	0.006044	365.608046	-6.950	1.69e-11***
VerbTense (present)	-0.067431	0.014370	17.720595	-4.693	0.000189***
PrecedingAdjective (Yes)	-0.093674	0.015721	23.833677	-5.959	3.88e-06***

**Table 7** Random effects statistics of the mixed-effects model (NounType not eliminated and interactions added)

	Variance	SD
Subject (Intercept)	0.0024122	0.04911
Item (Intercept)	0.0006729	0.02594
Residual	0.0075423	0.08685

**Table 8** Fixed effects statistics of the mixed-effects model (NounType not eliminated and interactions added)

	Estimate	SE	df	t value	Pr(> t )
(Intercept)	-1.118073	0.020185	26.888001	-55.393	<2e-16***
NounType (R)	0.017679	0.026079	19.435153	0.678	0.505812
Log10SpeechRate_z	-0.039746	0.007702	457.262882	-5.160	3.68e-07***
VerbTense (present)	-0.059584	0.020033	17.522743	-2.974	0.008304**
PrecedingAdjective (Yes)	-0.086685	0.021721	23.124181	-3.991	0.000571***
NounType (R)* Log10SpeechRate_z	-0.004549	0.009426	522.684441	-0.483	0.629584
NounType (R)* PrecedingAdjective (Yes)	-0.013833	0.030097	21.766699	-0.460	0.650358
NounType (R)* VerbTense (present)	-0.015944	0.028371	17.610578	-0.562	0.581200

which was identical to the preceding one, the only two differences being that the fixed effect of interest, NounType, remained in the model until the end, irrespective of whether or not it fulfilled the three criteria from Plag et al. (2017:194), and that we added two-way interactions with NounType and all significant factors after all non-significant factors had been excluded.<sup>13</sup> The overall picture did not change, as can be seen in Tables 7 and 8.

<sup>13</sup>The exclusion order was the same as previously mentioned (with the exception of NounType).



### 3.6 Intermediate summary and power analysis

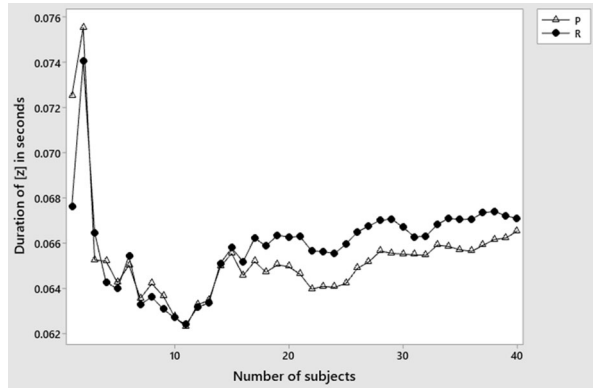
Our analyses revealed that the *s* suffix becomes longer if the speech rate decreases, which is unsurprising. Further, if no adjective precedes the noun, the *s* is longer in comparison to if an adjective appears in prenominal position. This might have prosodic reasons, but one should take into account that the sentences with and those without a modifying adjective contain different nouns and are overall different, which might also play a role here. Similarly, the finding that the *s* is longer if the verb in the sentence is in the past tense could have two origins. For one, the sentences with a present and those with a past tense verb differed and this might have caused the duration variation. Alternatively, speakers possibly compensate the absence of overt morpho-syntactic agreement in fine acoustic detail. While the present tense verbs functioned as a second plurality marker in the sentence, the only sign of plurality in most of the sentences containing a past verb is the *s*, and this might be enforced. It is beyond the scope of the present paper to investigate the robustness of this agreement effect in a well-controlled environment, but such a finding, if replicated and expanded, would be in line with the literature showing that rising syntagmatic predictability can trigger phonetic reduction (see Sect. 2.4; Rose 2017). That is, due to the presence of another plurality marker in sentences with overt number agreement, the *s* might potentially be shorter since it is more predictable in this particular context than in a context without overt agreement. Overall, and crucially, the observations made for speech rate, adjective presence and verb tense are not the core of the current analysis and do not interact with the central variable NounType. The latter, in turn, did not show the effect one might expect (see Sect. 2.4): the *s* of PTN did not differ in duration from the *s* of RPN.

