AJELS can be found on EBSCOhost databases
https://doi.org/10.59960/12.1.a4

Subphonemic detail and foreign language learning: Word-final sibilants in Korean learners of L2 English and L3 German

Eva Maria Luef

Received: November 6, 2023
Reviewed: January 14, 2024
Revised: February 3, 2024
Accepted: March 4, 2024

Corresponding author
Eva Maria Luef
University of Hamburg, Institute of English and American Studies, Germany; Charles University, Department of English and ELT Methodology, Czech Republic
evamaria.luef@ff.cuni.cz

Additional information is available at the end of the article

Abstract

A variety of studies have documented durational differences in homophonous word-final phonemic and morphemic alveolar sibilants in English ([s,z]). The present study investigates Korean learners of English as a foreign language (L2), whose first language does not show acoustic release of word-final sibilants, in their production of English phonemic and morphemic [s] and [z] in word-final position. In addition, Korean learners' production of word-final [s] in German as a third language (L3) are analyzed and compared to how first language speakers of German acoustically realize different functional variants of homophonous alveolar sibilants. Results show that Korean learners are able to incorporate subphonemic detail of word-final [s] and [z] that they implicitly learn in their first (English) and second (German) foreign language. The findings support the view that Korean listeners/ speakers are attentive to durational variation in word-final sibilants in foreign languages. The findings can contribute to the growing field of subphonemic studies in foreign language learning.

Keywords: sibilants; morpho-phonetics; Korean; English as a foreign language; German as a foreign language; DaF; subphonemic detail

1. Introduction

Morphological processes are generally assumed to produce different types of word-internal boundaries, which can have various phonological consequences (Chomsky & Halle, 1968; Inkelas, 2014). Numerous recent studies have demonstrated that phonetically gradient processes may also play a role, with fine phonetic detail (but no change in phonological category) being conditioned upon morphological boundaries (Losiewicz, 1995; Plag et al., 2017; Song et al., 2013; Tomaschek et al., 2019; Walsh & Parker, 1983; Zimmermann, 2016). In English, subphonemic features of sibilants [s] and [z] have been reported to be different, with homophonous word-final sibilants being phonetically varied according to the morphological boundary strength that precedes them. First language (L1) speakers of English distinguish word-final sibilant phonemes and morphemes, such as in *lapse–laps*, in terms of their durations (Losiewicz, 1995; Walsh & Parker, 1983). While both types of /s/ in the example are technically phonemic, the term is used here (and in related research) to refer to /s/ phonemes that do not have a morphemic function (as in *lapse*) as opposed to /s/ phonemes that have morphemic functions (as in plural *laps*; in the following, all variants of alveolar sibilants will be referred to as |s|). Plag et al. (2017) reported phonemic |s| to be longest, and different |s| morphemes – corresponding to different morphological boundary strengths – varying by duration, with clitic |s| being shortest. Tomaschek et al. (2019) and Zimmermann (2016, for New Zealand English) found corresponding results. Different results have been reported in studies by Walsh and Parker (1983) and Song and colleagues (2013), who found shorter durations for phonemic than for morphemic word-final segments. While these contradictory results do not bode well for the assumption that morphemic and phonemic word-final |s| are systematically varied by length, the inconclusive findings may be a result of experimental and/ or statistical differences (see arguments in Plag et al., 2017; Tomaschek et al., 2019).

In general, findings on durational differences between |s| phonemes and morphemes run counter to traditional notions of speech production where phonetic and morphological processing are strictly separated (e.g., Chomsky & Halle, 1968; Levelt et al., 1999). This theoretical conundrum has been difficult to explain. Tomaschek et al. (2019) provide a comprehensive explanation for the durational |s| effects by describing them in the context of associative learning between the articulatory gesture of |s| and the semantics of the morphological function. They suggest that the durational differences are influenced by the degree to which phonological and collocational properties of words are able to discriminate between the different functions represented by the word-final |s| (see Tucker et al., 2019, for a similar argument). According to Tomaschek et al. (2019), the probability of word-final |s| to serve as a descriminative cue is related to its duration; specifically, if more uncertainty about the intended function is created, its duration decreases. In their own words, “energy is not invested in a signal that creates confusion instead of clarity” (p. 154). This is related to the concept of ‘functional load’, where phonological contrast is lost when functional load of an item decreases (Wedel et al., 2013).

A question that has remained unexplored to date is how this subphonemic variation in word-final sibilants is acquired by foreign language learners of English. Especially languages where word-final |s| release is absent, such as Korean, can provide a window into the integration of subphonemic detail in learned languages. In Korean, word-final phonemic |s| (i.e., <ㅅ>) has no acoustic surface realization, as all word-final coronal obstruents in the language are neutralized to unreleased [t̚] (Kim & Jongman, 1996). Some insights into Korean learners’ attention to subphonemic detail of English can be gleaned by studying loanword phonology. Korean distinguishes the two sibilant fricatives <ㅅ> and <ㅆ>, which differ by duration and may appear as released surface forms in word-initial and -internal positions. For instance, the words *buy* 사다 [sa̠da̠] and *inexpensive* 싸다 [s͈a̠da̠] represent two examples of the different sibilants. According to the standard view of Korean phonology, <ㅅ> [s] is classified as phonologically lax (Cho & Whitman, 2019) and characterized by a duration of under 140 ms (Kim & Curtis, 2002). The phonologically tense variant [s̤] or [s*]/ <ㅆ> is characterized by longer than 140 ms duration as well as its effect on succeeding vowels (even though there is no evidence for production differences, listeners associate higher f0 with /s*/, see Lee & Katz, 2016). The English durational frication variation in relation to the phonological environment of an obstruent – |s| being shorter in consonant clusters, with about 123–145 ms and longer as singletons, with about 145–162 ms (see Haggard, 1972) – seems to register with Korean listeners, as they incorporate this information into the phonetics of their English loanwords (Ahn & Iverson, 2004; Iverson & Lee, 2006; Kim, 1999). Loan phonology suggests that English singleton |s| is borrowed with tense [s*] (e.g., *sale* as [s*ɛil], *pass* as [pʰæs*ɨ]) while |s| in consonant clusters is borrowed with lax [s] in Korean preconsonantally (e.g., *star* as [sɨtʰa] and *fast* as [pʰɛsɨtʰɨ]). Cross-linguistic perception studies involving naïve listeners demonstrate that Korean tense [s*] can be phonetically similar to English [s] across different segmental position (Cheon & Anderson, 2008; Schmidt, 1996, 2007). Even though Korean phonology has two sibilant types which can both correspond to English /s/, they do not appear in word-final position, and Korean learners of English have to fill an acoustic gap when acquiring English phonetics.

