Research Article

The duration of word-final /s/ differs across morphological categories in English: Evidence from pseudowords

Dominic Schmitz[a]*, Dinah Baer-Henney[b], Ingo Plag[a]

[a] English Language and Linguistics, Heinrich Heine University, Düsseldorf, Germany

[b] Linguistics and Information Science, Heinrich Heine University, Düsseldorf, Germany

Short Title: The duration of word-final /s/ in English: Evidence from pseudowords

Corresponding Author:

Dominic Schmitz

English Language and Linguistics

Heinrich Heine University

Universitätsstraße 1

40225 Düsseldorf, Germany

Tel: +49 211 81-15126

E-mail: Dominic.Schmitz@uni-duesseldorf.de

Number of Tables:

13

Number of Figures:

6

Word count:

10,458

Keywords:

morphology, speech production, subphonemic difference, pseudoword paradigm

Abstract

Previous research suggests that different types of word-final /s/ and /z/ (e.g. non-morphemic or plural) in English show realizational differences in duration. However, there is no agreement on the nature of these differences. That is, experimental studies provide evidence for durational differences of the opposite direction as results by corpus studies.

The present study focuses on four types of word-final /s/ in English, i.e. non-morphemic, plural, and *is*- and *has*-clitic /s/. Adopting a pseudoword-paradigm, a production study with native speakers of Southern British English was carried out. The results show significant durational differences between the types of /s/ under investigation. That is, non-morphemic /s/ is longer than plural /s/, which in turn is longer than clitic /s/, while there is no durational difference between the two clitics. This is fully in line with previous corpus studies.

Thus, the morphological category of a word-final /s/ appears to be a robust predictor for its phonetic realization. Hence, morphological information may influence speech production in such a way that systematic subphonemic differences arise. This calls for revisions of current models of speech production in which morphology does not play a role in later stages of production.

1. Introduction

Recent research on the acoustic properties of seemingly homophonous elements has shown unexpected effects on their realization by their morphological structure. For words, studies have found evidence for seemingly homophonous lexemes to actually differ in phonetic details such as vowel quality or length (e.g. Gahl, 2008; Drager, 2011). For stems, Kemps et al. (2005a, b) found that free and bound variants of a stem differ acoustically, and that listeners make use of such phonetic cues in speech perception. For prefixes, Ben Hedia & Plag (2017) and Ben Hedia (2019) showed that the more segmentable a prefix the longer the duration of its nasal for *un-*, *in-* (negative) and *in-* (locative).

On the level of individual segments, several studies have shown that the phonetic realization of word-final /s/ and /z/ in English (henceforth S) depends on its morphological category. In corpus studies, Zimmermann (2016), Plag et al. (2017), and Tomaschek et al. (2019) found non-morphemic word-final S to have longer durations than morphemic word-final S, with suffixes showing longer durations than clitics. Experimental studies (e.g. Walsh & Parker, 1983; Li et al., 1999; Seyfarth et al., 2017; Plag et al., 2019) also found seemingly identical word-final S to be realized differently depending on its morphological category. However, their results are not as clear as those by the previously mentioned corpus studies. One major drawback of all previously conducted studies are the potentially confounding effects of the lexical and contextual properties of the items under investigation, e.g. potential storage effects (e.g. Caselli et al., 2016).

Most importantly, as traditional models of speech production assume that phonetic processing does not have information on morphological makeup (e.g. Levelt & Wheeldon, 1994; Levelt et al., 1999), such findings pose a serious challenge, calling for an explanation on how morphological information would come to influence articulation.

The present study also addresses realizational differences on the level of individual segments based on different types of word-final S in English. We investigate whether different types of word-final S, i.e. non-morphemic, plural, and *is-* and *has-*clitic S, show differing phonetic realizations in terms of duration. This, for the first time, will be done within a pseudoword paradigm in order to provide further insight into subphonemic realizational differences beyond lexical and contextual properties. We suggest that if systematic differences can also be found within pseudoword paradigms, one can assume realizational differences between seemingly identical segments to be of a robust nature rather than a by-product of confounding lexical factors. This calls for a revision of models on the relationship between morphology, phonology and phonetic realization.

The paper is structured as follows. In the next section we will take a closer look at the interplay of morphological structure and the phonetic signal. Section 3 will present our methodology. The analysis and results of our study are presented in Section 4 and 5, followed by a discussion and conclusion in Section 6.

2. Morphology and phonetic realization

In English, a number of morphological categories can take the phonological form of /s/, i.e. plural, genitive, genitive plural, 3rd person singular, as well as the clitics of *is*, *has*, and *us*. As such, there is nothing in the segmental representation of the morphological categories that accounts for systematic realizational differences on the phonetic level between different S morphemes, or between morphemic and non-morphemic S. One possible source of such phonetic differences could lie in the prosodic structure, however. In the framework of Prosodic Morphology, there is a complex mapping of morphological structure onto prosodic structure (e.g. Nespor & Vogel 2007), since prosodic boundaries may correlate with particular phonetic properties, segments at such boundaries may show systematic differences in phonetic implementation (see, for example, Keating 2006). Phonetic differences between two phonologically homophonous affixes could therefore result from a difference in the prosodic structure that goes with the two affixes.

All types of S, morphemic and non-morphemic, are treated in a similar way in standard feed-forward formal theories of morphology-phonology interaction (e.g. Chomsky & Halle, 1968; Kiparsky, 1982). In the case of morphological word-final S, a process called 'bracket erasure' is said to remove all morphological information from a pertinent word form once retrieved from the lexicon during the stage of 'lexical phonology' and leaves speech production without an insight into the morphological makeup at the stage of 'post-lexical phonology'. Once retrieved, there is no informational difference between word-final morphemic and non-morphemic types of S. Thus, there is nothing in such a system that could account for realizational differences, e.g. different durations, between phonologically identical suffixes, clitics, and non-morphemic segments.

Such a distinction of lexical and post-lexical processing is also an integral part of established theories in psycholinguistics. According to models of speech production such as the one proposed by Levelt et al. (1999), morphemic S would not differ in their realization from corresponding non-morphemic realizations of S. In such models, meanings are stored in the mental lexicon with their forms being represented phonologically. The module called 'articulator' uses these phonological forms for speech production, hence, has no information on the lexical origin of particular segments. Such a strict modular feed-forward model cannot explain durational differences between different types of word-final S.

Yet, there is more evidence that suggests that models as those by Kiparsky (1982) and Levelt et al. (1999) may be insufficient. For homophonous lexemes, Gahl (2008) and Lohmann (2018) investigated acoustic realizations of seemingly homophonous word pairs such as *time* and *thyme*, and found the more frequent member of each pair to be of shorter duration. This indicates a number of possible consequences. First, a separate storage entry for each member of a word pair appears to be evident. Second, separate entries come with individually stored frequencies, thus, influencing the realization of the pertinent lemma. Third, this may be evidence in favour of exemplar models (e.g. Goldinger, 1998; Bybee, 2001; Pierrehumbert, 2001, 2002; Gahl & Yu, 2006), assuming members of homophone pairs have individual lexical entries with accompanying features and information.

Further evidence for differing acoustic realizations of supposedly homophonous lexemes was found by Drager (2011). Drager compared realizations of *like* as adverb, verb, discourse particle, and as part of the quotative *be like*. Differences found surface in several phonetic parameters. Similar effects were found for function words such as *four* and *for* and different versions of words such as *to*, which were investigated by Lavoie

(2002) and Jurafsky et al. (2002). Such fine realizational differences indicate that at the phonetic level two or more phonologically homophonous lemmas may differ in their realization.

Similarly, evidence for seemingly homophonous elements below the word level having different phonetic realizations has also been found. Kemps et al. (2005a, b) found that in Dutch and English segmentally identical free and bound variants of a base (e.g. *help* without a suffix versus *help* in *helper*) differ acoustically, and that listeners make use of such phonetic cues in speech perception. Sugahara & Turk (2004, 2009) found phonetic differences between the final segments of a monomorphemic stem as compared to the final segments of the same stem if followed by a suffix, e.g. in *mist rain* versus *missed rain* (Sugahara & Turk, 2004). The stem had slightly longer rhymes if followed by certain suffixes. Seyfarth et al. (2017) found that for words ending in fricatives the durations of a word's morphological relatives influence the realization of that word. In their study, stems of morphologically complex words showed longer durations than similar strings of segments in homophonous simple words (e.g. *free* in *frees* vs. *freeze*). They concluded that the durational targets of the morphologically complex word's relatives influence the word's duration to such an extent that a durational difference between the pertinent complex word and its homophonous simple counterpart arise.

For prefixes, Smith et al. (2012) found systematic realizational differences for *dis-* and *mis-* between prefixed and so-called pseudo-prefixed words (e.g. *discolour* vs. *discover*). Prefixed words showed longer durations and longer voice onset times, among other things. Ben Hedia & Plag (2017) and Ben Hedia (2019) showed that the more segmentable a prefix the longer the duration of its nasal for *un-*, *in-* (negative) and *in-* (locative) prefixes.

On the articulatory level, Cho (2001) found evidence for the variability of intergestural timing between identical strings in heteromorphemic and tautomorphemic contexts. In their electropalatographic study, Cho showed that the timing of the gestures for [ti] and [ni] in Korean shows more variation when the sequence is heteromorphemic, thus indicating that morphological structure is reflected in articulatory gestures, which in turn may lead to correlates in the acoustic signal.

Thus, it seems that there is vast evidence for seemingly homophonous elements, i.e. lexemes, bases and affixes, to differ on the level of speech production. Differences on the level of segments have been reported as well. Previous corpus studies on word-final S in English found realizational differences between non-morphemic, suffix and clitic variants, while previous experimental studies differ in their findings. Zimmermann (2016) on New Zealand English (data from QuakeBox corpus; Walsh et al., 2013), and Plag et al. (2017) as well as Tomaschek et al. (2019) on North American English (data from Buckeye Corpus of Conversational Speech; Pitt et al., 2007) find that non-morphemic S showed longer durations than suffix and clitic S. In turn, suffix S was also of longer duration than clitic S. While these results draw a clear picture of S duration across morphological categories (including the non-morphemic S), they are subject to unbalanced data sets due to the nature of corpora. That is, corpus data may contain a huge number of confounding and moderator variables that experimental data can be controlled for (Gries, 2015).

Turning to previous experimental studies, we find less clear results. Walsh & Parker (1983) carried out a production experiment with three homophonous word pairs (e.g. *Rex* and *wrecks*). They measured the duration of the word-final S in both the monomorphemic and the complex word of each pair in three different conditions which where each produced by eight to ten participants. Condition I consisted of an unambiguous context; condition II consisted of a semantically neutral context; Condition

III consisted of a semantically anomalous context (Walsh & Parker, 1983: 201-202). While in two of these conditions there is a small difference of 9 ms in the means of the different types of S, there is none in the third condition. Still, they conclude that ‘speakers of English systematically lengthen morphemic /s/’ (Walsh & Parker, 1983: 204). However, their study shows several flaws. The analysed data set only consisted of about 110 observations of a mixture of common and proper nouns, no phonetic covariates were integrated in their analysis nor were appropriate inferential statistical methods applied. Therefore, there are many reasons to be sceptical of their results.

In another study, Li et al. (1999) measured S duration in child-directed speech with data originally elicited for another study (see Swanson & Leonard, 1994, on vowel durations in function words). Their study found plural S to be longer than third singular S. However, as the study originally was not designed for this endeavour, half of all plural items occurred sentence-finally, while almost all third person singular items occurred sentence-medial. The durational difference found between the suffixes may hence be attributed to effects of phrase-final lengthening (e.g. Klatt, 1976; Wightman et al., 1992) rather than to inherent phonetic differences due to morphological categories.

In a more recent study, Seyfarth et al. (2017) conducted a production experiment to collect data on non-morphemic, plural, and third singular /s/ and /z/ durations. They found the non-morphemic variant to be shorter than the morphemic instances. However, they did not find differences between the voiced and the voiceless allomorphs during their analysis. This may be a worrisome result especially due to the small number of items with voiceless allomorphs (n = 6) as compared to the high number of items with voiced allomorphs (n = 20) in their data.

Most recently, Plag et al. (2019) found plural and genitive plural S to be of different durations. In their study, the genitive plural suffix showed significantly longer durations as compared to the plural suffix. An overview of the durational differences found in the aforementioned studies is given in Table 1.

Table 1. Overview of durational differences of word-final S found in previous studies.

Study	**Findings**
Zimmermann, 2016; Plag et al., 2017; Tomaschek et al., 2019	non-morphemic > plural > clitics
Walsh & Parker, 1983	plural > non-morphemic
Li et al., 1999	plural > 3rd singular
Seyfarth et al., 2017	plural > non-morphemic
Plag et al., 2019	genitive plural > plural

In sum, there is evidence that there may be durational differences between different types of S. However, while results of corpus studies are in line with each other, they might be flawed due to imbalanced data sets. Previous experimental studies, on the other hand, often rely on small data sets, and lack phonetic covariates, appropriate statistical methods, or a proper distinction of voiced and voiceless segments. Another crucial difference between corpus and experimental studies is the use of homophones. While all previous experimental studies restrict their data to homophone pairs, corpus studies take into consideration all words. The limitation to homophones and the resulting competition between their representations might be a problem in itself as it appears to be unclear how members of homophone pairs are stored and connected to their respective frequencies (see section 2.2.). In all cases, previous results were subject to potentially confounding effects of the lexical properties (e.g. potential storage effects, see e.g. Caselli et al., 2016) and contextual effects (e.g. phrase final lengthening, see e.g. Klatt, 1976; Wightman et al., 1992) of the items under investigation. Also, so far, no experimental study included clitics in their analysis.

A study is therefore called for that investigates the durational nature of different types of word-final S in English, preferably with carefully controlled data avoiding potentially confounding effects. This paper presents such a study investigating word-final S in English by means of a pseudoword production task. In this task, we elicited three types of word-final S: monomorphemic, plural, and clitic S (with the auxiliaries *is* and *has*). We will address some the issues of previous studies. That is, the use of pseudowords prevents potential lexical effects to confound our findings, while our highly controlled task evades the influence of contextual effects. Even though our data will also contain homophones to a certain extent, the individual members do not have lexical representations. That is, we can rule out effects of competition between homophonous lexical entries due their similar representations.

To answer the questions on durational differences between different types of S we test the two null hypotheses given in (1) and (2).

(1) Null Hypothesis 1
There is no durational difference between non-morphemic and morphemic word-final S in English.

(2) Null Hypothesis 2
There is no durational difference between different types of morphemic word-final S in English.

3. Method

3.1. Speakers and recordings

Forty native speakers of Southern British English took part in the experiment. Twenty-six of them were female and sixteen were male. Their mean age was 28.7 years, ranging from 19 to 58. Eight speakers were bi- or multilingual, and twenty-five speakers were from London while the other fifteen speakers were from other places in South Britain. The participants had no background in linguistics.