A general and well-known danger of null effects is that the experiment could potentially be underpowered. In order to minimize this risk for our study, we conducted a power analysis illustrated in Durvasula and Liter (2020:197–198). In Fig. 6, we plot the cumulative mean suffix durations (y axis) of our 40 subjects. That is, for instance, “30” on the x axis refers to the mean of the average values of the first 30 subjects, “40”, in turn, refers to the mean of the average values of all 40 subjects. We see two lines, one with triangles for PTN (called P) and one with circles for RPN (called R). Using this graph, we see how the two groups of interest, PTN and RPN, develop as more subjects are added to the experiment and can then hypothesize what would happen if even more data was part of the study. We observe that the values of PTN and RPN develop in a comparable manner very early on and over the 40 subjects, and are even getting closer to each other at the high end of the x axis. This shapes our expectations, that is, on the basis of Fig. 6, we have no reason to expect that the picture would change if more data was added to the study since the two lines develop in a rather stable and consistent manner.

## 4 Summary and discussion

It is by now well known that the role of fine acoustic detail is greater than several established theories in morphology, phonology and psycholinguistics assume. Different variables have been proven to affect the duration of words, segments and other

**Fig. 6** Cumulative mean *s* durations in seconds



elements. The present contribution introduced another case, for which acoustic variation seemed plausible from an informative, a paradigmatic and a psycholinguistic angle. It turned out, however, that the PTN and RPN do not differ in terms of *s* duration. Our study does therefore not provide further support for theories of language (production) that allow a flexible interaction between the acoustic output and core linguistic domains such as morphology. We do not believe either, however, that feed-forward and other models denying a direct interplay of phonetics and higher-order areas tell the entire truth; too many other studies revealed that they are inadequate in several respects. Instead, we believe that our results support and emphasize the morpho-syntactic similarities of the items in focus. Despite the differences between PTN and RPN, the two are linguistically also similar.

RPN typically occur in their singular form in larger complex expressions like compounds. Gordon (1985), for instance, found that children prefer structures like *rat-eater*, rather than *rats-eater*. Even though this is not generally true for PTN, there is slight evidence that at least some PTN appear without the *s* in compounds (see, e.g., Gordon 1985). Two examples are *goggle-eyed* and *goggle repair kit* (Gregory Stump p.c.). This shows that PTN partially behave like RPN; however, to assess this aspect more precisely and to estimate how far these observations are compatible with our null effect, we would need more data and systematic analyses.

Earlier, we considered Beyersmann et al.'s (2015) suggestion that plural-dominant plurals might be more prone to whole-word retrieval than singular-dominant plurals. Our data do not support this idea: although PTN are more frequent and RPN less frequent than their stems, this variation did not lead to duration differences in our study, which, in turn, might speak for two things. One is a similar retrieval of the two. Several studies showed that the *s* in English nouns is articulated differently if its nature is non-affixal in comparison to if it has an affixal function (e.g. Plag et al. 2017; Seyfarth et al. 2018). Cohen (2014), among others, connects the acoustic distinction to different retrieval strategies, arguing that items with non-affixal *s* are retrieved as wholes, while the affixed nouns are composed upon usage. We concentrated on another comparison and found no acoustic deviation. If the retrieval strategies for PTN and RPN were indeed comparable, a logical follow-up question would be whether whole-word retrieval or concatenation of stem and suffix was more plausible for our

test items. Our acoustic measurements are hard to interpret in one or the other direction and we leave it to future research to examine the precise mechanisms more closely.<sup>14</sup> Alternatively, it might simply be the case that no difference in duration between PTN and RPN emerged despite differences in retrieval.