Phonetic learning in a second language (“L2”) can have various roots. The complex relationship between perception in production-related speech processes in L2 has been debated (e.g., Baese-Berk & Samuel, 2016; Wang et al., 2003), and it is unclear whether L2 speech is characterized by more (Wade et al., 2007) or less (Vaughn et al., 2019) phonetic variation. Any phonetic variability encountered by L2 learners in the L1 signal has to be categorized as either representing allophonic variation or meaningful for morpho-phonetic contrast (as in the case of word-final |s|). In the context of the present study the question is whether Korean learners of English (implicitly) recognize the fine-grained morpho-phonetic variation associated with word-final |s| in English and incorporate this information into their own speech patterns. If Korean learners create new phonetic categories for different types of word-final |s|, it could indicate that they are able to process subtle phonetic detail (which is likely below the level of conscious perception in both L1 and L2 speakers) and adapt their speech output in a second language accordingly. Loan phonology hints in this direction.

Not much is known about durational variation in homophonous phonemic and morphemic word-final |s| in languages other than English or about second language learners. The present study was designed to investigate word-final |s| durations in Korean learners of English as a second language and German as a third language (“L3”). Specifically, the durational |s| variations measured in L1 speaker groups of the languages will be compared to those measured in Korean learners. The following two research questions are posed:

(1) Do Korean learners of L2 English show similar durational variation in their word-final released |s| as L1 English speakers?

(2) Do Korean learners of L3 German show similar durational variation in their word-final released |s| as L1 German speakers?

As opposed to English, Koreans are not regularly exposed to German loanwords or the German language in general, resulting in less phonetic experience with the language. While German also displays different types of morphemic word-final |s| (in addition to phonemic), they differ from the English ones in terms of function and frequency. German word-final |s| morphemes include plural |s|, singular and plural genitives, clitics of *das* 'the' and *es* 'it' (e.g., *fürs* [fyːɐs] 'for the', *mach's* [maxs] 'do it'), inflections (*kleines* [klaɪnəs] 'small', neutr.nom. sg. or neutr.acc.sg), and derivations (*morgens* [mɔɐgəns] 'in the morning'). The insights into word-final |s| acoustics in German will not only help contextualize the effect more broadly within the Germanic language family, but the comparison with Korean learners of German will provide information relating to L2 and L3 phonetic learning in L1 Korean users.

2. Method

2.1 Participant, materials, and procedures

First language speakers of Korean (*N* = 22, 17 female/ 5 male, mean age = 22.7 years) were recruited in Seoul (Republic of Korea) to participate in a sentence-reading task in their foreign languages English and German between November and December 2018. Their dialectal background can be described as greater area Seoul. Participants gave written consent and were paid 10,000 Korean Won for their participation in the phonetic experiment. They self-reported their proficiency levels in English (B2 and higher) and in German (A2 and higher), according to the Common European Framework of References for Languages, an international standard for describing language ability (Council of Europe, 2018). A2 refers to the advanced beginners' stage, where students have mastered basic foreign-language skills. B2 is an advanced intermediate level, with students being able to communicate easily and spontaneously in a variety of situations (more information can be found on the website of the Council of Europe: https://www.coe.int/en/web/common-european-framework-reference-languages). Participants had learned English and German primarily in Korean classroom settings, and had spent less than one year in an English- or German-speaking country. Average age of onset was grade 3 (age: 8 years) for English and grade 10 (age: 15 years) for German. American English is the predominant learner input variety in South Korean English classrooms (Ahn, 2011), while central and northern German is the main input variety for German.

Each participant was asked to read 90 English sentences (mean sentence length: 6.3 ± 2.1 words) and 90 German sentences (mean sentence length: 6.2 ± 1.9 words) once at a preferred speed. 55% of participants chose to start the procedure by reading the English sentence list first. Carrier sentences containing words ending in different types of |s| (i.e., phonemic, morphemic) were interspersed with filler sentences. Sentences were phonologically and semantically unrelated. After each sentence, participants rated syntactical difficulty by stating the Sino-Korean terms for 'easy', 'medium' and 'difficult' (*ha*/ 하, *jung*/ 중, *sang*/ 상). This procedure was meant to minimize priming effects (Kinoshita et al., 2018). Mispronounced items (e.g., [ʤɪl] instead of [gɪl]) were removed from the sample.

First language speakers of English and German took part in the same experiments. Seven L1 speakers of American English (6 female and 1 male university students, mean age = 26 years) were recruited between October and December 2022 to read the English stimuli. Ten L1 speakers of Austrian German (female university students, mean age = 24 years) were recruited in 2018 to read the German stimuli. All participants gave their written consent and were paid 10 Euros for their contribution to a language corpus. The identical set of target words were used for the English analyses (L1, L2), and for the German analyses (L1, L3). The English and German carrier sentences can be found in Supplementary Table S1.