The recordings took place at Chandler House, University College London. The acoustic data were recorded with a Røde NT1-a microphone using an RME Fireface UC audio interface and sampled at 44.1 kHz, 16 bit.

3.2. Speech material

In total, 48 pseudowords adopting Berko-Gleason's (1958) pseudoword-paradigm were used in the production experiment. Following her reasoning, we assume phonetic effects found in pseudoword-paradigms to mirror linguistic reality. Our pseudowords followed the phonotactic constraints of English (Clements & Keyser, 1983) and contained a complex onset consisting of a plosive and an approximant (/pl/, /bl/, /kl/, /gl/, /pr/), and either a short vowel (/ɪ/, /ʌ/), a long vowel (/i:/, /u:/), or a diphthong (/aʊ/, /eɪ/) as nucleus. One half of the pseudowords had simple codas (/p/, /t/, /k/, /f/), while the other half had an additional voiceless alveolar fricative (/ps/, /ts/, /ks/, /fs/). The set of coda consonants preceding the S was chosen in such a way that the voiceless realization of the S allomorphs was elicited. Our study is restricted to the voiceless realization as clearest results have emerged from literature for voiceless S. Pseudowords with complex codas were used to elicit non-morphemic S, while pseudowords with simple codas were used to elicit morphemic types of S. The pseudowords used in the experiments are given in Table 2.

One issue when constructing pseudowords is their spelling. For vowels, orthographic representations were chosen following the highest phonotactically legal grapheme-phoneme probabilities (Gontijo et al., 2003). The aforementioned coda consonants, however, showed a variety of possible orthographic representations to choose from. That is, /p/ may be represented by <p> or <pp>, /t/ may be represented by <t> or <tt>, /k/ may be represented by <k>, <c>, or <ck>, and /f/ may be represented by <f> or <ph>. When combined with a coda-internal /s/, some additional options can be observed: /ks/ may not only be represented as <ks>, <cs> or <cks> but also as <x>, /ps/ may be represented as <ps>, <pps>, and <pse>, and /ts/ may be represented as <ts>, <tts>, and <tz>. The choice of orthographic representation is important for two reasons. First, when comparing two kinds of words variable representations add another source of variation of unclear consequences and should be avoided. Second, studies on the influence of number of letters on spoken language production have found that increasing the number of letters to represent a single sound may go together with longer durations in speech (e.g. Brewer, 2008). Based on these considerations, the following orthographic representations were chosen for all word-final clusters: /ks/ is represented uniformly in spelling as <ks>, /ps/ is represented uniformly as <ps>, /ts/ is represented uniformly as <ts>, and /fs/ is represented uniformly as <fs>.

A second potential problem with the pseudowords constructed for this study is their phonotactics. All our pseudowords are phonotactically legal, and their final consonant clusters (with /s/ as the second consonant) are not uncommon in multimorphemic words. However, in monomorphemic words these clusters are rarer, or, in the case of /fs/, even unattested (e.g. in CELEX, Baayen et al., 1995). The different phonotactic probabilities of these clusters could potentially influence the pronunciation of /s/ in our nonce words, especially when spoken in the contexts where these words receive a monomorphemic interpretation. We have included two measures in our regression models to control for phonotactic probability. First, we included the biphone probability sum (Vitevitch & Luce, 2004) as a general measure of phonotactic probability of the whole word-form. Second, to assess the potential effect of phonotactics on the difference between monomorphemic and suffixed words we used the biphone probability of the final clusters in monomorphemic words as a covariate in our regression models (all variables are explained in detail in section 4.1). The rationale behind biphone probability in monomorphemic words as a covariate is this: If the differences in phonotactic probability of the clusters between monomorphemic and multimorphemic words lead to differences in the production of the two kinds of words we should find a significant interaction between phonotactic probability on the one hand and type of word (monomorphemic vs. suffixed/cliticized words) on the other in the regression models.

Table 2. Orthographic representation of the completed stimuli set.

	ɪ	i:	u:	ʌ	aʊ	eɪ
items for morphemic S elicitation	glip	pleep	cloop	prup	bloup	glaip
	glit	pleet	cloot	prut	blout	glait
	glik	pleek	clook	pruk	blouk	glaik
	glif	pleef	cloof	pruf	blouf	glaif
items for non-morphemic S elicitation	glips	pleeps	cloops	prups	bloups	glaips
	glits	pleets	cloots	pruts	blouts	glaits
	gliks	pleeks	clooks	pruks	blouks	glaiks
	glifs	pleefs	cloofs	prufs	bloufs	glaifs

To elicit the types of S under investigation, 48 contexts and accompanying questions for S elicitation were created. The verbs directly following the pseudowords in these contexts were chosen in such a way that out of twelve verbs in total, three each started with a voiceless plosive (/p/, /k/), a vowel (/ɑ/, /i:/, /ə/, /eɪ/), a nasal (/m/, /n/), and an approximant (/w/, /l/). Examples are given in (3) to (6) with verbs in bold print (see Appendix A for all contexts). This was done to control for possible coarticulatory effects of either of these segmental classes with the preceding S.

(3) Every day, the *glips* **plays** with the cloops.
(4) Two days ago, the *glips* **ate** their lunch together.
(5) Tonight, the *glip's* **meeting** the cloop for a drink.
(6) The *glip's* **written** a love letter to the cloop.

To keep priming effects to a minimum, pseudowords were split into two groups. Each group consisted of 24 pseudowords, with 12 pseudowords used for morphemic S elicitation and 12 pseudowords used for non-morphemic S elicitation. This way we ensured that no single participant encountered a phonologically identical pseudoword as both morphologically simple and complex, i.e. no participant was to encounter /glɪps/ as both singular and plural/clitic item. Participants were distributed equally across both groups.

Each speaker produced 12 pseudowords for non-morphemic S elicitation and 12 different pseudowords for the elicitation of plural, *is*- and *has*-clitic S (see Table 3).

Table 3. Number and type of S elicitations per speaker.

non-morphemic S	plural S	is-clitic S	has-clitic S	total number of trials per speaker
12	12	12	12	48

To ensure that each pseudoword was elicited within each context, i.e. with each verb for each type of S, twelve pseudorandomized lists were created. The same twelve lists were used for both groups to keep them comparable. Additionally, types of S were alternated in such a way that no type of S was elicited twice in a row. This was done to keep priming effects to a minimum.

3.3. Procedure

First, participants were introduced to the idea of a recently discovered far away planet. They were told that the inhabitants of this planet at first might appear bizarre, but engage in activities known to the participants, and not to worry about the unfamiliar names of the creatures. Second, the trial structure was explained, i.e. for each slide there will be pictures and names of alien creatures, a short explanation of a situation, and a question relevant to the situation which is to be answered aloud. Participants were then told to proceed in a natural pace and to take as much time as necessary to read and understand the aliens' names as well as the situations. To avoid possible confusion due to the simplicity of the task at hand, participants were made believe that they are part of a control group of an experiment originally designed for children. Before starting practice trials, participants were reminded to use the aliens' names instead of pronouns when answering. Then, a practice set of four contexts (see Appendix B) was used to familiarize the participants with the experimental procedure itself.

Figure 1

For each trial, the screen proceeded similarly (see Figure 1 as well as examples (7) to (10)): First, the pertinent pseudoword(s) were introduced. Two different pseudowords were introduced in non-morphemic, *is*- and *has*-clitic elicitation contexts, while only one pseudoword was introduced in plural settings. In either case, two images (van de Vijver & Baer-Henney, 2014) representing the pseudowords were used to create familiarity with the items under investigation. In all cases but plural, two images of different creatures were given, while in plural contexts two images of the same creature were used. Second, a context was introduced. Third, a question was given to elicit an answer with the pertinent type of S while the context slowly faded out. The fading out of the question forced the participants not to rely on the reading-aloud of the given context. This open format was chosen in order to elicit speech that is as natural as possible. By choosing such an open format one obviously runs the risk of eliciting a large proportion of responses that do not contain the desired forms. This drawback of our design was countered by having a large number of trials and participants. This strategy resulted in a sufficient number of observations. The experiment was carried out in a self-paced fashion; participants were instructed to progress in a contextually appropriate manner and at a speaking rate they considered to be normal.

(7) non-morphemic context
Introduction: This is a glaits. # And this is a pleeps.
Context: Every day, the glaits plays with the pleeps.

Question: What happens every day?
Answer: The glaits plays with the pleeps.

(8) plural context
Introduction: This is a glait. # And this is another one.
Context: Two days ago, the glaits ate their lunch together.
Question: What happened two days ago?
Answer: The glaits ate their lunch together.

(9) *is*-clitic context
Introduction: This is a glait. # And this is a pleep.
Context: Tonight, the glait's meeting the pleep for a drink.
Question: What's happening tonight?
Answer: The glait's meeting the pleep for a drink.

(10) *has*-clitic context
Introduction: This is a glait. # And this is a pleep.
Context: The glait's written a love letter to the pleep.
Question: What's happened?
Answer: The glait's written a love letter to the pleep.

3.4. Labels and measurements

As a first step, all recordings were manually transcribed on the utterance level. Using the freely available WebMAUS Basic system (Schiel, 1999; Kisler, et al., 2017), a phonetic transcription and segmentation based on the manual transcription was created. This automated segmentation was then manually checked by six trained annotators using the software Praat (Boersma & Weenink, 2020). Boundaries marking the beginning of an item or S were moved to the nearest zero crossing where both spectrogram and waveform indicated the initiation of the gesture for the respective segment, following laid out segmentation criteria based on features of specific sounds as described in the phonetic literature (e.g. Ladefoged, 2003). In the case of S, the boundaries were set to the zero crossing closest to the onset and offset of the friction visible in the waveform (see Figure 2). If a pause followed the S, the boundary was set to the point where the friction of the S dropped to silence.

Figure 2

The reliability of the segmentation criteria was verified by trial segmentations, in which it was ensured that all annotators placed boundaries with only very small variations. Each annotator worked on a disjoint set of items; segmentation criteria were regularly re-verified in meetings of the annotators. After the segmentation process, a Praat script was used to extract the item, its phonetic transcription and its duration, as well as the S duration itself. If applicable, the duration of the following pause was also extracted. Additionally, the preceding and the following word were extracted as well.

3.5. Pre-processing

A part of the 1920 (40 participants * 48 utterances) recorded data points had to be excluded from analysis for one or more of the following reasons. If an utterance did not include a word-final S, this utterance was discarded (n=599). A high number of failures to produce final S was expected especially with the clitics since participants could use a different tense form, or the full form of the auxiliary. It was also expected that participants would produce wrong pronunciations (including those with the final S) of the newly encountered written word-forms, as the participants had to retrieve them from short-term memory after the fading out of the context. Additionally, utterances containing stutter or hesitation (n=29), or replacement of pseudowords by pronouns (n=15) were excluded as well. Some utterances were ungrammatical (n=9), while other utterances contained pseudowords that were not part of the original set of pseudowords (n=8). Cases where the interpretation of the final /s/ was ambiguous presented another problem (n=114). An example of such a case is given in (11) where a *has*-clitic was expected. Note that two pseudowords without a non-morphemic word-final S were introduced, while either a non-morphemic S or *has*-clitic S was produced for the item under investigation, and most likely a non-morphemic word-final S for the second pseudoword. As for regular inflected verbs there was no way to decide which type of S had been produced in such cases, such utterances were discarded.

(11) Introduction: This is a glait. # And this is a pleep.
Context: The glait's attended concerts with the pleep many times.
Question: What's happened many times?
Answer: The glaits attended many concerts with the pleeps many times.

After exclusions, 1146 data points remained in the final data set. The final data set as well as the analysis and results discussed in the following sections can be found at https://osf.io/j4wxc/?view_only=b5399ef1adae4b679c4100d4b8ea6011.

4. Analysis

4.1. Covariates

The set of covariates chosen for the present study is similar to that of other studies on phonetic effects of morphological structure (e.g. Pluymaekers et al., 2005b, 2010; Hanique et al., 2013; Plag et al., 2017). In the following, covariates are briefly discussed.

SPEAKINGRATE. As speaking rate is a self-evident variable affecting segment durations, this was controlled for. Speaking rate was computed as the number of syllables in an utterance divided by the duration of the utterance and finally centred (Robinson & Schumacker, 2009; Afshartous & Preston, 2011; Winter, 2019). The computation was done automatically in Praat (de Jong & Wempe, 2008). This way of computing speaking rate is similar to those utilized in previous studies (e.g. Plag et al., 2017).

BASEDURLOG. Indicating a more local speaking rate (e.g. Plag et al., 2017), base duration was measured as well. Base duration in this case is equal to the summed duration of all word-internal segments preceding the S under investigation. That is, the stem of complex items and the segmental string without the final S of morphologically simple items is henceforth considered as base. We log-transformed and centred the base duration and called this variable BASEDURLOG.

PAUSEDUR & PAUSEBIN. In order to account for final-lengthening effects, all stretches of silence between the offset of the word-final S and the onset of the following word were measured. Silence of 50 ms and above was considered as pause (Lee & Oh, 1999; see also Zvonik & Cummins, 2003, and Krivokapić, 2007, on short pause durations in-between short phrases). The closure durations of following plosives were taken into account by subtracting the mean closure duration of the pertinent plosive (mean values for /p, t, k/ adopted from Yao, 2007) from the measured stretch of silence. Only if the resulting duration was above the aforementioned threshold, it was considered a pause. Pause measurements were included as the continuous variable PAUSE as well as the binary variable PAUSEBIN (with the levels `pause` and `no_pause`).

ITEM & TRANSCRIPTION. Pseudowords were sometimes produced with varying segmental make-up. We therefore included both the orthographic representation of the pseudoword, and a phonological transcription of the word as spoken as two variables. These covariates were labelled ITEM and TRANSCRIPTION.

NEIGHBOURHOODDENSITY & NEIGHBOURHOODFREQUENCY. Neighbourhood densities and frequencies were included as covariates as the number of neighbours may influence phonetic reduction (e.g. Gahl et al., 2012). Both neighbourhood measures were taken from the CLEARPOND database (Marian et al., 2012). That is, NEIGHBOURHOODDENSITY describes the number of words differing in one segment from the item in question (Marian et al., 2012: 3), while NEIGHBOURHOODFREQUENCY describes the mean frequency (per million) of these neighbouring words.