PTN and RPN share much more important characteristics, however: they take the plural suffix *s* and control plural agreement. This morphological and syntactic plurality of PTN is even preserved if abbreviations are used, as is evident in *knicks*, a compressed form of *knickers* (Wickens 1992:24). An experimental study further reflects the similarities between PTN and RPN. Nenonen and Niemi (2010) conducted a kind of lexical-decision experiment in Finnish, in which participants had to state whether the item they were exposed to is associated with one or many referents. While RPN were categorized as “Many” in 92 percent of the cases, PTN were so in about 36 percent. This is rather unsurprising as most PTN probably occur more frequently with the meaning “one item/referent”, although they are grammatically plural. Interestingly and despite this first result, however, if subjects stated that a PTN has many referents, they responded more quickly in comparison to if they stated that a PTN has one referent. It seems that the morphological plurality expressed by the PTN accelerated the “Many” judgments but impeded the “One” decisions. We consider this empirical finding, for Finnish, to reflect the morpho-syntactic similarity between PTN and RPN, which is a likely reason why the plural suffix does not differ in duration in our data. It seems that this strong similarity between PTN and RPN outweighs their differences discussed in Sect. 2.4 in speech production.

## 5 Conclusion

English affixal and non-affixal word-final *s* differ in duration, as has been shown in various studies. There is also evidence for a difference between the types of affixal *s*. It made good sense, therefore, to examine the final segment of PTN and RPN. Previous work gave plausible reasons why we might find an effect here. And yet, having carefully controlled for as many potentially confounding variables as possible, our experiment found no difference in duration for the final segment of PTN and RPN. We attribute the effect to the similarity of PTN and RPN, as shown by the fact that both control plural agreement and believe that these similarities outweigh the differences between PTN and RPN in speech production. Our result precedes the theory, in that while morphologists have been very interested in the degree of semantic predictability of PTN, they have given little attention to their morphological structure. The lack of a durational difference strongly suggests that English PTN

---

<sup>14</sup>The role of morphological structure has been extensively examined in psycholinguistics and a discussion of arguments for and against specific routes in retrieval (or processing and representation more generally) is beyond the scope of this paper. Although, as mentioned above, we cannot provide arguments for a specific route, we believe that whole-word retrieval of our test items seems less likely if PTN and RPN are really retrieved in the same way. Concatenating stem and suffix seems generally attractive for RPN and other regularly and productively formed inflectional forms, possibly except for complex items of higher frequency (see, e.g., Sereno and Jongman 1997; Stemberger and MacWhinney 1986). Our PTN and RPN had an average frequency of 4.2 occurrences per million words and are therefore not extremely frequent (see van Heuven et al. 2014:1180).

have a regular plural affix, and control agreement regularly. All of which makes the maintenance of their defective status more surprising.

## Appendix: Test sentences used

*His yearnings appear to worry his therapist more and more.*  
*His earnings appear to worry his therapist more and more.*

*The pods eventually dropped.*  
*The odds eventually dropped.*

*These new browsers unexpectedly caused a stir.*  
*These new trousers unexpectedly caused a stir.*

*These old screens obviously need replacing.*  
*These old jeans obviously need replacing.*

*These small freezers always cause problems.*  
*These small tweezers always cause problems.*

*The extra beers ended up in the shed.*  
*The extra shears ended up in the shed.*

*Our brass gongs amazingly sold for several pounds at auction.*  
*Our brass tongs amazingly sold for several pounds at auction.*

*The fires eased the job of getting rid of the rubbish.*  
*The pliers eased the job of getting rid of the rubbish.*

*The toggles appear to be broken and they're useless.*  
*The goggles appear to be broken and they're useless.*

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11525-021-09381-x>.

**Acknowledgements** Versions of this paper were given at the 19th International Morphology Meeting (Vienna, 2020) and at the 26th Architectures and Mechanisms for Language Processing Conference (Potsdam, 2020). We are grateful to those present for helpful comments, to Ingo Plag for insightful discussion and to two referees for valuable suggestions. We thank Penny Everson and Lisa Mack for their help in recruiting participants. Schlechtweg is the principal author.