Speech of L1 English speakers, L1 German speakers, and Korean foreign-language learners of L2 English and L3 German was recorded with a Sennheiser ME67 microphone attached to a ZoomH4n digital audio recorder and sampled at a rate of 44.1 kHz with 16-bit depth. Target words containing word-final |s| were manually identified and cut to be saved as separate .wav files. The .wav files were then loaded into *Penn Phonetics Lab Forced Aligner for English* (Yuan & Liberman, 2009), which produced a Textgrid file for use in *Praat* (Boersma & Weenink, 2019). As recommended by the programers, post-alignment editing was done in Praat (version 6.0.46) in order to remove faulty alignments. 7% of the manually corrected files were randomly selected (using the RANDOM function in Excel on the file list and selecting the top 7%) to be checked by a second rater and interrater reliability (Pearson's r, root mean square error/ RMSE) were calculated on overall word durations (as these were identified to contain errors). Excellent annotation agreement was established (r = .99, RMSE = .023). Lastly, overall word durations and durations of the word-final |s| were extracted with a script from the Textgrid files in Praat (see Figure 1).

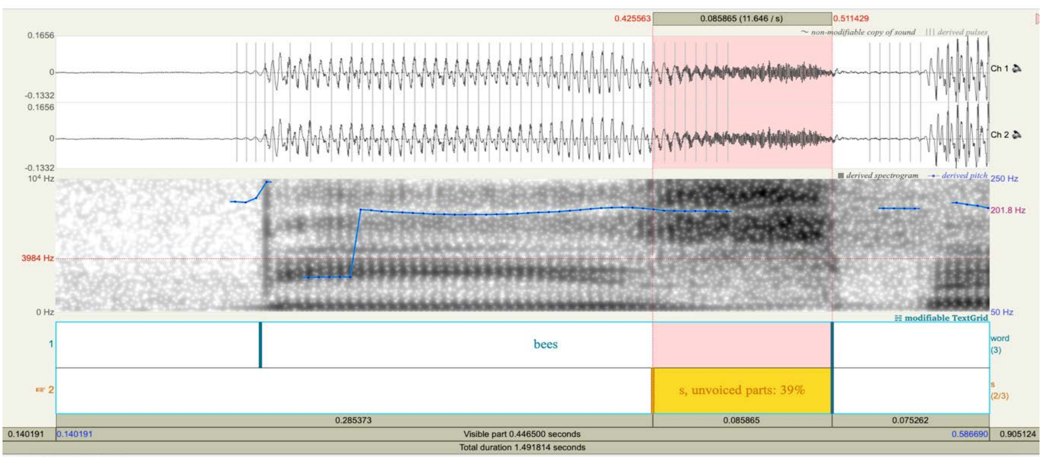

Figure 1. ***Example of a Praat annotation, indicating word duration on tier 1 and /s/ duration (with voiced portion) on tier 2.***

The majority of word-initial phonemes were plosive consonants and the start of the respective target words was marked at the burst of the stop (Abramson & Whalen, 2017). The acoustic start of approximants, liquids, vowels, fricatives, and nasals were determined by the start of cycling on the waveform, with the nearest zero crossing being defined as the beginning. The end point of words (i.e., the word-final |s|) was marked when the friction had ceased (i.e., the nearest zero crossing) and the sound had completely faded.

The beginning of word-final |s| was also read from the waveform and the beginning of the aperiodic noise in the high frequency range was defined as the initiation of |s|. This was acoustically clear in cases when vowels, liquids, and nasals preceded the word-final |s|. Word-final plosive+|s| combinations proved difficult to separate at times. Generally, Korean learners of English and German, as well as L1 English and L1 German speakers, showed visible acoustic onset of the |s|, however, in about 20% of the cases no clear acoustic distinction between the plosive and the following |s| was visible. In addition, Korean learners would sometimes drop plosives and merge two |s| in word-final consonant clusters involving plosives and |s| and pronounce words such as "tests" as [tɛss]. All cases of unclear |s| separation or consonant deletion before word-final |s| were removed from the sample.

Spectrograms of the Korean learners were inspected for the presence of an epenthetic vowel immeditately following the word-final |s|. This occurred twice in the L2 English sample (one case each of "disease" and "geese") and never in the L3 German sample. Due to this negligible number of occurrences, the cases with epenthetic vowels were removed from the sample and presence or absence of epenthetic vowel was not included as a variable. Absence of |s| was recorded in 9 cases of L2 English (e.g., "way" instead of "ways") and in 6 cases of L3 German (e.g., "morgen" instead of "morgens"). These examples were removed from the data sets before further analysis.

The absolute duration of |s| in milli-seconds served as the dependent variable. Following Plag et al. (2017), the absolute values were Box-Cox transformed (Box & Cox, 1964). Table 1 presents the optimal lambda values per data set as calculated with the function *lambda* of the R package 'MASS' (Venables & Ripley, 2002).

Table 1
Lambda values used for the Box-Cox transformation of the dependent variables in the L1, L2, and L3 data sets.

Data set	Lambda used for Box-Cox transformation according to $y'_\lambda = (y^\lambda-1)/\lambda$
L1 English	-0.3434343
L2 English	-0.1010101
L1 German	-0.06060606
L3 German	-0.3838384

2.2 Confounding variables in L1, L2, L3

Duration effects in morphologically complex words can have a variety of causes (Hanique et al., 2013; Plag et al., 2017; Pluymaekers et al., 2010), and relevant confounding variables for word-final |s| were taken into account when designing the present study.

First, the functional categories of word-final |s| were determined for the languages. In English, |s| were divided into (a) phonemes, (b) plural morphemes, (c) 3rd person morphemes, and (d) *is*, the latter being very frequent and potentially skewing classification of phonemes. In German, the following |s| types were recorded: (a) phonemes, (b) plural morphemes, (c) inflections (non-genitive), (d) derivations, (e) genitives, (f) clitics, and (g) *es* ‘it’ (which is very frequent in German).