BIPHONEPROBSUM & BIPHONEPROBSUMBIN. A potential factor influencing the duration of a word in running speech is its predictability in context. The more predictable, the shorter the duration (e.g. Pluymaekers et al., 2005a; Bell et al., 2009; Torreira & Ernestus, 2009). Such a word bigram frequency, however, is not applicable to pseudowords for obvious reasons. Instead, the summed biphone probability was used analogously as a comparable measure. The summed biphone probability for each pseudoword and its phonological variants was calculated by the Phonotactic Probability Calculator (Vitevitch & Luce, 2004). Additionally, a binary covariate based on the summed biphone probability was created. The threshold for low vs. high summed biphone probability for biphoneProbSumBin was the mean of the continuous covariate. That is, all values below the mean were considered to be `low`, while all values above the mean were taken as `high`.

list & slideNumber. To account for possible durational differences due to priming and similar effects, the list number (1-12) and the point of occurrence during the experiment of the individual item were also included.

preC. It has been shown that the consonant preceding word-final S may influence the duration of word-final /s/ (e.g. Umeda, 1977: 853). In particular, Umeda (1977: 853) finds that /s/ becomes shorter after plosives, and longer after the fricative /θ/ (and this presumably also holds for /s/ after the fricative /f/). We therefore included the consonant preceding the final /s/ as a covariate, preC.

biphoneProb. For the reasons outlined in section 3.2 we included the probability of the final biphones /fs/, /ks/, /ps/ and /ts/ in monomorphemic words as a covariate. biphoneProb was computed on the basis of the transcriptions of all monomorphemic words in CELEX (Baayen et al., 1995).

folSeg & folType. To account for potential effects of the following word on the duration of S (e.g. Klatt, 1976; Umeda, 1977), these were included in regard to their onset segment adjacent to the word-final S. This was included in its phonological representation in folSeg (i.e. `k` for cooked) as well as in its segmental class by folType (i.e. approximant `APP` for *listen*, fricative `F` for *find*, nasal `N` for *know*, plosive `P` for *cook*, vowel `V` for *eat*).

speaker / gender / age / location / monoMultilingual. Speaker ID was included to account for inter-speaker differences in production. gender, age, and information on the place in which the bigger part of a participant's live was spent (location) were included as well as they may influence phonetic realizations. Additionally, participants who were early bilinguals were categorized as multilingual, while all other participants were categorized as monolingual in monoMultilingual[1].

4.2. Collinearity

One issue to address when fitting a model to a multitude of similar covariates is collinearity (e.g. Tomaschek et al., 2018). To avoid such issues, covariates were tested for correlation using the languageR package (Baayen & Shafaei-Bajestan, 2019).

Correlation was checked for item and transcription (*rho*=0.82, *p*<0.001, Spearman), pauseDur and pauseBin (*rho*=0.87, *p*<0.001, Spearman), neighbourhoodDensity and neighbourhoodFrequency (*rho*=0.86, p<0.001, Spearman), biphoneProbSum and biphoneProbSumBin (*rho*=0.87, p<0.001, Spearman), preC and biphoneProb (rho=0.38, p<0.001, Spearman), and for folSeg and folType (*rho*=-0.74, p<0.001, Spearman).

To avoid collinearity, the following procedure was adopted. For each pair of variables with a correlation of *rho*>0.5, two models containing only one of two variables were created and compared. This allowed us to decide which of the covariates under discussion was a stronger predictor for our dependent variable. This covariate was then

[1] Psycholinguistic experiments are standardly done with monolingual speakers (mostly of English, and mostly in the U.S.). In the multicultural context of a large European city like London, experiments with student populations necessarily involve speakers that are multilingual (with varying degrees of competence). To control for this potential confound, we added the variable monoMultilingual. While there are studies of phonetic duration in bilingual speech (e.g. Mack, 1982; Lee et al., 2012) the effect of mono-/multilingualism on the duration on word- final S has not been explored yet.

kept while the other one was no longer used. This led to the exclusion of ITEM, PAUSEDUR, NEIGHBOURHOODFREQUENCY, BIPHONEPROBSUM, and FOLSEG.

4.3. Statistical Analysis

Differences in consonant duration may play out as differences in absolute duration or as differences in relative duration (e.g. with gemination, e.g. Ridouane & Hallé, 2017; Ben Hedia, 2019; Oh & Redford, 2012). Some previous analyses of the duration of S (Plag et al., 2017) have therefore looked at both absolute and relative duration, and the present paper will also present these two types of analyses. In the first analysis (section 5.1) we used absolute duration of S as the dependent variable, whereas in the second analysis (section 5.2), the duration of S relative to the duration of the whole word is used as the dependent variable. Relative duration (i.e. the variable PROPORTIONOFS) was calculated by dividing the absolute duration of the S by the duration of the whole word.

In order to analyse our data, models were fitted using linear mixed-effects regression in R (R Core Team, 2019) using RStudio (RStudio Team, 2018) and as implemented by lme4 (Bates et al., 2015), lmerTest (Kuznetsova et al., 2017), and LMERConvenienceFunctions (Tremblay et al., 2015).

The dependent variable, duration of S, was log-transformed and centred following standard procedures to reduce the potentially harmful effect of skewed distributions in linear regression models (Winter, 2019). The name of this variable is SDURLOG. PROPORTIONOFS did not have a skewed distribution and no transformation was necessary.

Following the standard backward stepwise selection process (e.g. Baayen, 2008), the first models containing the explanatory variable TYPEOFS (with levels `nm` = non-morphemic; `pl` = plural; `is` = *is*-clitic; `has` = *has*-clitic) alongside all covariates provided in section 4.1. (with the exception of those excluded in 4.2.) were included. Random intercepts were included for SPEAKER, TRANSCRIPTION, LIST, SLIDENUMBER and AGE. Following the 'keep it maximal' policy of Barr et al. (2013), we initially also included a random slope for TYPEOFS by SPEAKER.

This full model was then continuously reduced through step-wise exclusion of non-significant factors. A factor was considered significant if it passed all of three tests. First, its F-value in the pertinent model had to yield a value below -2 or above 2. Second, the Akaike information criterion (AIC) of the model including the variable had to be lower than the AIC of a comparable model without the variable under discussion. Third, the results of a log-likelihood test comparing the model with to a model without the variable had to yield a *p*-value below the 0.05 threshold, indicating a significant improvement of the model containing the pertinent variable. This process was verified by using the step function of R, resulting in an identical model. We also eliminated all random intercepts and slopes that did not significantly improve the model in a log-likelihood test. Thus, we aimed for a meaningfully reduced random effect structure, following the criticism by Matuschek et al. (2017).

At the last stage of the model fitting process, the final model needed trimming of the residuals (e.g. Baayen & Milin, 2010). We removed data points with residuals larger than 2.5 standard deviations to ensure a satisfactory residual distribution. This resulted in a loss of 9 data points (0.8 %) and led to a satisfactory distribution of the residuals.

4.4. Overview of the data

An overview of all variables and their distribution is given in

Table 4 and Table 5.

Table 4. Summary of the dependent variable and numerical predictors in the final data set.

Dependent variable	Mean	St. Dev.	Min	Max
sDurLog	0.002	0.388	- 1.201	1.098
Numerical predictors	**Mean**	**St. Dev.**	**Min**	**Max**
speakingRate	-0.000	0.899	2.250	3.540
baseDurLog	0.072	0.194	0.000	3.559
pauseDur	0.072	0.193	0.000	3.559
neighbourhoodFrequency	27.345	84.645	0.000	412.027
biphoneProbSum	0.013	0.007	0.005	0.031
biphoneProb	0.001	0.002	0.000	0.004
age	28.740	9.743	19.000	58.000

Table 5. Summary of categorical predictors and the dependent variable in the final data set.

Categorical predictors	Levels							
item	48							
transcription	67							
NeighbourhoodDensity	0: 419	1: 238	2: 165	3:107	4: 14	5: 114	6: 32	7: 30
pauseBin	no: 777	yes: 342						
biphoneProbSumBin	low: 856	high: 263						
list	24							
slideNumber	48							
preC	f: 273	k: 292	p: 281	t: 273				
folSeg	18							
folType	APP: 299	F: 12	N: 230	P: 300	V: 278			
speaker	40							
gender	2							
location	London: 636	elsewhere: 483						
monoMultilingual	monolingual: 871	multilingual: 248						
Explanatory variable	**Levels**							
typeOfS	nm: 308	pl: 373	is: 284	has: 154				

5. Results

5.1. Absolute Duration

Figure 3 shows the distribution of the observed durations of non-morphemic, plural, *is*- and *has*-clitic S in the untrimmed data set. On average, non-morphemic S duration is 134 ms, which is about 13 ms longer than plural S with a mean duration of 121 ms. The mean duration of the *is*-clitic is 103 ms and the mean duration of the *has*-clitic is 94 ms.

Figure 3

While this may be an interesting result in itself, a multivariate analysis as described in the previous sections should be used to control for the many potentially intervening influences of the described covariates mentioned in section 4.1.

In our final model, fitted according to the procedure described above, we find main effects of type of S (typeOfS), speaking rate (speakingRate), base duration (baseDurLog), pause (pauseBin), biphone probability sum (biphoneProbSumBin), preceding consonant (preC), biphone probability (biphoneProb), following segmental type (folType), and mono-/multilingualism (monoMultilingual). There was no significant interaction between typeOfS and biphoneProb (F=0.4627, p=0.71, ANOVA of full model).

Regarding the random effects, only speaker-specific random intercepts turned out to significantly improve the model fit. The *p*-values for the analysis of variance of the final model are given in Table 6.

Table 6. *p*-values of fixed effects in the final model, fitted to the log-transformed durations of S.

	Sum Sq	Mean Sq	NumDF	DenDF	F.value	Pr (> F)
typeOfS	5.312	1.771	3	1089.66	33.338	0.000
speakingRate	0.230	0.230	1	1117.09	4.324	0.038
baseDurLog	9.466	9.466	1	1079.58	178.220	0.000
pauseBin	6.970	6.970	1	1110.28	131.235	0.000
biphoneProbSumBin	0.398	0.398	1	1082.26	7.492	0.006
biphoneProb	0.338	0.338	1	1079.25	6.360	0.012
preC	0.623	0.208	3	1080.29	3.910	0.009
folType	2.677	0.669	4	1081.55	12.598	0.000
monoMultilingual	0.345	0.345	1	37.37	6.498	0.015

The marginal *R*-squared value of the model is 0.46, that is fixed effects explain 46 percent of the variation in our data. The variance explained by the entire model is 61 percent as obtained by the conditional *R*-squared value of 0.61 (for marginal and conditional *R*-squared value computation see Nakagawa et al., 2017; values were computed with the MuMIn package, Barton, 2019).

The estimates of the final model and their *p*-values are given in Table 7. The reference levels for the categorical predictors are: for typeOfS it is non-morphemic S, for pauseBin it is `no-pause`, for biphoneProbSumBin it is `low`, for preC it is `t`, for folType it is `approximant`, and for monoMultilingual it is `monolingual`. All coefficients can be interpreted as changes relative to these reference levels.

Table 7. Fixed-effect coefficients and p-values as computed by the final model (mixed-effects model fitted to the log-transformed and centred durations of S).

	Estimate	Std. Error	df	*t*-value	Pr (> |t|)
(Intercept)	0.096	0.034	98.81	2.814	0.000
typeOfSpl	-0.114	0.019	1094.00	-6.062	0.000
typeOfSis	-0.178	0.020	1096.00	-8.839	0.000
typeOfShas	-0.196	0.024	1091.00	-8.14	0.000
speakingRate	-0.021	0.010	1117.00	-2.079	0.038
baseDurLog	0.586	0.044	1080.00	13.35	0.000
pauseBinpause	0.206	0.018	1110.00	11.456	0.000
biphoneProbSumBinhigh	0.047	0.017	1082.00	2.737	0.006
biphoneProb	0.069	27.53	1079.00	2.522	0.012
preCf	0.061	0.020	1081.00	-3.044	0.003
preCk	0.055	0.020	1082.00	-0.303	0.006
preCp	0.050	0.020	1079.00	2.522	0.012
folTypeF	0.012	0.070	1084.00	0.171	0.864
folTypeN	-0.036	0.021	1079.00	-1.764	0.078
folTypeP	-0.045	0.019	1080.00	-2.384	0.017
folTypeV	-0.136	0.020	1082.00	-6.85	0.000
monoMultilingualmultilingual	-0.152	0.059	37.37	-2.549	0.015

Effect size of individual predictors was checked by fitting models that lacked a particular predictor, and comparing their marginal *R*-squared values to those of the final model. The results are reflected in the hierarchy given in (12a). The decrease in *R*-squared is greatest when removing baseDurLog, followed by pauseBin, and so forth.

(12) (a) baseDurLog >> pauseBin >> typeOfS >> monoMultilingual >> folType >> speakingRate >> biphoneProbSumBin >> PreC >> biphoneProb

(b) typeOfS, folType, baseDurLog >> pauseBin >> biphoneProbSumBin >> preC >> monoMultilingual >> speakingRate >> biphoneProb

Additionally, we used ANOVAs to check whether a model that lacked a certain predictor performed better than a model that lacked a different predictor. The hierarchy in (12b) reflects the results. Models that either lack typeOfS, folType or baseDurLog do not show a significant difference in pair-wise ANOVAs. Models lacking either one of these predictors perform significantly worse than models lacking biphoneProbSumBin. Additionally, models lacking pauseBin do perform significantly better than models lacking baseDurLog, but perform significantly worse than models lacking biphoneProbSumBin. Models lacking biphoneProbSumBin perform significantly worse than models lacking preC, while models lacking preC perform significantly worse than models lacking monoMultilingual. speakingRate appears to perform worse than all other predictors with the exception of biphoneProb, which is the weakest of all predictors. Overall, the morphological status of an S appears to be a rather strong predictor of its acoustic duration.

Figure 4 shows the effect of the numerical variables included in the final model on S duration. The estimated values of the dependent variable and the base duration are back-transformed into seconds. Speaking rate and base duration show effects in the expected direction. With faster speech, S becomes shorter (panel A), while longer base durations also come with longer S durations (panel B). Higher biphone probability leads to longer S durations (panel C).

Figure 4

The partial effects of the categorical variables included in the final model are illustrated in Figure 5. S duration is longer if the S is followed by a pause (panel A), which can be interpreted as a clear case of phrase-final lengthening (e.g. Cooper & Danly, 1981). Higher biphone probability leads to longer S durations (panel B). There is also an effect of the preceding consonant: the plosive /t/ goes together with significantly shorter S durations than /k/ and /f/ (panel C). S duration is significantly shorter when followed by a vowel, while all other differences between following consonants are minor in nature (panel D). Lastly, monolingual speakers produce longer S durations than bi- or multilingual speakers (panel E).

Figure 5

The effect of the variable of interest, i.e. TYPEOFS, is plotted in Figure 6. As above, the values of the dependent variable are back-transformed into seconds.

Figure 6

We can see that there are durational differences between the different types of S. The results of pair-wise comparisons of the predicted means using Tukey contrasts (as implemented by the multcomp package for R, Hothorn et al., 2008) are summarized in Table 8.