**Funding Note** Open Access funding enabled and organized by Projekt DEAL.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Acquaviva, P. (2008). *Lexical plurals: a morphosemantic approach*. Oxford: Oxford University Press.
- Ahmed, S. K., Andersson, S., & Vaux, B. (2021). English phonology and morphology. In B. Aarts, A. McMahon, & L. Hinrichs (Eds.), *The handbook of English linguistics* (2nd edn., pp. 345–364). Hoboken: Wiley Blackwell.
- Anderson, S. R. (1992). *A-morphous morphology*. *Cambridge studies in linguistics: Vol. 62*. Cambridge: Cambridge University Press.
- Baayen, R. H. (1992). Quantitative aspects of morphological productivity. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology 1991* (pp. 109–149). Dordrecht: Kluwer Academic.
- Baayen, R. H., & Schreuder, R. (1999). War and peace: morphemes and full forms in a noninteractive activation parallel dual-route model. *Brain and Language*, *68*(1–2), 27–32.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. Version 1.1.26. *Journal of Statistical Software*, *67*(1), 1–48.
- Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., Gregory, M., & Gildea, D. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *Journal of the Acoustical Society of America*, *113*(2), 1001–1024.
- Bell, A., Brenier, J. M., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, *60*, 92–111.
- Bermúdez-Otero, R. (2018). Stratal phonology. In S. J. Hannahs & A. R. K. Bosch (Eds.), *The Routledge handbook of phonological theory* (pp. 100–134). New York, NY: Routledge.
- Beyersmann, E., Dutton, E. M., Amer, S., Schiller, N. O., & Biedermann, B. (2015). The production of singular- and plural-dominant nouns in Dutch. *Language, Cognition and Neuroscience*, *30*(7), 867–876.
- Biedermann, B., Beyersmann, E., Mason, C., & Nickels, L. (2013). Does plural dominance play a role in spoken picture naming? A comparison of unimpaired and impaired speakers. *Journal of Neurolinguistics*, *26*(6), 712–736.
- Bock, K., Eberhard, K. M., Cooper Cutting, J., Meyer, A. S., & Schriefers, H. (2001). Some attractions of verb agreement. *Cognitive Psychology*, *43*, 83–128.
- Boersma, P., & Weenink, D. (2019). Praat: doing phonetics by computer (Version 6.1.06). [Computer program]. Retrieved from <http://www.praat.org>.
- Butterworth, B. (1983). Lexical representation. In B. Butterworth (Ed.), *Language production: Vol. 2. Development, writing and other language processes* (pp. 257–294). London: Academic Press.
- Caramazza, A., Miceli, G., Silveri, M. C., & Laudanna, A. (1985). Reading mechanisms and the organization of the lexicon: evidence from acquired dyslexia. *Cognitive Neuropsychology*, *2*(1), 81–114.
- Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition*, *28*(3), 297–332.
- Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York, NY: Harper & Row.
- Clahsen, H., & Almazan, M. (2001). Compounding and inflection in language impairment: evidence from Williams syndrome (and SLI). *Lingua*, *111*(10), 729–757.
- Clopper, C. G., & Turnbull, R. (2018). Exploring variation in phonetic reduction: linguistic, social, and cognitive factors. In F. Cangemi, M. Clayards, O. Niebuhr, B. Schuppler, & M. Zellers (Eds.), *Phonology and phonetics: Vol. 25. Rethinking reduction: interdisciplinary perspectives on conditions, mechanisms, and domains for phonetic variation* (pp. 25–72). Berlin: De Gruyter Mouton.
- Cohen, C. (2014). Probabilistic reduction and probabilistic enhancement: contextual and paradigmatic effects on morpheme pronunciation. *Morphology*, *24*, 291–323.
- Cohen, C., & Kang, S. (2018). Flexible perceptual sensitivity to acoustic and distributional cues. *Mental Lexicon*, *13*(1), 38–73.
- Cohn, A. C., Brugman, J., Crawford, C., & Joseph, A. (2005). Lexical frequency effects and phonetic duration of English homophones: an acoustic study. *The Journal of the Acoustical Society of America*, *118*, 2036.
- Conwell, E. (2018). Token frequency effects in homophone production: an elicitation study. *Language and Speech*, *61*(3), 466–479.
- Corbett, G. G. (2019). Pluralia tantum nouns and the theory of features: a typology of nouns with non-canonical number properties. *Morphology*, *29*, 51–108.

- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93(3), 283–321.
- Demuth, K. (2011). The acquisition of phonology. In J. Goldsmith, J. Riggle, & A. C. L. Yu (Eds.), *The handbook of phonological theory* (2nd ed., pp. 571–595). Malden, MA: Wiley Blackwell.
- Drager, K. K. (2011). Sociophonetic variation and the lemma. *Journal of Phonetics*, 39, 694–707.
- Durvasula, K., & Liter, A. (2020). There is a simplicity bias when generalising from ambiguous data. *Phonology*, 37, 177–213.
- Engelhardt, P. E., & Ferreira, F. (2014). Do speakers articulate over-described modifiers differently from modifiers that are required by context? Implications for models of reference production. *Language, Cognition and Neuroscience*, 29(8), 975–985.
- Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. Los Angeles, CA: Sage.
- Frank, A. F., & Jaeger, T. F. (2008). Speaking rationally: uniform information density as an optimal strategy for language production. In *Proceedings of the annual meeting of the Cognitive Science Society* (Vol. 30, pp. 939–944).
- Frauenfelder, U. H., & Schreuder, R. (1992). Constraining psycholinguistic models of morphological processing and representation: the role of productivity. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology 1991* (pp. 165–183). Dordrecht: Kluwer Academic.
- Fromkin, V. A. (1973). The non-anomalous nature of anomalous utterances. In V. A. Fromkin (Ed.), *Janua linguarum: Vol. 77. Speech errors as linguistic evidence* (pp. 215–242). The Hague: Mouton. Reprinted from *Language* 47(1), 27–52 (1971).
- Gahl, S. (2008). Time and thyme are not homophones: the effect of lemma frequency on word durations in spontaneous speech. *Language*, 84(3), 474–496.
- Gahl, S., & Garnsey, S. M. (2004). Knowledge of grammar, knowledge of usage: syntactic probabilities affect pronunciation variation. *Language*, 80(4), 748–775.
- Giraud, H., & Grainger, J. (2000). Effects of prime word frequency and cumulative root frequency in masked morphological priming. *Language and Cognitive Processes*, 15(4/5), 421–444.
- Gordon, P. (1985). Level-ordering in lexical development. *Cognition*, 21, 73–93.
- Harley, T. A. (1984). A critique of top-down independent levels models of speech production: evidence from non-plan-internal speech errors. *Cognitive Science*, 8, 191–219.
- Hartley, T., Sharoff, S., Stephenson, P., Wilson, J., Babych, B., & Thomas, M. (2011). *IntelliText*. <http://corpus.leeds.ac.uk/itweb/htdocs/Query.html#>.
- Hay, J. (2001). Lexical frequency in morphology: is everything relative? *Linguistics*, 39(6), 1041–1070.
- Hay, J. (2003). *Causes and consequences of word structure*. New York, NY: Routledge.
- Hay, J., & Baayen, H. (2001). Parsing and productivity. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology 2001* (pp. 203–235). Dordrecht: Kluwer Academic.
- van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: a new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology*, 67(6), 1176–1190.
- Hornby, A. S. (2005). *Oxford advanced learner's dictionary* (2nd edn.). Oxford: Oxford University Press.
- Hsieh, L., Leonard, L. B., & Swanson, L. (1999). Some differences between English plural noun inflections and third singular verb inflections in the input: the contributions of frequency, sentence position, and duration. *Journal of Child Language*, 26(3), 531–543.
- Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. D. (2001). Probabilistic relations between words: evidence from reduction in lexical production. In J. L. Bybee & P. J. Hopper (Eds.), *Typological studies in language: Vol. 45. Frequency and the emergence of linguistic structure* (pp. 229–254). Amsterdam: Benjamins.
- Jurafsky, D., Bell, A., & Girand, C. (2002). The role of the lemma in form variation. In C. Gussenhoven & N. Warner (Eds.), *Phonology and phonetics: Vol. 4–1. Laboratory phonology: Vol. 7* (pp. 3–34). Berlin: Mouton de Gruyter.
- Karlssohn, F. (2000). Defectivity. In G. Booij, C. Lehmann, & J. Mugdan (Eds.), in collaboration with Wolfgang Kesselheim & Stavros Skopeteas. *Handbooks of linguistics and communication science: Vol. 17.1. Morphology: an international handbook on inflection and word-formation: Vol. 1* (pp. 647–654). Berlin: de Gruyter.
- Kemps, R. J. J. K., Ernestus, M., Schreuder, R., & Baayen, R. H. (2005a). Prosodic cues for morphological complexity: the case of Dutch plural nouns. *Memory & Cognition*, 33(3), 430–446.
- Kemps, R. J. J. K., Wurm, L. H., Ernestus, M., Schreuder, R., & Baayen, R. H. (2005b). Prosodic cues for morphological complexity in Dutch and English. *Language and Cognitive Processes*, 20(1–2), 43–73.