Phrase- or utterance-final lengthening describes the phonetic mechanism responsible for longer segment durations at the end of words, phrases, or sentences (Berkovits, 1993), and is a well-documented fact for Germanic languages (Delattre, 1966). Even though it has been linked to stress-timed isochronous patterns (Snow, 1994), utterance-final lengthening can also be found in Korean (Jeon & Nolan, 2013). Final |s| occurring in mid-sentence position (“mid”) were distinguished from those occurring in sentence-final (“final”) position in English and German.

The phonological environments of word-final |s| was also specified. The segments immediately preceding or succeeding it have important implications for its duration, with |s| in consonant clusters showing shorter durations (approximately 40% in clusters with plosives, and 15% in clusters with sonorants, see Klatt, 1974). Segments immediately following a word can also have a durational effect on the final segment of that word (Klatt, 1976; Umeda, 1977). In both English and German, segments immediately preceding the word-final |s| were classified into (a) plosives (e.g., *gigs, Dachs* [daks] ‘badger’), (b) nasals (e.g., *cans, meistens* [maɪstns] ‘mostly’), and (c) vowels (e.g., *kiss, Haus* [haus] ‘house’). In the English sample, the options also included (d) fricatives (e.g., *tariffs*), and (e) liquids (e.g., *animals*). In addition, the segment immediately following the word-final |s| were coded (for English: approximant, fricative, liquid, nasal, plosive, vowel, none; for German: fricative, plosive, vowel, none).

Another pertinent variable for word-final |s| duration is voicing, with voiced fricatives being shorter on average (Jongman et al., 2000; Plag et al., 2017). In English, unvoiced [s] and voiced [z] are recognized in word-final position, while in German only unvoiced [s] can appear in that position due to final devoicing. In the present study, phonetic voicing of all |s| were measured using the voice-report algorithm in Praat which bases its measurements on glottal pulsing. This yielded a continuum of voicing possibilities ranging from 0 to 100.

Additionally, graphemic representations of word-final |s| vary in English and in German. This may have an effect on how Korean learners of the languages produce the |s| segments, especially in the light of different writing systems between the first and the foreign languages (Mathieu, 2015). English graphemes for each word-final |s| were classified as (a) <s>, (b) <ss>, and (c) <c> depending on the |s|-inducing grapheme and regardless of whether an "e" followed the graphemes (e.g., *ice*). German graphemes were classified as (a) <s>, (b) <ss>, (c) <ß> (referred to as 'Eszett' or "sz"), with the graphemes always in word-final position.

Lexical frequency rates of target words have a known impact on segment durations, with more frequent words being characterized by more phonetic reduction across segments (Brown, 2009; Gahl, 2008). Lexical frequency rates of the L1 English and L1 German target words were determined with the help of Clearpond, which is based on the SUBTLEX (English) and the SUBTLEX-DE (German) corpora of film and TV subtitles (Marian et al., 2012). The frequency rates of base words plus |s|-affixes were calculated (e.g., English *cats* or German *kleines* 'small'). All lexical frequency rates were log transformed. Due to the presence of 0 values in lexical frequency rates (when words did not appear in the corpora), the constant value of 1 was added to all lexical frequency rates before taking a log transformation [log(x+1)].

Two second language corpora were used to estimate the frequency of word-final |s| types in L2 English and L3 German of Korean learners. For Korean English, the ICNALE Corpus, a collection of English as a foreign language data by Asian learners of different origins (Ishikawa, 2013) was used. The student turns of Korean speakers in the spoken corpus were searched for instances of phonemic and morphemic word-final |s|. Overall occurrences were divided by total number of tokens of the spoken corpus (95,381 tokens). For L3 German, written German texts by Korean learners of German (26 texts from 26 learners, total word count: 29,860) were extracted from the *Second Language Learner Corpus* (available at https://sites.google.com/gl-sec.com/korean-learners-german-corpus/home). The frequency rates of word-final phonemic and morphemic |s| were calculated.

Lexical frequency rate is known to be correlated with the word length, and the number of syllables per word is an important determiner for segment duration in stress-timed languages, such as English and German (Pamies Bertrán, 1999). The number of syllables per English and German target word were manually counted.

An individual global speech rate was calculated for each participant. Six filler sentences of the experimental reading tasks were analyzed in terms of the number of uttered syllables per second and the average number served as the variable "speech rate" per participant.

Base duration, defined as the duration of the target word minus the |s| duration was also added as a control variable. It can serve an indicator of local speech tempo and influenced |s| durations in the Plag et al. (2017) study.

Lastly, the release likelihood of word-final obstruents in English loanwords in Korean was considered. Kang (2003) describes various phonological scenarios which are particularly conducive to word-final |s| release, including monosyllabicity of target words, long vowels preceding the obstruents, and voicing of obstruents, although none of these is categorical. Thus, the variable "release likelihood" was composed of these three components, with each factor contributing equally to it. Values for "release likelihood" ranged from 0 (no feature present) to 3 (all features present) in English, and from 0 to 2 in German (succeeding long vowels, monosyllabic word; but excluding voicing as all word-final |s| were devoiced).

Table 2 gives an overview of the predictors and their characteristics.

Table 2
Overview of predictors and their characteristics.

Type	Predictor	Characteristics	
		English	**German**
Categorical	Type of |s|	phoneme, plural, 3rd person, is	phoneme, plural, inflection, derivation, genitive, clitic, es
	Sentence position	Mid-sentence, sentence-final	
	Preceding segment	plosive, nasal, vowel, fricative, liquid	plosive, nasal, vowel
	Following segment	approximant, fricative, liquid, nasal, plosive, vowel, none	fricative, plosive, vowel, none
	Grapheme	s, ss, c	s, ss: ß
	Syllables	1, 2, 3	
	Release likelihood	0, 1, 2, 3	0, 1, 2
Continuous	Voicing	Percentage of glottal pulses	
	Lexical frequency	Token frequency	
	Base duration	Token duration minus word-final |s| duration (local speech rate)	
	Speech rate	Global speech rate measured with filler sentences	

The variables used for the L1 English and L1 German models were identical to the L2 English and L3 German models, with the exception of the variable "release likelihood", which was entered only into the L2 and L3 models.