Table 8. Multiple comparisons of means of duration of S (Tukey contrasts). Significant codes: ‘***’ p < 0.001, ‘**’ p < 0.01, ‘*’ p < 0.05.

	Estimate	Std. Error	*z*-value	Pr (> |z|)	
pl – nm	-0.114	0.019	-6.062	< 0.001	***
is – nm	-0.018	0.020	-8.839	< 0.001	***
has – nm	-0.196	0.024	-8.140	< 0.001	***
is – pl	-0.064	0.019	-3.294	0.005	**
has – pl	-0.082	0.023	-3.503	0.003	**
has – is	-0.018	0.023	-0.766	0.868	

Based on the Tukey tests, the comparison of the different types of S yields the significant contrasts shown in Table 9. If we look at the different durations given in Table 10, the following hierarchy emerges: non-morphemic > plural > *is*-/*has*-clitic.

Table 9. Significant contrasts in duration between different types of S. Significant codes: ‘***’ p < 0.001, ‘**’ p < 0.01, ‘*’ p < 0.05.

	nm	pl	*is*	*has*
nm	n.a.	***	***	***
pl		n.a.	**	**
is			n.a.	
has				n.a.

Table 10. S durations as estimated by the final model using non-centred data. All values are back-transformed to seconds. Values given are estimated for items without following pause, high biphone sum probability, monolingual speakers, and across all preceding and following segment types.

TYPEOFS	Mean
non-morphemic	0.224
plural	0.200
is-clitic	0.187
has-clitic	0.184

To summarize, the durational differences between non-morphemic and all other types of S, as well as the durational difference between plural and the clitics are significant, while there is no significant durational difference between both clitics. Non-morphemic S is longest in duration, followed by plural S, which in turn is followed by clitic S.

5.2. Relative Duration

The results for relative duration are very similar to those of absolute duration. The p-values for the analysis of variance of the final model are given in Table 11. Table 12 shows the coefficients for the final model. All effects go in the same direction as in the analysis of absolute duration. The only predictors that have lost significance when compared to the model for absolute duration are biphoneProb, preC and speakingRate.

Table 11. p-values of fixed effects in the final model, fitted to the relative durations of S.

	Sum Sq	Mean Sq	NumDF	DenDF	F.value	Pr (> F)
typeOfS	0.161	0.054	3	1070.68	25.510	0.000
pauseBin	0.186	0.186	1	1101.26	88.518	0.000
biphoneProbSumBin	0.015	0.015	1	36.32	6.917	0.012
folType	0.071	0.018	4	1063.31	8.389	0.000
monoMultilingual	0.010	0.010	1	37.81	4.561	0.039

Table 12. Fixed-effect coefficients and p-values as computed by the final model (mixed-effects model fitted to the relative durations of S).

	Estimate	Std. Error	df	*t*-value	Pr (> |t|)
(Intercept)	0.299	0.007	89.73	45.827	0.000
typeOfSpl	-0.019	0.004	1085.00	-5.157	0.000
typeOfSis	-0.031	0.004	1070.00	-7.651	0.000
typeOfShas	-0.035	0.005	1067.00	-7.260	0.000
pauseBinpause	0.033	0.004	1101.00	9.408	0.000
biphoneProbSumBinhigh	0.013	0.005	36.32	2.630	0.012
folTypeF	0.001	0.014	1068.00	0.086	0.931
folTypeN	-0.006	0.004	1061.00	-1.409	0.159
folTypeP	-0.007	0.004	1056.00	-1.708	0.088
folTypeV	-0.022	0.004	1063.00	-5.568	0.000
monoMultilingualmultilingual	-0.024	0.011	37.81	-2.136	0.039

The differences in the means show the same pattern as in the analysis of absolute duration, as can be seen in Table 13.

Table 13. Multiple comparisons of means of relative duration of S (Tukey contrasts). Significant codes: '***' p < 0.001, '**' p < 0.01, '*' p < 0.05.

	Estimate	Std. Error	*z*-value	Pr (> |z|)	
pl – nm	-0.019	0.004	-5.157	< 0.001	***
is – nm	-0.031	0.004	-7.651	< 0.001	***
has – nm	-0.035	0.005	-7.260	< 0.001	***
is – pl	-0.011	0.004	-2.936	0.017	*
has – pl	-0.015	0.005	-3.300	0.005	**
has – is	-0.004	0.005	-0.854	0.827	

The analysis of relative duration thus fully supports the results for absolute duration.

6. Discussion

Following in the footsteps of previous studies on durational differences between different types of S, we tested whether the morphological category of word-final S has an influence on its acoustic duration in speech production. In order to avoid imbalanced data as in the case of corpus studies, we used a production experiment, i.e. speech material elicited by the means of highly controlled contexts of a production task. For the first time in this context, pseudowords instead of real words were used to eliminate potentially confounding lexical effects.

We started out from two null hypotheses. The first null hypothesis stated that there is no durational difference between non-morphemic and morphemic word-final S. The second null hypothesis stated that there is no durational difference between different types of morphemic word-final S. Investigating these hypotheses with the elicited data, we find that both null hypotheses need to be rejected as type of S is indeed a significant predictor for S duration. That is, there are significant durational differences between non-morphemic and morphemic types of word-final S, with morphemic types of S being significantly shorter in duration than non-morphemic S. Also, there are significant durational differences between the plural suffix and the *is*- and *has*-clitic S, with plural S being significantly longer than clitic S and with no significant difference between the two clitics. Hence, type of S emerged as a strong, significant predictor of segmental duration.

6.1. Comparison of results to other studies

How do our results on word-final S in pseudoword context relate to the findings of previous studies? Let us first compare the results of the present paper to those of corpus studies on the same matter. The studies of Zimmermann (2016) on New Zealand English, and Plag et al. (2017) and Tomaschek et al. (2019) on North American English found significant differences in duration between non-morphemic and morphemic word-final S, with morphemic S being shorter than non-morphemic S. Additionally, the corpus studies also found suffix S to be significantly longer in duration as compared to clitic S. The results from the corpus studies on North American English and New Zealand English are identical, and the present study's results are completely in line with these studies. In sum, the same effects occur in three varieties in English, including Southern British English, and they occur in two corpus studies and one experimental pseudoword study.

Turning to previous experimental studies, we find differing results. Walsh & Parker (1983) also found differences between non-morphemic and morphemic S. However, their results go into the opposite direction, i.e. non-morphemic S was found to be shorter than morphemic S. Yet, as their study lacks inferential statistical methods and the inclusion of phonetic covariates, we cannot tell whether the small difference between mean values found by Walsh & Parker (1983) is actually meaningful.

Seyfarth et al. (2017) found durational differences between non-morphemic and morphemic S. However, similar to the findings of Walsh & Parker (1983), in their data non-morphemic S was shorter than morphemic S. That is, their results go into the opposite direction from the present findings. One has to note, though, that in their study only six words with word-final /s/ were used as against a majority of twenty words with word-final /z/. Even though they do not find voicing to be a significant predictor in their post-hoc analysis, one might suggest the small number of /s/ items and thus the lacking statistical power to be one plausible reason for this.

Results of both experimental studies (Walsh & Parker, 1983; Seyfarth et al., 2017) are subject to potentially confounding effects of the lexical and contextual properties of the items under investigation. Their finding of non-morphemic S being shorter than morphemic S may well be an artefact of such properties. The items used in the present study, however, are much less prone to be subject to such effects as they are pseudowords with no established representations in the speakers' mental lexicons.

Concerning clitic S production, our findings are in line with those by previous corpus studies, i.e. clitic S is shorter than non-morphemic and plural S. However, we cannot compare our results to previously reported ones by other experimental studies, as all previously conducted experimental studies did not investigate clitic S production.

6.2. Explanations and implications

This study's results raise important questions for established theories. Most evidently, it is unclear why there are durational differences between types of S at all. Why should non-morphemic S be longer than suffix S, which in turn is longer than clitic S? Which theory could account for such findings? These are fundamental questions, especially as the influence of unbalanced distributions as well as the confounding effects of lexical and contextual properties were ruled out as determining factors.

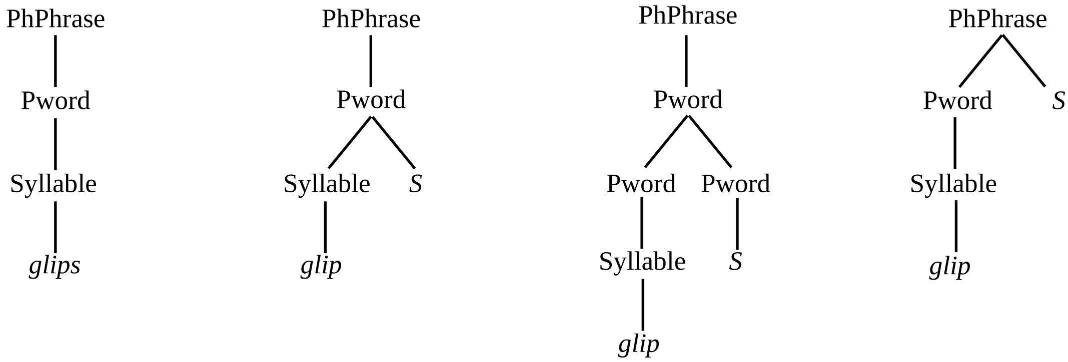
At an abstract level, our findings can be interpreted as evidence for morphological information in the phonetic signal, i.e. in post-lexical stages of speech production. This calls into question the distinction between lexical and post-lexical phonology, which has been an integral part of standard feed-forward formal theories of morphology-phonology interaction. Processes like bracket erasure are meant to erase any morphological information from retrieved segmental strings, such as one cannot trace any information about a sound's structural status in the acoustic signal. The findings of the present paper challenge such central tenets of lexical phonology and morphology as proposed by Kiparsky (1982).

Turning to psycholinguistic models of speech production, well-established models seem equally unable to account for our results. Levelt et al. (1999) assume that meaning is stored in the mental lexicon while the phonological makeup is composed of individually stored segments and syllables. These segments and syllables are retrieved in production by the articulator module, and do not account for differences in meaning. That is, morphology dependent phonetic detail is not part of such a model, as it is not part of the representation of a lexical entry and cannot be accommodated by the elements available to the articulator module. Hence, such an account is ruled out by our findings.

However, there are some alternative explanations imaginable. Let us first discuss a prosodic approach. In prosodic phonology (e.g. Booij, 1983), different types of word-final S are analysed as having different positions in the hierarchical prosodic configuration. These configurations co-determine the degree of integration of an S to the word it belongs to. These different degrees of integration might then emerge as durational differences between types of S in speech production.

Applying Selkirk's (1996) approach, non-morphemic S, uncontroversially, is an integral part of the prosodic word, as shown in (13). Goad (1998) analyses plural S as an 'internal clitic', which is adjoined to the highest prosodic constituent below the prosodic word, as shown in (14). In Goad (2002), however, plural S is analysed as an 'affixal clitic', like third person singular S in Goad et al. (2003) and Goad & White (2019), as shown in (15). The prosodic status of the cliticized auxiliary S is not entirely clear, but presumably it is best analysed as 'free clitic', as in (16).

(13) non-morphemic S | (14) plural S ‘internal clitic’ | (15) plural S ‘affixal clitic’ | (16) clitic S ‘free clitic’



The prosodic phonology approach thus posits a structural prosodic difference between non-morphemic S, plural S and clitic S. This prosodic difference might be mirrored in durational differences. It is, however, not so clear, what particular phonetic effects this approach would predict, and by which processing mechanism the structural prosodic differences would be translated into different articulations. The most plausible prediction would be that closer integration into the prosodic word would correlate with shorter durations. What we find is quite the opposite: the more prosodically integrated the S, the longer it becomes, which is the opposite pattern that one would expect under the prosodic approach.

Another possible explanation for our findings lies within exemplar-based models (e.g. Goldinger, 1998; Bybee, 2001; Pierrehumbert, 2001, 2002; Gahl & Yu, 2006). In such models, lexemes are linked to a frequency distribution over their phonetic outcomes as experienced by the individual speaker. These distributions are updated with each new experience: experienced subtle subphonemic differences then may result in representations mirroring these properties. While such an account may explain how durational differences between different types of word-final S may result from stored phonetic representations, it leaves open the question of how such systematic differences between clouds of exemplars come about in the first place.

One possible explanation for our findings can be found within the computational modelling framework of naïve discriminative learning (NDL) which is based on simple but powerful principles of discriminative learning theory (Rescorla, 1988; Ramscar & Yarlett, 2007; Ramscar et al., 2010; Baayen et al., 2011). According to this theory, learning results from exposure to informative relations among events in the individual’s environment. Individuals use the associations between these events to create cognitive representations of their environment. Most importantly, associations and their resulting representations are updated constantly on the basis of new experiences. Associations are built between features (‘cues’, e.g. biphones) and classes or categories (‘outcomes’, e.g. different types of S) that co-occur in events in which the learner is predicting the outcomes from the cues (Tomaschek et al., 2019: 11). The relation between cues and outcomes is modelled mathematically by the so-called Rescorla-Wagner equations (Rescorla & Wagner, 1972; Wagner & Rescorla, 1972; Rescorla, 1988). Following these equations, an association strength or ‘weight’ increases every time a cue and an outcome co-occur, while it decreases if a cue occurs without the outcome in a learning event. This results in a continuous recalibration of association strengths, which is a crucial part of discriminative learning. So far, several studies have shown that NDL can successfully model various morphological phenomena, e.g. reaction times in studies on morphological processing (e.g. Baayen et al., 2011; Blevins et al., 2016).

Tomaschek et al. (2019) find the same patterning as Plag et al. (2017) in their data set (the complete Buckeye Corpus). They show that the different durations of S can be understood as following from the extent to which words' phonological and collocational properties can discriminate between the inflectional functions expressed by the S. The input features (cues) for their discriminative network were the words ('lexomes' as pointers to the meaning of the forms) in a five-word window centred on the S-bearing word and the biphones in the phonological forms of these words. These cues are associated with the inflectional functions of the S. Two main measurements emerged as significant predictors of S duration. The so-called 'activation' is a measure of an outcome's baseline activation, i.e. of how well an outcome is entrenched in the lexicon. The other measure is 'activation diversity', which quantifies the extent to which the cues in the given context also support other targets. The general pattern now is the following: When the uncertainty about the targeted outcome increases, the acoustic duration of S decreases. In other words, stronger support (both from long-term entrenchment and short-term from the context) for a morphological function leads to a longer, i.e. enhanced, acoustic signal.