- Kiparsky, P. (1982). From cyclic phonology to lexical phonology (part 1). In H. van der Hulst & N. Smith (Eds.), *The structure of phonological representations* (pp. 131–176). Dordrecht: Foris.
- Krasheninnikova, E. A. (1979). Phonetic aspects of lingua-informatics. In H. Hollien & P. Hollien (Eds.), *Current issues in linguistic theory: Vol. 9. Current issues in the phonetic sciences. Proceedings of the IPS-77 congress*, Beach, Florida, 17–19 December 1977 (pp. 71–76). Amsterdam: Miami. John Benjamins.
- Kuperman, V., Pluymaekers, M., Ernestus, M., & Baayen, H. (2007). Morphological predictability and acoustic duration of interfixes in Dutch compounds. *The Journal of the Acoustical Society of America*, 121(4), 2261–2271.
- Kurumada, C., & Grimm, S. (2017). Communicative efficiency in language production and learning: optional plural marking. In *Proceedings of the 39<sup>th</sup> annual meeting of the cognitive science society*, CogSci 2017, London.
- Kurumada, C., & Jaeger, T. F. (2015). Communicative efficiency in language production: optional case-marking in Japanese. *Journal of Memory and Language*, 83, 152–178.
- Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B., & Jensen, S. P. 2020. Package lmerTest. Tests in linear mixed effects models. Version 3.1.3. <https://github.com/runehaubol/lmerTestR/issues>.
- Ladefoged, P. (2003). *Phonetic data analysis: an introduction to fieldwork and instrumental techniques*. Malden, MA: Blackwell.
- Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Malden, MA: Blackwell.
- Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. New York, NY: Routledge.
- Levelt, W. J. M. (1989). *Speaking: from intention to articulation*. Cambridge, MA: The MIT Press.
- Levelt, W. J. M. (1995). The ability to speak: from intensions to spoken words. *European Review*, 3(1), 13–23.
- Levelt, W. J. M., Roelofs, A., & Meyer, S. A. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1–75.
- Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H&H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modelling* (pp. 403–439). Dordrecht: Kluwer.
- Loewen, S., & Plonsky, L. (2016). *An A – Z of applied linguistics research methods*. London: Palgrave.
- Lohmann, A. (2018a). *Cut* (N) and *cut* (V) are not homophones: lemma frequency affects the duration of noun-verb conversion pairs. *Journal of Linguistics*, 54, 753–777.
- Lohmann, A. (2018b). *Time* and *thyme* are NOT homophones: a closer look at Gahl's work on the lemma frequency effect, including a reanalysis. *Language*, 94(2), e180–e190.
- Lohmann, A. (2020). No acoustic correlates of grammatical class: a critical re-examination of Sereno and Jongman (1995). *Phonetica*, 77, 429–440.
- Machač, P., & Skarnitzl, R. (2009). *Principles of phonetic segmentation*. Prague: Epocha Publishing House.
- Manelis, L., & Sharp, D. A. (1977). The processing of affixed words. *Memory & Cognition*, 5(6), 690–695.
- Matthews, P. H. (1997). *The concise Oxford dictionary of linguistics*. Oxford: Oxford University Press.
- Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing type I error and power in linear mixed models. *Journal of Memory and Language*, 94, 305–315.
- Minitab (2019). Minitab 19 [Computer program]. [www.minitab.com](http://www.minitab.com).
- Moore-Cantwell, C. (2013). Syntactic predictability influences duration. In *Proceedings of meetings on acoustics* (Vol. 19).
- Nononen, M., & Niemi, J. (2010). Mismatches between grammatical number and conceptual numerosity: a number-decision experiment on collective nouns, number neutralization, pluralia tantum, and idiomatic plurals. *Folia Linguistica*, 44(1), 103–125.
- Norcliffe, E., & Jaeger, T. F. (2016). Predicted head-marking variability in Yucatan Maya relative clause production. *Language and Cognition*, 8, 167–205.
- Payne, J., & Huddleston, R. (2002). Nouns and noun phrases. In R. Huddleston & G. K. Pullum (Eds.), *The Cambridge grammar of the English language* (pp. 323–523) Cambridge: Cambridge University Press.
- Pierrehumbert, J. B. (2001). Exemplar dynamics: word frequency, lenition and contrast. In J. L. Bybee & P. J. Hopper (Eds.), *Typological studies in language: Vol. 45. Frequency and the emergence of linguistic structure* (pp. 137–157). Amsterdam: John Benjamins.
- Pierrehumbert, J. B. (2002). Word-specific phonetics. In C. Gussenhoven & N. Warner (Eds.), *Phonology and phonetics: Vol. 4–1. Laboratory phonology: Vol. 7* (pp. 101–140). Berlin: Mouton de Gruyter.
- Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York, NY: Springer.