2.3 Statistics

As a first step, an ANOVA was calculated on the dependent variable "duration of |s|" (absolute duration, Box-Cox transformed) with s-types as groupings to see whether there are any differences among the different types of word-final |s|. For this, the R function *aov* was used. This was followed by pair-wise comparisons of the means using Tukey contrasts for unbalanced group sizes (Herberich et al., 2010), implemented with the R packages 'multcomp' (Hothorn et al., 2008) and 'sandwich' (Zeileis, 2006).

Next, linear mixed effects models (Bates et al., 2014) were calculated with the dependent variable "duration of |s|" (absolute values, Box-Cox transformed) and the following fixed effects: type of |s|, sentence position, preceding segment, following segment, voicing, grapheme, syllables, lexical frequency, speech rate, and base duration. The L2 and L3 models also included “release likelihood”. As random effect, “ID” of the participants was included, with maximal random slope structure (all fixed effects) in order to minimize type I error rate (Barr et al., 2013). In order to detect collinearity in the models that might be introduced when using the same predictor as a fixed and as a random effect (see Tomaschek et al., 2018, for a discussion), correlations between the dependent and the independent variables were calculated to glean the direction of an effect and compare that result to the regression results.

Full models were constructed for each language (L1 English, L1 German, L2 English, L3 German) and further investigated in terms of the best fit variables. The selection of variables to be entered into the mixed effects models was based on Baayen’s et al. (Baayen et al., 2008) suggestions and considered t-values of specific variables (only include those with values over 2 and below -2), AIC values between models (include a variable if it lowers AIC of a model), and a significant ANOVA p-value when comparing the full model to one without the variable under investigation. All models were implemented in R with the function *lmer* of the package ‘lme4’ (Bates et al., 2014). The significance of individual fixed effects in the final model (i.e., containing only the selected variables) was determined with the function *summ* of the package ‘jtools’ (Long, 2022). Model diagnostics were run with the R package ‘car’ (Fox & Weisberg, 2019). Graphs of results were plotted with the R packages ‘ggplot2’ (Wickham, 2016). Correlations among the numerical variables were calculated with the *cor* function in R, set for “Pearson”; correlations among the categorical variables were calculated with Cramer’s V using the R package ‘rcompanion’ (Mangiafico, 2019).

The sample size of the L1 English models was 485 tokens, 79 types, and 7 speakers; the L2 English models contained 1655 tokens, 79 types, and 22 speakers. The L1 German models consisted of 267 tokens, 23 types, and 10 speakers; the sample size for the L3 German models was 537 tokens, 23 types, and 22 speakers.

2.3.1 Model selection

<u>*L1 English*</u>

Following the variable exclusion procedure outlined earlier (see Baayen et al., 2008), the variables “sentence position”, “lexical frequency”, “grapheme” and “speech rate” were removed from the L1 English models. Correlations among the independent variables (see Table 3) yielded no significant correlations that would have prompted the removal of a variable.

Table 3
Correlation matrix of the independent variables in the L1 English model.

	Preceding segment	Following segment	Syllables	Base duration
Type of s	0.41	0.28	0.26	
Preceding segment		0.27	0.23	
Following segment			0.26	
Voicing				-0.12

Variance inflation did not seem to be a problem with any of the variables (see Sheather, 2009). Table 4 lists the variance inflaction factors per variable as well as the AIC values per model that lacked a specified fixed effect ("variable excluded") to be compared to the AIC of the full model containing all variables.

Table 4
Final variables included in L1 English model with variance inflation factors and AIC values.

Variable	VIF	AIC full model: -1415 AIC when variable excluded
Type of s	1.25	-1344
Preceding segment	1.17	-1401
Following segment	1.19	-1395
Voicing	1.36	-1404
Syllables	1.44	-1398
Base duration	1.46	-1408

The final L1 English model included the following predictors: "|s|-type", "preceding segment", "following segment", "voicing", "syllables", and "base duration".

<u>*L2 English*</u>
The selection procedure mandated the removal of the following variables from the L2 English model: "sentence position", "syllables", "lexical frequency", "speech rate", and "grapheme". Correlations are presented in Table 5.

Table 5
Correlation matrix of the independent variables in the L2 English models.

	Preceding segment	Following segment	Release likelihood	Base duration
Type of s	0.4	0.26	0.34	
Preceding segment		0.3	0.29	
Following segment			0.21	
Voicing				-0.12

See Table 6 for variance inflation factors and AIC values of the models.

Table 6
Final variables included in L2 English model with variance inflation factors and AIC values.

Variable	VIF	AIC full model: -1350 AIC when variable excluded
Type of s	1.23	-1346
Preceding segment	1.18	-1311
Following segment	1.2	-1320
Voicing	1.06	-1307
Base duration	1.4	-1322
Release likelihood	1.7	-1334

The L2 English model was finalized with the following predictors: "|s|-type", "preceding segment", "following segment","voicing", "base duration", and "release likelihood".

L1 German

The variables "syllables", "lexical frequency", "base duration", and "sentence position" were removed from the original model. The high correlation between "type of |s|" and "preceding segment" led to the latter variable being removed from the final model (see Table 7).

Table 7
Correlation matrix of the independent variables in the L1 German models.

	Preceding segment	Following segment	Grapheme	Speech rate
Type of s	0.68	0.48	0.33	
Preceding segment		0.35	0.1	
Following segment			0.35	
Voicing				-0.07

Variance inflation factors and AIC values can be found in Table 8.

Table 8
Final variables included in L1 German model with variance inflation factors and AIC values.