This effect seems to run counter to the predictions of information theoretic accounts and probabilistic theories, according to which words and segments are realized shorter when they are less informative (Aylett & Turk, 2004; Jaeger, 2010; Cohen Priva, 2015). However, the effects are in line with studies showing that duration increases with increasing paradigmatic certainty (Kuperman et al. 2007; Cohen 2014; Tucker et al. 2019). For instance, Kuperman and colleagues found that the duration of a given interfix in Dutch compounds increases with increasing probability of this interfix (as against its competitors) in the left constituent family of the compound. With English S, the competing morphological functions constitute the paradigm within which the support for a particular function is gauged. It is these paradigmatic forces that play a key role in the acoustics of final S, and it appears that pseudowords produced in a natural speech context are subject to the same discriminative effects.

6.3. Pseudoword structure and its influence on results

While the use of pseudowords in phonetic experiments comes with a number of benefits (see section 3.2), it also raises some questions. First, there is the issue of phonotactic probability raised in section 3.2. Two measures of phonotactic probability (one for the whole word, the other for the final cluster) were included to address this issue. It turned out that phonotactic probability has a say in the productions of our pseudowords, as it has for real words. Crucially, there was no interaction between the type of S and the biphone probability of the cluster in monomorphemic words. This means that speakers produced these clusters in the same way, no matter whether the cluster occurred in the monomorphemic words or straddled the morphemic boundary between the stem and the S. The main effects of the phonotactic probability variables turned out to be rather weak, and, crucially, were properly controlled for in the regression analysis. In sum, the phonotactic probability of the final cluster does not seem to have unduly influenced the results.

Second, there might have been a problem with another aspect of the phonological structure of the pseudowords in the experiment, i.e. long-distance agreement of phonological features (Coetzee, 2005; 2008). Such OCP-effects might have arisen with pseudowords such as *pleep* (in which initial /p/ and final /p/ share all features) or *glik* (in which the initial and final sounds share the dorsal feature. Following the findings by Coetzee (2008), we coded a new variable to test this effect post-hoc empirically as an additional covariate and as an interacting term of TYPEOFS with the following levels: 'not well-formed' for pseudowords in which the initial and final consonant share all features (n=836), 'moderately well-formed' for pseudowords in which the initial and final consonant share the dorsal feature (n=147), and 'well-formed' for all remaining pseudowords (n=145). There was no significant main effect of this variable on the duration of S, nor a significant interaction with TYPEOFS. OCP effects thus cannot explain our results.

Third, after having carried out the experiments, it came to our attention that some of our pseudowords have real word relatives that are spelled differently but are phonologically identical. That is, *glits* corresponds to *glitz*, *glaiks* corresponds to *Gleicks*, *glif(s)* corresponds to *glyph(s)*, and *pleet(s)* corresponds to *pleat(s)*. These words might have unduly influenced our results and should perhaps not have been included into the statistical analysis. To check whether these items had any influence on the results, we created a data set containing all data but the three potentially offending items. Fitting the final model (as done in section 4.3) to this new dataset resulted in basically the same findings, i.e. TYPEOFS was still a significant predictor for S duration showing the same significant differences between non-morphemic, plural, and clitic items as presented in Table 9.

It has recently been shown that the notion of pseudoword is problematic in a more general way. The notion of pseudoword itself is usually based on the idea of the lexicon as a community construct. When talking about the mental lexicon, however, it is clear that what is an existing word and what is an unknown pseudoword is a matter of the individual speaker's mental lexicon. All participants in our experiment denied knowing any of the pseudowords used in this experiment when asked afterwards. At the community level, Google frequencies of pseudowords have been shown to be a robust predictor of reaction times in lexical decision tasks (e.g. Hendrix & Sun 2020). To test whether Google frequency had an effect on our results, the covariate GOOGLEFREQ was created containing the number of Google search hits for each pseudoword. The addition of this covariate as either fixed effect or interacting term to TYPEOFS resulted in its exclusion during the model simplification procedure.

6.4. Directions for future research and conclusion

The results of the present study may bring up further questions. First, assuming the durational differences found here and in previous studies are indeed systematic, one would also like to know whether language users are able to perceive them. This automatically leads to questions of whether all differences are perceivable or only some of them given our knowledge on the perception of differences in fricative durations (e.g. Klatt & Cooper, 1975). Secondly, if the durational differences are perceivable, another question naturally suggests itself: do users of a language not only perceive but also make use of such differences? These questions call for highly controlled perception and comprehension studies.

To summarize, this paper was first to investigate durational differences of different types of word-final S in English in pseudowords. The analysis yielded important evidence on the question of realizational differences between phonologically identical segments, showing that phonologically identical /s/ segments, such as non-morphemic and morphemic S, can indeed be phonetically distinct. Additionally, it also showed that there are realizational distinctions between different phonologically identical morphemic types of S. As these results were found using pseudowords, one can most likely exclude confounding effects of lexical properties, hence, durational differences between different types of S appear to be of a robust nature rather than a by-product of confounding factors. This leads to the conclusion that differences in S durations are due to the processing of the morphological information encoded in the pertinent type of S. In other words, morphological information may influence speech production in such a way that systematic subphonemic differences arise. This calls for revisions of current models on the relationship between morphology, phonology, and phonetic realization.

Acknowledgements

The authors are grateful to the members of the DFG Research Unit FOR2373 and the audience of several conferences (19th International Morphology Meeting, February 2020, Vienna; LabPhon 17, July 2020, Vancouver; UK Cognitive Linguistics Conference 2020, July 2020, Birmingham; 16. Phonetik und Phonologie Tagung, September 2020, Trier) for valuable input, to James White and Andrew Nevins allowing the production experiment to take place at Chandler House, University College London, and to Andrew Clark for his technical support. The usual disclaimers apply.

Funding Sources Statement

This research was funded by the Deutsche Forschungsgemeinschaft (Research Unit FOR2373 'Spoken Morphology', Project BA6523/1-1 'Final S in English: The role of acoustic detail in morphological learning'), which we gratefully acknowledge.

Statement of Ethics

The research performed in this paper has ethic approval of the ethics committee of the Linguistic Society of Germany and of the University College London (LING-2018-8-01). All participants signed a written informed consent form before participating in the production study and were provided with detailed information sheets.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Author Contributions Statement

Dominic Schmitz, Ingo Plag, and Dinah Baer-Henney conceived of the presented idea and planned the experiment. Dominic Schmitz carried out the experiment and, with Ingo Plag, performed the statistical analysis with input from Dinah Baer-Henney. Dominic Schmitz wrote the manuscript; it was proofread by all authors. All authors provided critical feedback and helped shape the research, analysis, and manuscript.

References

Afshartous, D., & Preston, R. A. (2011). Key results of interaction models with centering. *Journal of Statistics Education*, 19. doi: 10.1080/10691898.2011.11889620

Aylett, M., & Turk, A. (2004). The Smooth Signal Redundancy Hypothesis: A Function Explanation for Relationships between Redundancy, Prosodic Prominence, and Duration in Spontaneous Speech. *Language and Speech, 47*, 31-56. doi: 10.1177/00238309040470010201

Baayen, R. H. (2008). *Analysing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, *3,* 12-28. doi: 10.21500/20112084.807

Baayen, R. H., & Shafaei-Bajestan, E. (2019). languageR [R package]. Version 1.5.0, retrieved August 2019. https://CRAN.R-project.org/package=languageR

Baayen, R. H., Milin, P., Filipović Durdević, D., Hendrix, P., & Marelli, M.. (2011). An amorphous model for morphological processing in visual comprehension based on naïve discriminative learning. *Psychological Review, 118*, 438–482. doi: 10.1037/a0023851

Baayen, R. H. , Piepenbrock, R., & Gulikers, L. (1995). The CELEX Lexical Database (CD-ROM). *Linguistic Data Consortium*. Philadelphia, PA: University of Pennsylvania,

Barr, J. D., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *3*, 255-278. doi: 10.1016/j.jml.2012.11.001

Barton, K. (2019). MuMIn: Multi-Model Inference [R package]. Version 1.43.6, retrieved August 2019. https://CRAN.R-project.org/package=MuMIn

Bates. D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*, 1-48. doi: 10.18637/jss.v067.i01

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. doi: 10.1016/j.jml.2008.06.003

Ben Hedia, S. (2019). *Gemination and degemination in English affixation. Investigating the interplay between morphology, phonology and phonetics*. Studies in Laboratory Phonology 8. Berlin: Language Science Press.

Ben Hedia, S., & Plag, I. (2017). Gemination and degemination in English prefixation: Phonetic evidence for morphological organization. *Journal of Phonetics*, *62*, 34-49. doi: 10.1016/j.wocn.2017.02.002

Berko-Gleason, J. (1958). The child’s learning of English morphology. *Word*, *14*, 150-177. doi: 10.1080/00437956.1958.11659661

Blevins, J. P., Ackerman, F., & Malouf, R. (2016). Morphology as an adaptive discriminative system. In H. Harley & D. Siddiqi (Eds.) *Morphological metatheory*. Amsterdam and Philadelphia: John Benjamins, 271-301.

Boersma, P., & Weenink, D. (2020). Praat:doing phonetics by computer [Computer program]. Version 6.0.49, retrieved March 2019. http://www.praat.org/

Booij, G. (1983). Principles and Parameters in Prosodic Phonology. *Linguistics*, *21*, 249-280. doi: 10.1515/ling.1983.21.1.249

Brewer, Jordan. (2008). *Phonetic Reflexes of Orthographic Characteristics in Lexical Representation* [PhD Thesis]. The University of Arizona.

Bybee, J. L. (2001). *Phonology and language use*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511612886
Caselli, N. K., Caselli, M. K., & Cohen-Goldberg, A. M. (2016). Inflected words in production: Evidence for a morphologically rich lexicon. *Quarterly Journal of Experimental Psychology, 69*, 434-454. doi: 10.1080/17470218.2015.1054847
Cho, T. (2001). Effects of Morpheme Boundaries on Intergestural Timing: Evidence from Korean. *Phonetica, 58*, 129-162. doi: 10.1159/000056196
Chomsky, N. & Halle, M. (1968). *The sound pattern of English*. New York: Harper and Row.1
Clements, G. N., & Keyser, S. J. (1983). *CV Phonology: A Generative Theory of the Syllable*. Cambridge, MA: MIT Press.
Coetzee, A. W. (2005). The Obligatory Contour Principle in the perception of English. In Frota, S., Vigario, M., & Freitas, M. J. (Eds.) *Prosodies*. New York: Mouton de Gruyter, 223-245.
Coetzee, A. W. (2008). Grammar is both categorical and gradient. In Parker, S. (Ed.) Phonological argumentation. Oakville, CT: Equinox Pub. Ltd, 9-42.
Cohen, C. (2014). *Combining structure and usage patterns in morpheme production: Probabilistic effects of sentence context and inflectional paradigms*. Berkeley: University of California, PhD Dissertation.
Cohen Priva, U. (2015). Informativity affects consonant duration and deletion rates. Laboratory Phonology, 6, 243-278. doi: 10.1515/lp-2015-0008
Cooper, W. E., & Danly, M. (1981). Segmental and Temporal Aspects of Utterance-Final Lengthening. *Phonetica, 38*, 106–115. doi:10.1159/000260017
de Jong, N., & Wempe, T. (2008). Praat Script Syllable Nuclei [Praat script]. Retrieved November 2019. https://sites.google.com/site/speechrate/Home/praat-script-syllable-nuclei-v2
Drager, K. (2011). Sociophonetic variation and the lemma. *Journal of Phonetics*, *39*, 694-707. doi: 10.1016/j.wocn.2011.08.005
Gahl, S. (2008). Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language*, *84*, 474-496. doi: 10.1353/lan.0.0035
Gahl, S., & Yu, A. C. L. (2006). Special issue on exemplar-based models in linguistics. *The Linguistic Review, 23*, 213-216. doi: 10.1515/TLR.2006.007
Gahl, S., Yao, Y., & Johnson, K. (2012). Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. *Journal of Memory and Language, 66*, 789-806. doi: 10.1016/j.jml.2011.11.006
Goad, H. (1998). Plurals in SLI: Prosodic deficit or morphological deficit? Language Acquisition, 7, 247-284. doi: 10.1207/s15327817la0702-4_6
Goad, H. (2002). Markedness in Right-edge Syllabification: Parallels across Populations. *Canadian Journal of Linguistics, 47*, 151-186. doi: 10.1017/S0008413100022933
Goad, H., & White, L. (2019). Prosodic effects on L2 grammars. *Linguistic Approaches to Bilingualism, 9*, 769-808. doi: 10.1075/lab.19043.goa
Goad, H., White, L., & Steele, J. (2003). Missing inflection in L2 acquisition: Defective syntax or L1-constrained prosodic representations? *The Canadian Journal of Linguistics/La revue canadienne de linguistique, 48*, 243-263. doi: 10.1017/S0008413100000669
Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105*, 251-279. doi: 10.1037/0033-295X.105.2.251

Gontijo, P. F. D., Gontijo, D., & Shillcock, R. (2003). Grapheme-phoneme probabilities in British English. *Behavior Research Methods, Instruments, & Computers, 35*, 136-157. doi: 10.3758/bf03195506

Gries, S. T. (2015). The most under-used statistical method in corpus linguistics: multi-level (and mixed-effects) models. Corpora, 10, 95-125. doi: 10.3366/cor.2015.0068

Hanique, I., Ernestus, M., & Schuppler, B. (2013). Informal speech processes can be categorical in nature, even if they affect many different words. *Journal of the Acoustical Society of America*, *133*, 1644-1655. doi: 10.1121/1.4790352

Hendrix, P., & Sun, C. (2020). A word or two about nonwords: frequency, semantic neighborhood density, and orthography-to-semantics consistency effects for nonwords in the lexical decision task. Journal of Experimental Psychology: Learning, Memory, and Cognition. doi: 10.1037/xlm0000819

Holling, H. (1983). Suppressor structures in the general linear model. *Educational and Psychological Measurement, 43*, 1-9. doi: 10.1177/001316448304300101

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous Inference in General Parametric Models. *Biometrical Journal, 50,* 346-363. doi: 10.1002/bimj.200810425

Jaeger, F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology, 61,* 23-62. doi: 10.1016/j.cogpsych.2010.02.002

Jurafsky, D., Bell, A., & Girand, C. (2002). The role of the lemma in form variation. In Warner, N., & Gussenhoven, C., (Eds.) *Papers in Laboratory Phonology, 7,* 3–34.

Kemps, R., Ernestus, M., Schreuder, R., & Baayen, R. H. (2005b). Prosodic cues for morphological complexity in Dutch and English. *Language and Cognitive Processes, 20*, 43-73. doi: 10.1080/01690960444000223

Kemps, R., Ernestus, M., Schreuder, R., & Baayen, R. H. (2005a). Prosodic cues for morphological complexity: The case of Dutch plural nouns. *Memory & Cognition, 33*, 430-446. doi: 10.3758/BF03193061

Kiparsky, P. (1982). Lexical morphology and phonology. In Yang, I. (Ed.) *Linguistics in the morning calm: Selected papers from SICOL*, 3-91. Seoul: Hanshin.