- Pinker, S. (1991). Rules of language. *Science*, 253(5019), 530–535.
- Pitt, M. A., Dilley, L., Johnson, K., Kiesling, S., Raymond, W., Hume, E., & Fosler-Lussier, E. (2007). *Buckeye corpus of conversational speech (2nd release)*. Columbus: Department of Psychology, Ohio State University.
- Plag, I., Homann, J., & Kunter, G. (2017). Homophony and morphology: the acoustics of word-final S in English. *Journal of Linguistics*, 53, 181–216.
- Plag, I., Hedia, S. B., Lohmann, A., & Zimmermann, J. (2020). An <s> is an <s'>, or is it? Plural and genitive-plural are not homophonous. In L. Körtvélyessy & P. Stekauer (Eds.), *Complex words: advances in morphology* (pp. 260–292). Cambridge: Cambridge University Press.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. London: Longman.
- R Core Team (2021). *R: a language and environment for statistical computing. R version 4.0.4*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org>.
- Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, 64, 249–284.
- Rose, D. E. (2017). *Predicting plurality: an examination of the effects of morphological predictability on the learning and realization of bound morphemes*. (Doctoral dissertation).
- Schlechtweg, M. (2018). *Memorization and the compound-phrase distinction: an investigation of complex constructions in German, French and English*. *Studia grammatica: Vol. 82*. Berlin: de Gruyter.
- Schlechtweg, M., & Härtl, H. (2020). Do we pronounce quotation? An analysis of name-informing and non-name-informing contexts. *Language and Speech*, 63(4), 769–798.
- Schlechtweg, M., & Heinrichs, M. (2020). *The acoustics of number: duration differences in singular-plural syncretism*. Unpublished manuscript.
- Schlechtweg, M., Heinrichs, M., & Linnenkohl, M. (2020). Differences in acoustic detail: the realization of syncretic nouns in German. In M. Schlechtweg (Ed.), *Trends in linguistics. Studies and monographs: Vol. 345. The learnability of complex constructions: a cross-linguistic perspective* (pp. 39–62). Berlin: De Gruyter Mouton.
- Schmitz, D., Baer-Henney, D., & Plag, I. (2021). The duration of word-final /s/ differs across morphological categories in English: evidence from pseudowords. Submitted.
- Schuppler, B., van Dommelen, W. A., Koreman, J., & Mirjam, E. (2012). How linguistic and probabilistic properties of a word affect the realization of its final /t/: studies at the phonemic and sub-phonemic level. *Journal of Phonetics*, 40, 595–607.
- Schwarzlose, R., & Bradlow, A. R. (2001). What happens to segment durations at the end of a word? *The Journal of the Acoustical Society of America*, 109, 2292.
- Sereno, J. A., & Jongman, A. (1995). Acoustic correlates of grammatical class. *Language and Speech*, 38(1), 57–76.
- Sereno, J. A., & Jongman, A. (1997). Processing of English inflectional morphology. *Memory & Cognition*, 25(4), 425–437.
- Seyfarth, S., Garellek, M., Gillingham, G., Ackerman, F., & Malouf, R. (2018). Acoustic differences in morphologically-distinct homophones. *Language, Cognition and Neuroscience*, 33(1), 32–49.
- Smith, R., Baker, R., & Hawkins, S. (2012). Phonetic detail that distinguishes prefixed from pseudo-prefixed words. *Journal of Phonetics*, 40, 689–705.
- Song, J. Y., Demuth, K., Evans, K., & Shattuck-Hufnagel, S. (2013). Durational cues to fricative codas in 2-year-olds' American English: voicing and morphemic factors. *The Journal of the Acoustical Society of America*, 133, 2931–2946.
- Stemberger, J. P., & MacWhinney, B. (1986). Frequency and the lexical storage of regularly inflected forms. *Memory & Cognition*, 14(1), 17–26.
- Sugahara, M., & Turk, A. (2009). Durational correlates of English sublexical constituent structure. *Phonology*, 26(3), 477–524.
- Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 14(6), 638–647.
- Tomaschek, F., Plag, I., Ernestus, M., & Baayen, R. H. (2021). Phonetic effects of morphology and context: modeling the duration of word-final S in English with naïve discriminative learning. *Journal of Linguistics*, 57, 123–161.
- Turk, A., Satsuki Nakai, S., & Sugahara, M. (2006). Acoustic segment durations in prosodic research: a practical guide. In S. Sudhoff, D. Lenertová, R. Meyer, S. Pappert, P. Augurzky, I. Mleinek, N. Richter, & J. Schließer (Eds.), *Language, context, and cognition: Vol. 3. Methods in empirical prosody research* (pp. 1–27). Berlin, Germany: de Gruyter.