Variable	**VIF**	**AIC full model: 55 AIC when variable excluded**
Type of s	1.23	52
Following segment	1.25	57
Voicing	1.03	64
Grapheme	1.17	57
Speech rate	1.0	61

The final L1 German model was constructed including the predictors "|s|-type", "following segment", "voicing", "grapheme", and "speech rate".

L3 German
First, the variables "sentence position" and "base duration" were removed. As in the L1 German model, "type of |s|" and "preceding phoneme" were highly correlated (see Table 9), leading to the removal of the latter variable.

Table 9
Correlation matrix of the independent variables in the L3 German models.

	Preceding segment	**Following segment**	**Grapheme**	**Syllables**	**Release likelihood**	**Lexical frequency**	**Speech rate**
Type of s	0.68	0.48	0.33	0.5	0.6		
Preceding seg.		0.35	0.1	0.46	0.44		
Following seg.			0.35	0.16	0.49		
Graphemes				0.27	0.26		
Syllables					0.5		
Voicing						0.14	0.25
Lexical frequency							-0.01

Table 10 shows the AIC values per model and the variance inflation factors per variable. The variable "release likelihood" was removed due to its poor estimation in the model (>5, see Sheather, 2009; Tomaschek et al., 2018).

Table 10
Variables included in L3 German model with variance inflation factors and AIC values.

Variable	VIF	AIC full model: -1803 AIC when variable excluded
Type of s	1.53	-1787
Preceding segment	2.26	-1787
Following segment	2.59	-1766
Voicing	1.04	-1774
Grapheme	1.52	-1795
Syllables	3.1	-1798
Lexical frequency	4.1	-1763
Speech rate	1.01	-1799
Release likelihood	8.84	-1773

The final L3 German model included the following predictors: "|s|-type", "preceding segment", "following segment", "voicing", "grapheme", "syllables", "lexical frequency", and "speech rate".

2.4 Descriptive overview

Descriptive comparisons of word-final |s| in Korean learners and L1 speakers of English and German yielded differences in frequency (see Figures 2 and 3). Korean speakers used overall fewer English words ending in phonemic and morphemic |s| than L1 English speakers. Especially the verb *is* was shown to be overused. The inverse correlation between the clitic and *is* in this context may reflect a tendency of learners to use the full verb form rather than its clitic form.

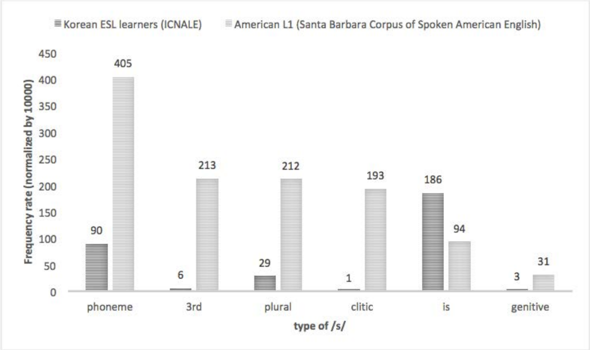

Figure 2. ***Normalized frequency of word-final |s| in Korean learners of English (International Corpus Network of Asian Learners of English) and L1 American speakers (Santa Barbara Corpus of Spoken American English).***

Korean learners of German resembled L1 German speakers more closely in their use of words ending in morphemic and phonemic |s|, which may be explained by the written nature of the L3 German corpus (no other corpus is available for Korean learners of German). Here, the genitive and plural markers were slightly overused in comparison to the Germans, while fewer phonemic |s| were recorded (see Figure 3).

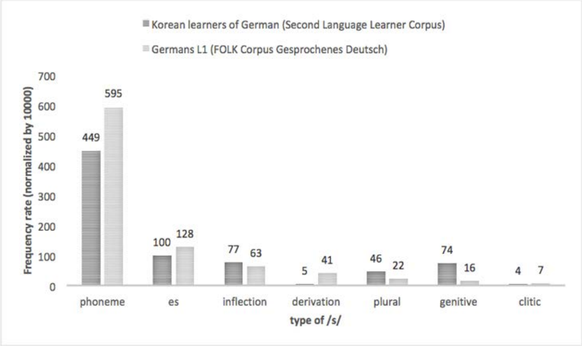

Figure 3. ***Normalized frequency of word-final |s| in Korean learners of German (Second Language Learner Corpus) and L1 German speakers (FOLK Corpus Gesprochenes Deutsch).***

Morphemic word-final |s| is more frequent in English, while German is generally characterized by a higher number of phonemic |s| in word-final position (see Figure 4). In terms of information bits associated with word-final |s| in English and German, a Shannon's entropy value of 2.31 for English indicates more uncertainty regarding the functional aspects of |s| than is the case in German (Shannon's entropy = 1.57).

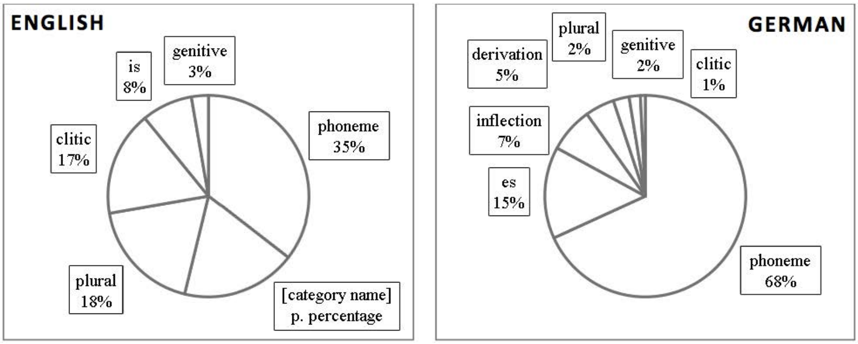

Figure 4. ***Frequencies of word-final |s| types in English (Santa Barbara Corpus) and German (FOLK Corpus).***

3. Results

3.1. English word-final |s| durations

3.1.1 L1 English

In L1 English, an ANOVA showed a significant effect of type of |s| ($F=35.79$, $p<0.001$) on the absolute duration of word-final |s|. The pair-wise Tukey comparisons identified various significant differences (see Table 11).