Kisler, T, Reichel, U. D., & Schiel, F. (2017). Multilingual processing of speech via web services. *Computer Speech & Language*, *45*, 326-347. doi: doi.org/10.1016/j.csl.2017.01.005

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustic Society of America, 59*, 1208-1221. doi: 10.1121/1.380986

Klatt, D. H., & Cooper, W. E. (1975). Perception of segment duration in sentence contexts. In Cohen, A., & Nooteboom, S. G., (Eds.) *Structure and process in speech perception*. Berlin: Springer, 69–89.

Krivokapić, J. (2007). Prosodic planning: Effects of phrasal length and complexity on pause duration. *Journal of Phonetics, 35*, 162-179. doi: 10.1016/j.wocn.2006.04.001

Kuperman, V., Pluymaekers, M., Ernestues, M., & Baayen, H. (2007). Morphological predictability and acoustic salience of interfixes in Dutch compounds. *journal of the Acoustical Society of America, 121*, 2261-2271.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. Journal of Statistical Software, 82, 1-26. doi: 10.18637/jss.v082.i13

Ladefoged, P. (2003). *Phonetic data analysis: An Introduction to Fieldwork and Instrumental Techniques*. Malden, MA: Blackwell.

Lavoie, L. (2002). Some influences on the realization of for and four in American English. *Journal of the International Phonetic Association, 32*, 175-202. doi:10.1017/S0025100302001032

Lee, S. A. S., & Iverson, G. K. (2012). Stop consonant productions of Korean-English bilingual children. *Bilingualism: Language and Cognition, 15*, 275-287. doi: 10.1017/S1366728911000083

Lee, S., & Oh, Y.-H. (1999). Tree-Based Modeling of Prosodic Phrasing and Segmental Duration for Korean TTS Systems. *Speech Communications, 28,* 283-300. doi: 10.1016/S0167-6393(99)00014-X

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. doi: 10.1017/s0140525x99001776

Levelt, W. J. M. & Wheeldon, L. R. (1994). Do speakers have access to a mental syllabary? *Cognition, 50*, 239-269. doi: 10.1016/0010-0277(94)90030-2

Li, H., L. Leonard & L. Swanson. (1999). Some differences between English plural noun inflections and third singular verb inflections in the input: The contribution of frequency, sentence position and duration. *Journal of Child Language, 26*, 531-543. doi: 10.1017/s030500099900392x

Lohmann, A. (2018). Time and thyme are NOT homophonous: a closer look at Gahl’s work on the lemma frequency effect including a reanalysis. *Language, 94,* e180-e190. doi: 10.1353/lan.2018.0032

Mack, M. (1982). Voicing-dependent vowel duration in English and French: Monolingual and bilingual production. *Journal of the Acoustical Society of America, 71*, 173-178. doi: doi.org/10.1121/1.387344

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-Linguistic Easy-Access Resource for Phonological and Orthographic Neighborhood Densities. *PLoS ONE, 7*, e43230. doi: 10.1371/journal.pone.0043230

Matuschek, H., Kliegl, R., Vasisth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language, 94,* 305-315. doi: 10.1016/j.jml.2017.01.001

Nakagawa, S., Johnson, P.C.D., Schielzeth, H. (2017). The coefficient of determination R^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface, 14*. doi: 10.1098/rsif.2017.0213

Oh, G. E., & Redford, M. A. (2012). The production and phonetic representation of fake geminates in English. *Journal of Phonetics, 40*, 82–91.

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In Bybee, J. & Hopper, P., (Eds) *Typological studies in language, Vol. 45. Frequency and the emergence of linguistic structure*, 137-157. doi: 10.1075/tsl.45.08pie

Pierrehumbert, J.B. (2002). Word-specific phonetics. In Gussenhoven, C., & Warner, N., (Eds.) *Papers in Laboratory Phonolog, 7*, 101-140.

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, OH: 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. doi: 10.1017/S0022226715000183

Plag, I., Lohmann, A., Ben Hedia, S., & Zimmermann, J. (2019). An <s> is an <s’>, or is it? Plural and genitive-plural are not homophonous. To appear in Körtvélyessy, L. & Stekauer, P. (Eds.) *Complex Words*. Cambridge: Cambridge University Press.

Pluymaekers, M., Ernestus, M., & Baayen, R. H. (2005a). Articulatory planning is continuous and sensitive to informational redundancy. *Phonetica, 62*, 146-159. doi: 10.1159/000090095

Pluymaekers, M., Ernestus, M., & Baayen, R. H. (2005b). Lexical frequency and acoustic reduction in spoken Dutch. *Journal of the Acoustical Society of American, 118*, 2564-2569. doi: 10.1121/1.2011150

Pluymaekers, M., Ernestus, M., Baayen, R. H., & Booij, Geert. (2010). Morphological effects in fine phonetic detail: The case of Dutch -igheid. In Fougeron, C., (Ed.) *Papers in Laboratory Phonology, 10,* 511-531. Berlin and New York: Mouton de Gruyter.

R Core Team. (2019). R: A Language and Environment for Statistical Computing. https://www.R-project.org

Ramscar, M. & D. Yarlett. (2007). Linguistic self-correction in the absence of feedback: A new approach to the logical problem of language acquisition. Cognitive Science, 31, 927-960. doi: 10.1080/03640210701703576

Ramscar, M., D. Yarlett, M. Dye, K. Denny & K. Thorpe. (2010). The effects of feature-label-order and their implications for symbolic learning. *Cognitive Science, 34*, 909–957. doi: 10.1111/j.1551-6709.2009.01092.x

Rescorla, R. A. (1988). Pavlovian conditioning. It's not what you think it is. *American Psychologist, 43,* 151-160. doi: 10.1037/0003-066X.43.3.151

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlocian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.) *Classical Conditioning II: Current research and theory*. New York: Appleton Century Crofts, 64-99.

Ridouane, R. & Hallé, P. (2017). Word-initial geminates: From production to perception. In H. Kubozono (Ed.) *The phonetics and phonology of geminate consonants, Vol. 2 (Oxford Studies in Phonology and Phonetics)*, 66–84. Oxford, UK: Oxford University Press.

Robinson, C., & Schumacker, R. E. (2009). Interaction effects: centering, variance inflation factor, and interpretation issues. *Multiple Linear Regression Viewpoints*, 35.

Rstudio Team. (2018). Rstudio: Integrated Development Environment for R. http://www.rstudio.com

Schiel, F. (1999). Automatic Phonetic Transcription of Non-Prompted Speech. In *Proceedings of the ICPhS*, pp. 607-610.

Selkirk, E. (1996). The prosodic structure of function words. In K. Demuth & J. Morgan (Eds.) *Signal to Syntax: Bootstrapping from speech to grammar in early acquisition*. Lawrence Erlbaum, 187-213.

Seyfarth, S., Garallek, M., Gillingham, G. , Ackermann, F., & Malouf, R.. (2017). Acoustic differences in morphologically-distinct homophones. *Language, Cognition and Neuroscience,* 1-18. doi: 10.1080/23273798.2017.1359634

Smith, R., Baker, R., & Hawkins, S. (2012). Phonetic detail that distinguishes prefixed from pseudo-prefixed words. *Journal of Phonetics, 40,* 689-705. doi: 10.1016/j.wocn.2012.04.002

Sugahara, M. & Turk, A. (2004). Phonetic Reflexes of Morphological Boundaries at a Normal Speech Rate. In B. Bel & I. Marlien (Eds.) *Speech Prosody 2004,* 353-356.

Sugahara, M. & Turk, A. (2009). Durational correlates of English sublexical constituent structure. *Phonology, 26,* 477-524. doi: 10.1017/S0952675709990248

Swanson, L.A., & Leonard, L. B. (1994). Duration of function-word vowels in mother's speech to young children. *Journal of Speech and Hearing Research, 37,* 1394-1405. doi: 10.1044/jshr.3706.1394

Tomaschek, F., Hendrix, P., & Baayen, R. H. (2018). Strategies for addressing collinearity in multivariate linguistic data. *Journal of Phonetics, 71,* 249-267. doi: 10.1016/j.wocn.2018.09.004

Tomaschek, F., Plag, I., Baayen, R. H., & Ernestus, M. (2019). Phonetic effects of morphology and context: Modeling the duration of word-final S in English with naïve discriminative learning. *Journal of Linguistics*, 1-39. doi: 10.1017/S0022226719000203

Torreira, F., & Ernestus, M. (2009). Probabilistic effects on French [t] duration. *INTERSPEECH*, 448-451.

Tremblay, A., & Ransijin, J. (2015). LMERConvenienceFunctions: Model Selection and Post-hoc Analysis for (G)LMER Models [R package]. Retrieved August 2019. https://CRAN.R-project.org/package=LMERConvenienceFunctions

Tucker, B. V., Sims, M., & Baayen, R. H. (2019). Opposing forces on acoustic duration. https://psyarxiv.com/jc97w. doi: 10.31234/osf.io/jc97w

Umeda, N. (1977). Consonant duration in American English. *Journal of the Acoustical Society of America, 61*, 846–858. doi: 10.1121/1.381374

van de Vijver, R., & Baer-Henney, D. (2014). Developing biases. *Frontiers in Psychology, 5*, Article 634. doi: 10.3389/fpsyg.2014.00634

Vitevitch, M.S. & Luce, P.A. (2004). A web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments, and Computers, 36*, 481-487. doi: 10.3758/BF03195594

Wagner, A. R., & Rescorla, R. A. (1972). Inhibition in Pavlovian conditioning: Application of a theory. In R. A. Boakes & M. S. Halliay (Eds.) *Inhibition and learning*. New York: Academic Press, 301-336.

Walsh, L., Hay, J., Bent, D., Grant, L., King, J., Millar, P., Papp, V., & Watson, K. (2013). The UC QuakeBox Project: creation of a community-focused research archive. *New Zealand English Journal, 27*, 20–32.

Walsh, T., & F. Parker. (1983). The duration of morphemic and non-morphemic /s/ in English. *Journal of Phonetics, 11*, 201-206. doi: 10.1016/S0095-4470(19)30816-2

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental duration in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America, 91*, 1707-1717.

Winter, B. (2019). *Statistics for Linguists: An Introduction Using R*. New York: Routledge.

Yao, Y. (2007). Closure duration and VOT of word-initial voiceless plosives in English in spontaneous speech. *UC Berkeley PhonLab Annual Report, 3*, 183-225.

Zimmermann, J. (2016). Morphological status and acoustic realization: Findings from NZE. In Carignanand, C., & Tyler, M. D., (Eds.) *Proceedings of the Sixteenth Australasian International Conference on Speech Science and Technology*, Parramatta, pp. 201-204.

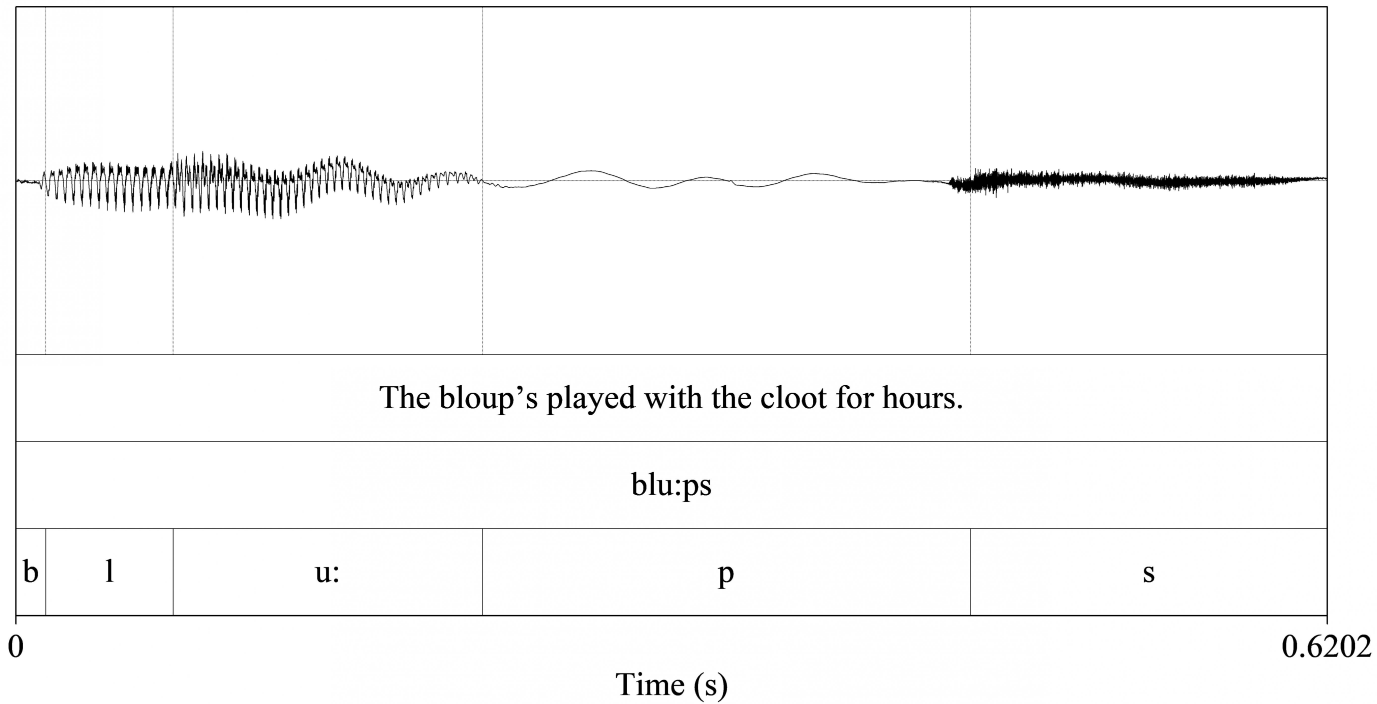
Zvonik, E., & Cummins, F. (2003). The effect of surrounding phrase lengths on pause duration. *Proceedings of Eurospeech*, Geneva, pp. 777-780.