- van de Vijver, R., & Tomaschek, F. (2021). Special issue: phonological and phonetic variation in spoken morphology. *Morphology*, *31*, 83–86.
- Walsh, T., & Parker, F. (1983). The duration of morphemic and non-morphemic /s/ in English. *Journal of Phonetics*, *11*(2), 201–206.
- Whalen, D. H. (1991). Infrequent words are longer in duration than frequent words. *The Journal of the Acoustical Society of America*, *90*, 2311.
- Wickens, M. A. (1992). *Grammatical number in English nouns. Current issues in linguistic theory: Vol. 76*. Amsterdam: John Benjamins.
- Wickham, H., & Bryan, J. (2019). Readxl: read excel files. R package. Version 1.3.1. <https://CRAN.R-project.org/package=readxl>.
- Wickham, H., et al. (2019). Welcome to the tidyverse. Version 1.3.0. *Journal of Open Source Software*, *4*(43), 1686.
- Winter, B. (2020). *Statistics for linguists: an introduction using R*. New York, NY: Routledge.
- Wisniewski, E. J. (2010). On using count nouns, mass nouns, and pluralia tantum: what counts? In F. J. Pelletier (Ed.), *New directions in cognitive science. Kinds, things, and stuff: mass terms and generics* (pp. 166–190). Oxford: Oxford University Press.
- Wu, L. (2010). *Mixed effects models for complex data. Monographs on statistics and applied probability: Vol. 113*. Boca Raton: CRC Press.
- Zimmermann, J. (2016). Morphological status and acoustic realization: findings from New Zealand English. In C. Carignan & M. D. Tyler (Eds.), *Proceedings of the sixteenth Australasian international conference on speech science and technology (SST-2016)*, Parramatta, Australia, 6–9 December 2016 (pp. 201–204). Canberra: ASSTA.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.