Table 11
Multiple pair-wise comparisons (Tukey contrasts). L1 English speakers.

	Estimate	Std. Error	t value	Pr(>|t|)	Sign.
Phonemic – plural	-0.11	0.01	-11.9	< 0.001	***
Phonemic –is	0.09	0.01	6.7	< 0.001	***
Phonemic – 3rd person	-0.14	0.02	-7.6	< 0.001	***
Plural – is	-0.02	0.01	-1.8	0.25	
Plural – 3rd person	-0.02	0.02	-1.35	0.52	
is – 3rd person	-0.05	0.02	-2.37	0.07	

Significance codes: "*"=<0.05, "**"=<0.01, "***"=<0.001

Phonemic |s| was the longest (see Figure 5).

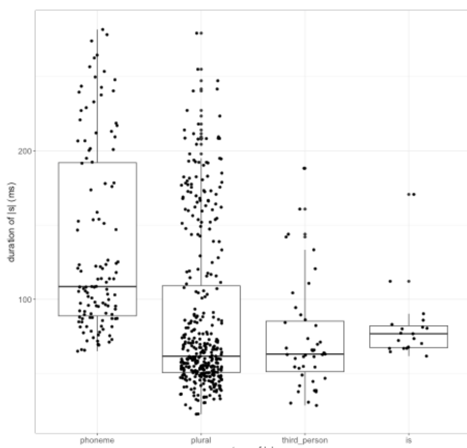

Figure 5. ***Durations of different types of |s| in L1 English speakers.***

Table 12 shows the result of the L1 English fixed effects model. Pseudo R-squared for the fixed effects equaled 0.77; pseudo R-squared for the total model equaled 0.78.

Table 12
Results of the L1 English mixed effects model.

Fixed effect	Estimate	St. Error	t-value	Pr(>|t|)	Signif.
Intercept	2.38	0.02	110.31	< 0.001	
Type of |s|: phoneme	-0.01	0.01	-1	0.32	
|s|: plural	-0.09	0.02	-5.77	< 0.001	***
|s|: third person	-0.09	0.02	-4.25	< 0.001	***
Preceding segment: vowel	-0.01	0.01	-0.95	0.35	
Preceding segment: nasal	-0.07	0.01	-8.74	< 0.001	***
Preceding segment: plosive	-0.04	0.01	-3.62	< 0.001	***
Following segment: none	0.13	0.01	8.64	< 0.001	***
Following segment: fricative	0	0.01	0.08	0.93	
Following segment: liquid	0.02	0.02	1	0.32	
Following segment: nasal	0.01	0.03	0.48	0.63	
Following segment: plosive	0.01	0.02	0.56	0.59	
Following segment: vowel	-0.03	0.01	-2.96	0.01	*
Voicing	-0	0	-6.29	< 0.001	***
Syllables	-0.04	0.01	-6.52	< 0.001	***
Base duration	0.2	0.04	4.69	< 0.001	***

Significance codes: "*"=<0.05, "**"=<0.01, "***"=<0.001

Word-final |s| was longest when it succeeded and shortest when it preceded a vowel. Monosyllabic words displayed the longest |s| durations, as did voiced |s| segments. In addition, word-final |s| increased in length together with base durations of words. The results for voicing and base duration are consistent with the correlation coefficients calculated with the dependent variable (Pearson's r=-0.63 and 0.02).

3.1.2 L2 English

In L2 English the ANOVA yielded an effect of type of |s| on the absolute durations of word-final |s| ($F=28.1$, $p<0.001$). Pair-wise contrasts are summarized in Table 13.

Table 13
Multiple pair-wise comparisons (Tukey contrasts). L2 English speakers.

	Estimate	Std. Error	t value	Pr(>|t|)	Sign.
Phonemic – plural	-0.12	0.01	-7.9	< 0.001	***
Phonemic –is	0.26	0.03	8.1	< 0.001	***
Phonemic – 3rd person	-0.15	0.02	-6.1	< 0.001	***
Plural – is	0.14	0.03	4.6	< 0.001	***
Plural – 3rd person	-0.03	0.02	-1.5	0.45	
is – 3rd person	0.11	0.04	2.9	0.016	*

Significance codes: “*”=<0.05, “**”=<0.01, “***”=<0.001

As displayed in Figure 6, phonemic |s| was the longest, followed by plural |s|.

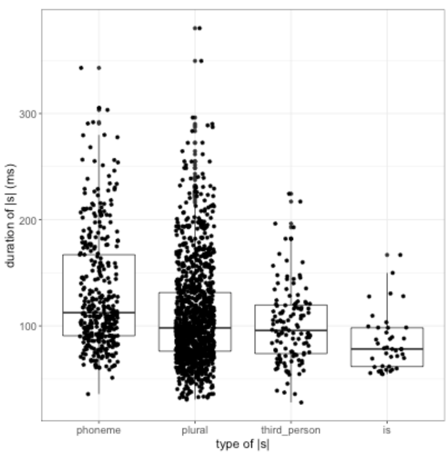

Figure 6. ***Durations of different types of |s| in L2 English speakers.***

Table 14 shows the result of the L2 English fixed effects model. Pseudo R-squared for the fixed effects equaled 0.6; pseudo R-squared for the total model equaled 0.67.

Table 14
Results of the L2 English mixed effects model.