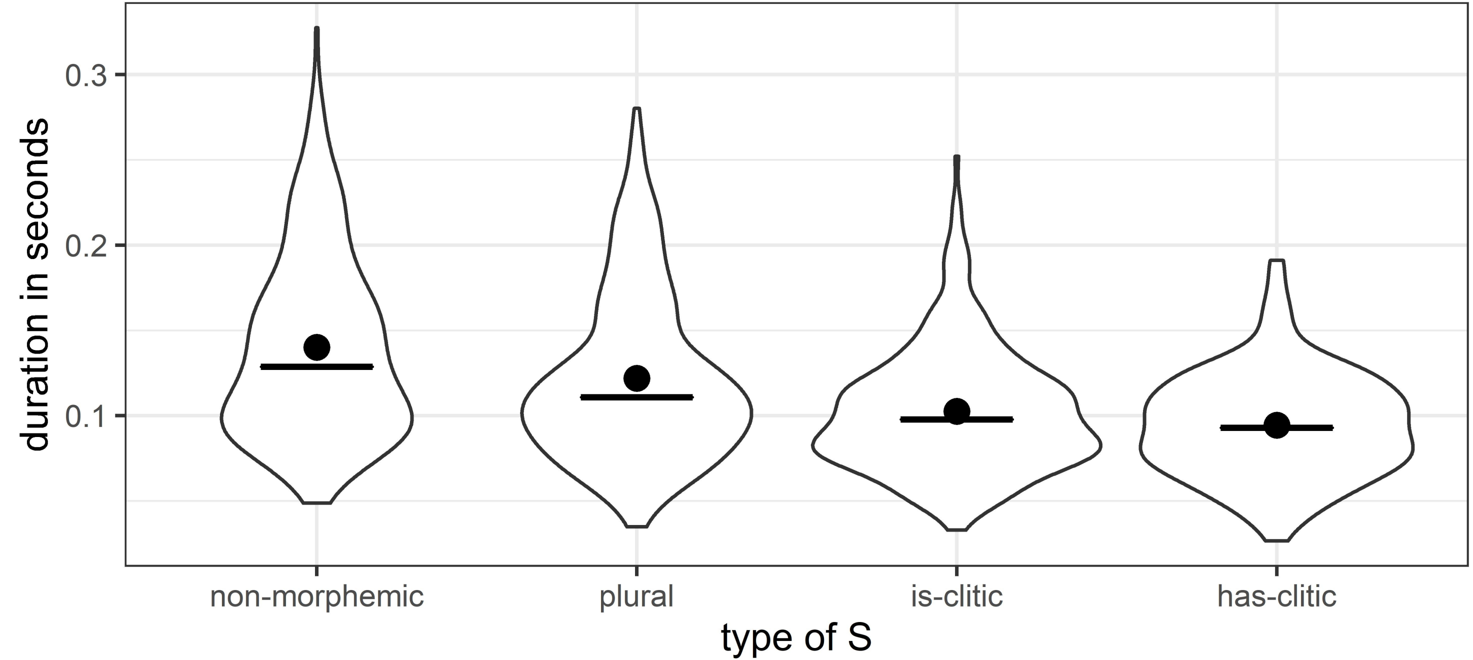
This is a bloup.

And this is a cloot.

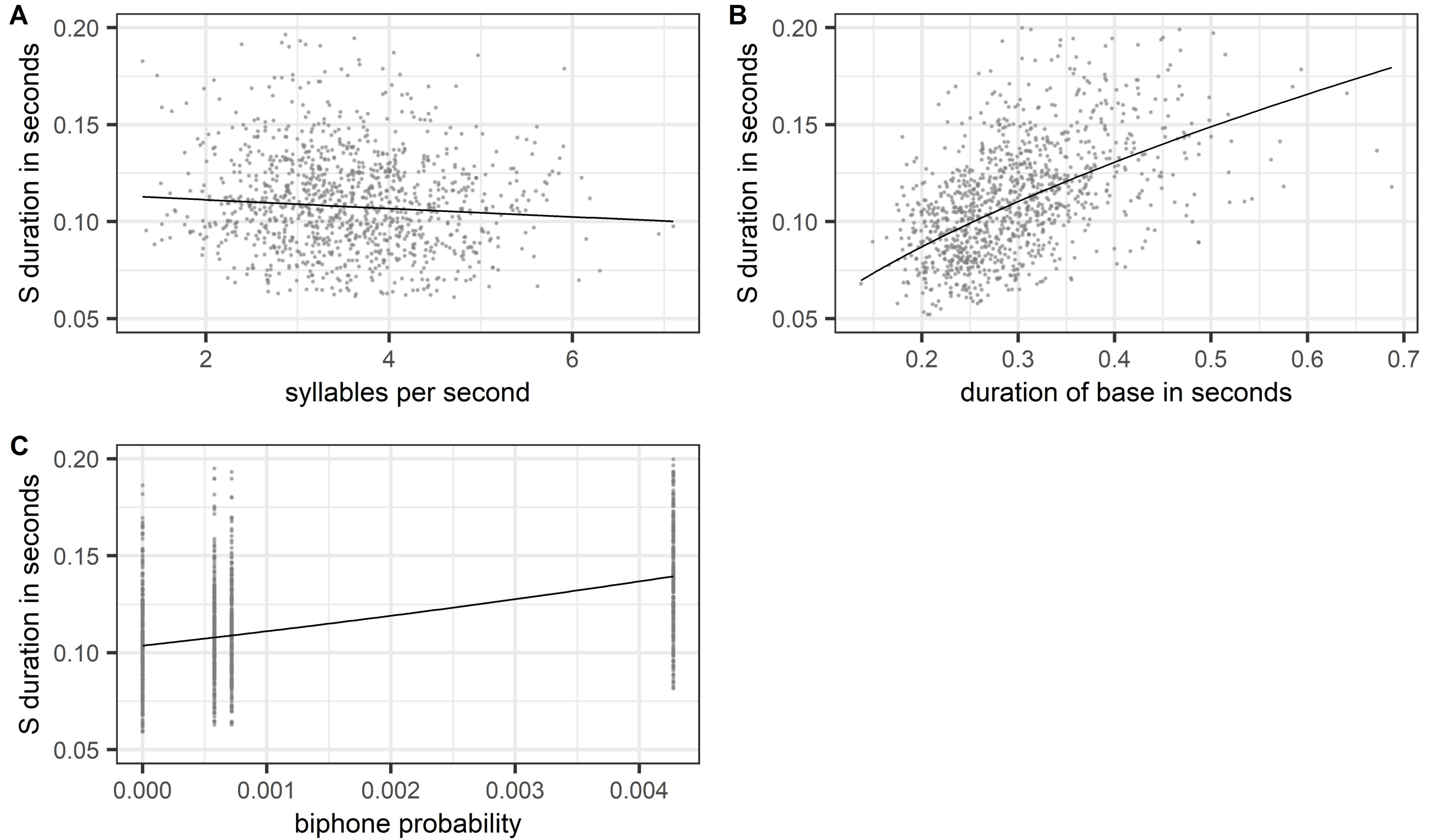
The bloup's played with the cloot for hours.

What's happened for hours?

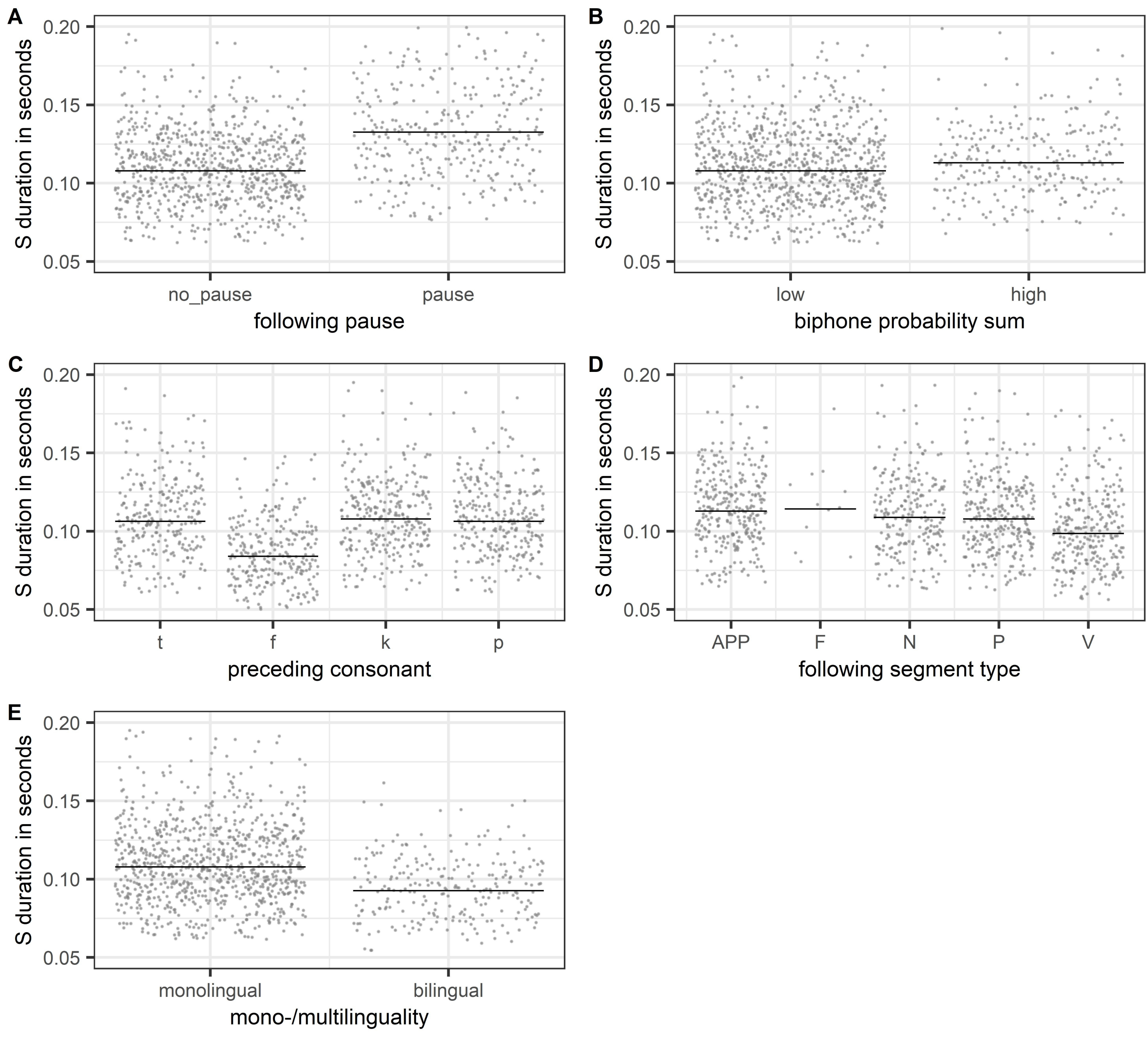

The bloup's played with the cloot for hours.
blu:ps
b
l
u:
p
s
0
0.6202
Time (s)

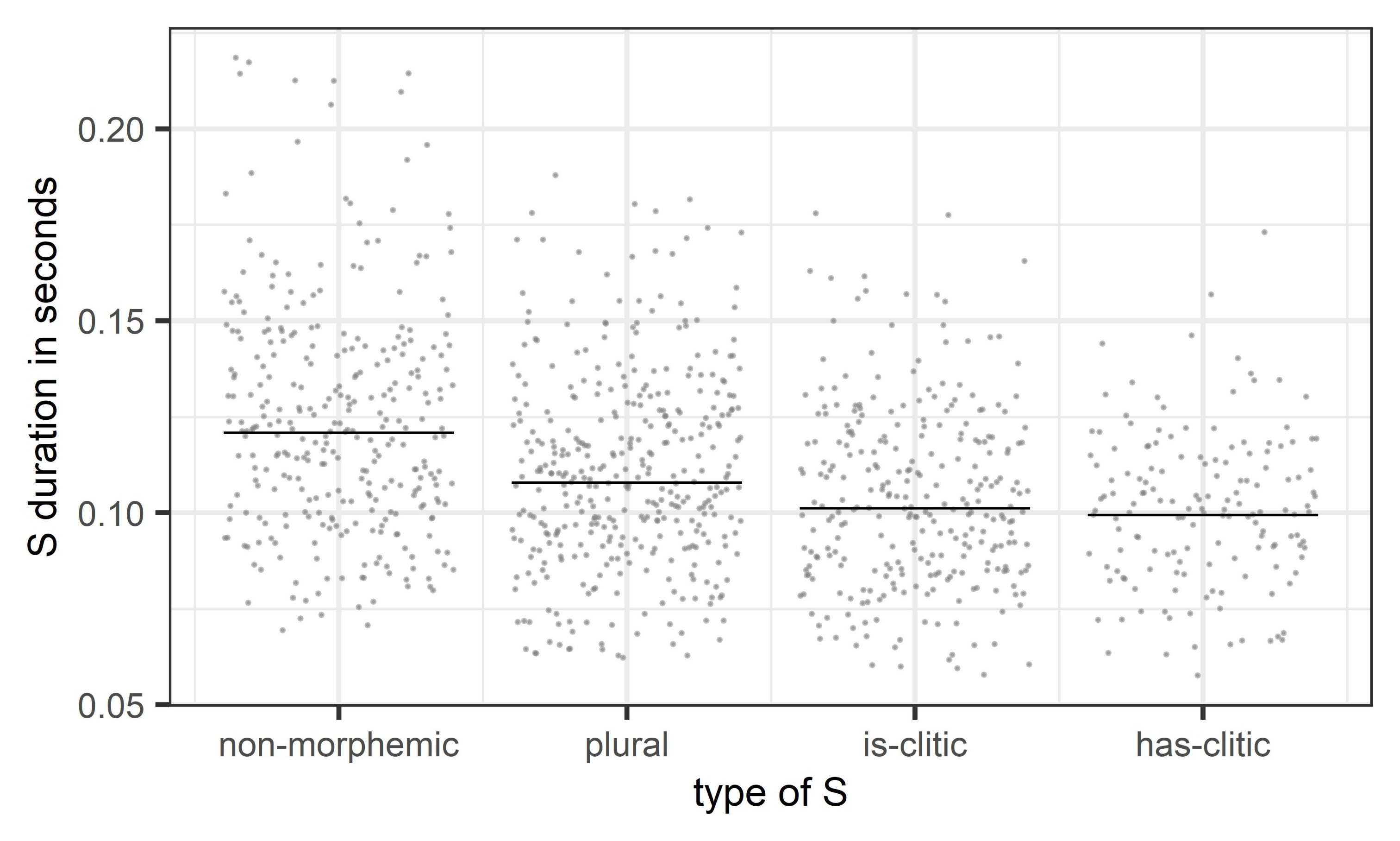


0.3
0.2
0.1
duration in seconds
non-morphemic
plural
is-clitic
has-clitic
type of S



A
S duration in seconds
0.20
0.15
0.10
0.05
2
4
6
syllables per second
B
S duration in seconds
0.20
0.15
0.10
0.05
0.2
0.3
0.4
0.5
0.6
0.7
duration of base in seconds
C
S duration in seconds
0.20
0.15
0.10
0.05
0.000
0.001
0.002
0.003
0.004
biphone probability


A
B
C
D
E
S duration in seconds
0.20
0.15
0.10
0.05
no_pause
pause
following pause
low
high
biphone probability sum
t
f
k
p
preceding consonant
APP
F
N
P
V
following segment type
monolingual
bilingual
mono-/multilinguality



S duration in seconds
0.20
0.15
0.10
0.05
non-morphemic
plural
is-clitic
has-clitic
type of S

Table 1. Overview of durational differences of word-final S found in previous studies.

Study	**Findings**
Zimmermann, 2016; Plag et al., 2017; Tomaschek et al., 2019	non-morphemic > plural > clitics
Walsh & Parker, 1983	plural > non-morphemic
Li et al., 1999	plural > 3rd singular
Seyfarth et al., 2017	plural > non-morphemic
Plag et al., 2019	genitive plural > plural

Table 2. Orthographic representation of the completed stimuli set.

	ɪ	iː	uː	ʌ	aʊ	eɪ
items for morphemic S elicitation	glip glit glik glif	pleep pleet pleek pleef	cloop cloot clook cloof	prup prut pruk pruf	bloup blout blouk blouf	glaip glait glaik glaif
items for non-morphemic S elicitation	glips glits gliks glifs	pleeps pleets pleeks pleefs	cloops cloots clooks cloofs	prups pruts pruks prufs	bloups blouts blouks bloufs	glaips glaits glaiks glaifs

Table 3. Number and type of S elicitations per speaker.

non-morphemic S	plural S	is-clitic S	has-clitic S	total number of trials per speaker
12	12	12	12	48

Table 4. Summary of the dependent variable and numerical predictors in the final data set.

Dependent variable	Mean	St. Dev.	Min	Max
sDurLog	0.002	0.388	- 1.201	1.098
Numerical predictors	**Mean**	**St. Dev.**	**Min**	**Max**
speakingRate	-0.000	0.899	2.250	3.540
baseDurLog	0.072	0.194	0.000	3.559
pauseDur	0.072	0.193	0.000	3.559
neighbourhoodFrequency	27.345	84.645	0.000	412.027
biphoneProbSum	0.013	0.007	0.005	0.031
biphoneProb	0.001	0.002	0.000	0.004
age	28.740	9.743	19.000	58.000

Table 5. Summary of categorical predictors and the dependent variable in the final data set.

Categorical predictors	Levels							
ITEM	48							
TRANSCRIPTION	67							
NEIGHBOURHOODDENSITY	0: 419	1: 238	2: 165	3:107	4: 14	5: 114	6: 32	7: 30
PAUSEBIN	no: 777	yes: 342						
BIPHONEPROBSUMBIN	low: 856	high: 263						
LIST	24							
SLIDENUMBER	48							
PREC	f: 273	k: 292	p: 281	t: 273				
FOLSEG	18							
FOLTYPE	APP: 299	F: 12	N: 230	P: 300	V: 278			
SPEAKER	40							
GENDER	2							
LOCATION	London: 636	elsewhere: 483						
MONOMULTILINGUAL	monolingual: 871	multilingual: 248						
Explanatory variable	**Levels**							
TYPEOFS	nm: 308	pl: 373	is: 284	has: 154				

Table 6. *p*-values of fixed effects in the final model, fitted to the log-transformed durations of S.

	Sum Sq	Mean Sq	NumDF	DenDF	F.value	Pr (> F)
typeOfS	5.312	1.771	3	1089.66	33.338	0.000
speakingRate	0.230	0.230	1	1117.09	4.324	0.038
baseDurLog	9.466	9.466	1	1079.58	178.220	0.000
pauseBin	6.970	6.970	1	1110.28	131.235	0.000
biphoneProbSumBin	0.398	0.398	1	1082.26	7.492	0.006
biphoneProb	0.338	0.338	1	1079.25	6.360	0.012
preC	0.623	0.208	3	1080.29	3.910	0.009
folType	2.677	0.669	4	1081.55	12.598	0.000
monoMultilingual	0.345	0.345	1	37.37	6.498	0.015

Table 7. Fixed-effect coefficients and p-values as computed by the final model (mixed-effects model fitted to the log-transformed and centred durations of S).

	Estimate	Std. Error	df	*t*-value	Pr (> |t|)
(Intercept)	0.096	0.034	98.81	2.814	0.000
TYPEOFSpl	-0.114	0.019	1094.00	-6.062	0.000
TYPEOFSis	-0.178	0.020	1096.00	-8.839	0.000
TYPEOFShas	-0.196	0.024	1091.00	-8.14	0.000
SPEAKINGRATE	-0.021	0.010	1117.00	-2.079	0.038
BASEDURLOG	0.586	0.044	1080.00	13.35	0.000
PAUSEBINpause	0.206	0.018	1110.00	11.456	0.000
BIPHONEPROBSUMBINhigh	0.047	0.017	1082.00	2.737	0.006
BIPHONEPROB	0.069	27.53	1079.00	2.522	0.012
PRECf	0.061	0.020	1081.00	-3.044	0.003
PRECk	0.055	0.020	1082.00	-0.303	0.006
PRECp	0.050	0.020	1079.00	2.522	0.012
FOLTYPEF	0.012	0.070	1084.00	0.171	0.864
FOLTYPEN	-0.036	0.021	1079.00	-1.764	0.078
FOLTYPEP	-0.045	0.019	1080.00	-2.384	0.017
FOLTYPEV	-0.136	0.020	1082.00	-6.85	0.000
MONOMULTILINGUALmultilingual	-0.152	0.059	37.37	-2.549	0.015

Table 8. Multiple comparisons of means of duration of S (Tukey contrasts). Significant codes: '***' $p < 0.001$, '**' $p < 0.01$, '*' $p < 0.05$.

	Estimate	Std. Error	*z*-value	Pr (> |z|)	
pl – nm	-0.114	0.019	-6.062	< 0.001	***
is – nm	-0.018	0.020	-8.839	< 0.001	***
has – nm	-0.196	0.024	-8.140	< 0.001	***
is – pl	-0.064	0.019	-3.294	0.005	**
has – pl	-0.082	0.023	-3.503	0.003	**
has – is	-0.018	0.023	-0.766	0.868	

Table 9. Significant contrasts in duration between different types of S. Significant codes: '***' $p < 0.001$, '**' $p < 0.01$, '*' $p < 0.05$.

	nm	pl	*is*	*has*
nm	n.a.	***	***	***
pl		n.a.	**	**
is			n.a.	
has				n.a.

Table 10. S durations as estimated by the final model using non-centred data. All values are back-transformed to seconds. Values given are estimated for items without following pause, high biphone sum probability, monolingual speakers, and across all preceding and following segment types.

TYPEOFS	Mean
non-morphemic	0.224
plural	0.200
is-clitic	0.187
has-clitic	0.184

Table 11. p-values of fixed effects in the final model, fitted to the relative durations of S.

	Sum Sq	Mean Sq	NumDF	DenDF	F.value	Pr (> F)
TYPEOFS	0.161	0.054	3	1070.68	25.510	0.000
PAUSEBIN	0.186	0.186	1	1101.26	88.518	0.000
BIPHONEPROBSUMBIN	0.015	0.015	1	36.32	6.917	0.012
FOLTYPE	0.071	0.018	4	1063.31	8.389	0.000
MONOMULTILINGUAL	0.010	0.010	1	37.81	4.561	0.039

Table 12. Fixed-effect coefficients and p-values as computed by the final model (mixed-effects model fitted to the relative durations of S).

	Estimate	Std. Error	df	*t*-value	Pr (> |t|)
(Intercept)	0.299	0.007	89.73	45.827	0.000
TYPEOFSpl	-0.019	0.004	1085.00	-5.157	0.000
TYPEOFSis	-0.031	0.004	1070.00	-7.651	0.000
TYPEOFShas	-0.035	0.005	1067.00	-7.260	0.000
PAUSEBINpause	0.033	0.004	1101.00	9.408	0.000
BIPHONEPROBSUMBINhigh	0.013	0.005	36.32	2.630	0.012
FOLTYPEF	0.001	0.014	1068.00	0.086	0.931
FOLTYPEN	-0.006	0.004	1061.00	-1.409	0.159
FOLTYPEP	-0.007	0.004	1056.00	-1.708	0.088
FOLTYPEV	-0.022	0.004	1063.00	-5.568	0.000
MONOMULTILINGUALmultilingual	-0.024	0.011	37.81	-2.136	0.039

Table 13. Multiple comparisons of means of relative duration of S (Tukey contrasts). Significant codes: '***' $p < 0.001$, '**' $p < 0.01$, '*' $p < 0.05$.

	Estimate	Std. Error	*z*-value	Pr (> |z|)	
pl – nm	-0.019	0.004	-5.157	< 0.001	***
is – nm	-0.031	0.004	-7.651	< 0.001	***
has – nm	-0.035	0.005	-7.260	< 0.001	***
is – pl	-0.011	0.004	-2.936	0.017	*
has – pl	-0.015	0.005	-3.300	0.005	**
has – is	-0.004	0.005	-0.854	0.827	

Appendix A

Contexts and questions used in the production task sorted by onset segment of the verb following the word-final S, and the type of word-final S. The pseudowords cloot/cloots and glaik/glaiks are used as examples.

1. Approximant onset verbs

1a. write

non-morphemic

Context: The cloots writes a letter to the glaiks every month.
Question: What happens every month?

plural

Context: Last week, the cloots wrote a letter to their mother.
Question: What happened last week?

is-clitic

Context: The cloot's writing a letter to the glaik.
Question: What's happening?

has-clitic

Context: The cloot's written a love letter to the glaik.
Question: What's happened?

1b. listen

non-morphemic

Context: Every day, the cloots listens to the glaik’s singing.
Question: What happens every day?

plural

Context: Last week, the cloots listened to each other's songs.
Question: What happened last week?

is-clitic

Context: The cloot's listening to the glaik sing.
Question: What's happening?

has-clitic

Context: The glaik's a famous singer. The cloot's listened to all of his songs.
Question: What's happened?

1c. watch

non-morphemic

Context: Every night, the cloots watches the glaiks' TV series.
Question: What happens every night?

plural

Context: Yesterday, the cloots watched TV together.
Question: What happened yesterday?

is-clitic

Context: The cloot's watching the glaik play football.
Question: What's happening?

has-clitic

Context: The glaik's a famous football player. The cloot's his biggest fan. He’s watched all of the glaik's matches.
Question: What's happened?

2. Nasal onset verbs

2a. move

non-morphemic

Context: They’re good friends and want to live close to each other. Therefore, the cloots moves into a new home.
Question: What happens?

plural

Context: Last year, the cloots moved into a new home.
Question: What happened last year?

is-clitic

Context: The cloot's moving in with the glaik.
Question: What's happening?

has-clitic

Context: The cloot's moved in with the glaik.
Question: What's happened?

2b. meet

non-morphemic

Context: Every Saturday, the cloots meets the glaiks for a drink.
Question: What happens every Saturday?

plural

Context: Last week, the cloots met for a drink.
Question: What happened last week?

is-clitic

Context: Tonight, the cloot's meeting the glaik for a drink.
Question: What's happening tonight?

has-clitic

Context: One year ago, the cloot's met the glaik for the first time.
Question: What's happened one year ago?

2c. knit

non-morphemic

Context: Every night, the cloots knits a blanket for the glaiks.
Question: What happens every night?

plural

Context: Last week, the cloots knitted a blanket together.
Question: What happened last week?

is-clitic

Context: The cloot's knitting a hat for the glaik's birthday.
Question: What's happening?

has-clitic

Context: The cloot's knitted ten scarfs for the glaik last winter.
Question: What's happened last winter?

3. Plosive onset verbs

3a. play

non-morphemic

Context: Every day, the cloots plays with the glaiks.

Question: What happens every day?

plural

Context: Last week, the cloots played a game.

Question: What happened last week?

is-clitic

Context: The cloot's playing with the glaik.

Question: What’s happening?

has-clitic

Context: The cloot's played with the glaik for hours.

Question: What’s happened for hours?

3b. call

non-morphemic

Context: Every night, the cloots calls the glaiks for a nice chat.

Question: What happens every night?

plural

Context: Yesterday, the cloots called each other to talk about their day.

Question: What happened yesterday?

is-clitic

Context: The cloot's calling the glaik to talk about their evening plans.

Question: What’s happening?

has-clitic

Context: The cloot's calling the glaik, but the glaik does not answer the phone. The cloot's called the glaik several times by now.

Question: What’s happened several times now?

3c. cook

non-morphemic

Context: Every Sunday, the cloots cooks lunch for the glaiks.

Question: What happens every Sunday?

plural

Context: Every Friday, the cloots cook dinner together.

Question: What happens every Friday?

is-clitic

Context: The cloot's cooking dinner for the glaik.

Question: What's happening?

has-clitic

Context: The cloot's a great cook. The cloot's cooked lunch for the glaik for many years.

Question: What's happened for many years?

4. Vowel onset verbs

4a. ask

non-morphemic

Context: Every Friday, the cloots asks the glaiks about his weekend.
Question: What happens every Friday night?

plural

Context: Last Friday, the cloots asked each other about their weekend.
Question: What happened last Friday?

is-clitic

Context: The cloot's asking the glaik about his weekend.
Question: What’s happening?

has-clitic

Context: They just met. The cloot's a curious thing. He’s asked the glaik many questions in the past couple hours.
Question: What’s happened in the past couple hours?

4b. eat

non-morphemic

Context: The cloots eats breakfast with the glaiks every day.
Question: What happens every day?

plural

Context: Two days ago, the cloots ate their lunch together.
Question: What happened two days ago?

is-clitic

Context: The cloot’s eating cake with the glaik.
Question: What’s happening?

has-clitic

Context: They are having lunch together. The cloot's really hungry. He’s eaten the glaik’s lunch as well.
Question: What’s happened?

4c. attend

non-morphemic

Context: Tonight, the cloots attends the glaiks' party.
Question: What happens tonight?

plural

Context: Yesterday, the cloots attended a ball together.
Question: What happened yesterday?

is-clitic

Context: Tomorrow, the cloot's attending the glaik's party.
Question: What happens tomorrow?

has-clitic

Context: They're big music fans. The cloot's attended concerts with the glaik many times.
Question: What's happened many times?

Appendix B

Practice material used in the production task. The pseudowords lope/lopes and feap/feaps were used in the practice trials.

non-morphemic

Context: The feaps is on holiday, therefore the lopes misses him a lot.

Question: What’s happening?

plural

Context: Two weeks ago, the feaps convinced their best friend to join their sports team.

Question: What happened two weeks ago?

is-clitic

Context: The lope’s late. He’s missing his appointment with the feap.

Question: What’s happening?

has-clitic

Context: The feap’s convinced the lope many times to play a game with him.

Question: What’s happened in the past couple hours?