Fixed effect	Estimate	St. Error	t-value	Pr(>|t|)	Signif.
Intercept	3.75	0.04	99.4	< 0.001	
Type of |s|: phoneme	0.06	0.03	95.7	0.04	*
|s|: plural	0.01	0.03	0.43	0.67	
|s|: third person	0.04	0.03	1.21	0.23	
Preceding segment: vowel	-0.02	0.01	-1.56	0.12	
Preceding segment: nasal	-0.11	0.01	-9.59	< 0.001	***
Preceding segment: plosive	-0.1	0.02	-4.71	< 0.001	***
Following segment: none	0.16	0.02	7.88	< 0.001	***
Following segment: fricative	-0.07	0.02	-3.67	< 0.001	***
Following segment: liquid	0.05	0.04	1.4	0.17	
Following segment: nasal	-0.12	0.03	-3.93	< 0.001	***
Following segment: plosive	-0.08	0.02	-4.2	< 0.001	***
Following segment: vowel	-0.06	0.02	-3.57	< 0.001	***
Voicing	-0.0	0.0	-11.49	< 0.001	***
Base duration	0.27	0.05	5.36	< 0.001	***
Release likelihood	0.04	0.01	6.1	< 0.001	***

Significance codes: "*"=<0.05, "**"=<0.01, "***"=<0.001

Nasals preceding the word-final sibilants led to a durational decrease, while plosives led to an increase. Word-final |s| were significantly lengthened when no following segment appeared (utterance-final, pre-pausal) and shortest when fricatives, nasals, or vowels followed. As with the L1 English sample, voiced |s| were longer than their unvoiced counterparts, and longer base durations indicated longer sibilants. The correlation coefficients calculated with the dependent variable yielded Pearson's r=-0.5 for voicing and r=-0.11 for base duration. The influence of "release likelihood on |s| duration can be seen in Figure 7. High release likelihood was associated with shorter durations.

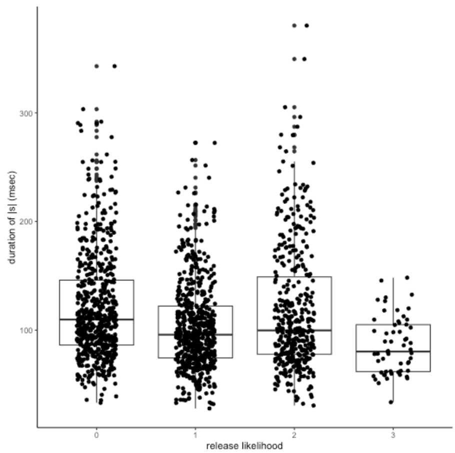

Figure 7. ***Highest release likelihood was associated with shortest |s| durations in L2 English speakers.***

3.2. German word-final |s| durations

3.2.1 L1 German

The L1 German ANOVA also showed a significant effect of type of |s| (F=5.1, p<0.001; see Figure 8), with pair-wise comparisons listed in Table 15.

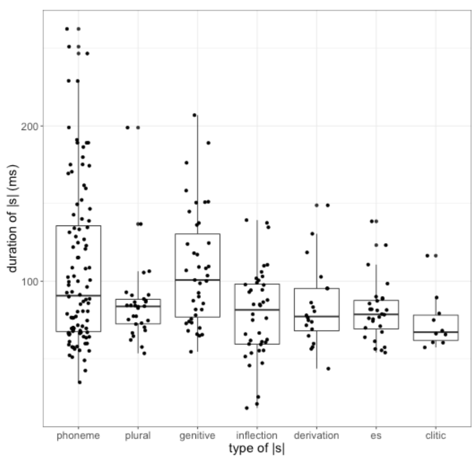

Figure 8. ***Durations of different types of |s| in L1 German (Austrian) speakers.***

Table 15
Multiple pair-wise comparisons (Tukey contrasts). L1 German speakers.

	Estimate	Std. Error	t value	Pr(>|t|)	Sign.
Phonemic – plural	-0.12	0.05	-2.5	0.18	
Phonemic –genitive	-0.04	0.05	-0.7	0.99	
Phonemic – inflection	0.23	0.07	3.4	0.012	*
Phonemic – derivation	0.15	0.06	2.37	0.21	
Phonemic – es	0.16	0.05	3.36	0.015	*
Phonemic – clitic	0.22	0.06	3.46	0.01	*
Plural – genitive	-0.16	0.05	-2.94	0.05	*
Plural – inflection	0.11	0.07	1.59	0.68	
Plural – derivation	0.03	0.06	0.43	0.99	
Plural – es	0.04	0.05	0.82	0.98	
Plural – clitic	0.1	0.07	1.59	0.68	
Genitive – inflection	-0.27	0.07	-3.77	0.004	**
Genitive – derivation	0.19	0.07	2.8	0.075	
Genitive – es	0.2	0.05	3.76	0.003	**
Genitive – clitic	0.26	0.07	3.83	0.003	**
Inflection – derivation	-0.08	0.08	-1.03	0.94	
Inflection – es	-0.07	0.07	-1.01	0.95	
Inflection – clitic	-0.003	0.08	-0.04	1.0	
Derivation – es	-0.01	0.06	-0.2	0.99	
Derivation – clitic	0.08	0.07	1.01	0.95	
Es – clitic	0.06	0.06	0.99	0.95	

Significance codes: "*"=<0.05, "**"=<0.01, "***"=<0.001

Table 16 shows the result of the L1 German model. Pseudo R-squared for the fixed effects equaled 0.54; pseudo R-squared for the total model equaled 0.6.

Table 16
Results of the L1 German mixed effects model.

Fixed effect	Estimate	St. Error	t-value	Pr(>|t|)	Signif.
Intercept	4.93	0.23	21.02	<0.001	
Type of |s|: phoneme	0.14	0.07	2.03	0.05	*
|s|: plural	0.18	0.08	2.26	0.03	*
|s|: genitive	0.08	0.08	1.07	0.29	
|s|: inflection	0.07	0.08	0.82	0.42	
|s|: derivation	0.13	0.08	1.62	0.11	
|s|: es	0.15	0.08	1.84	0.07	
Following segment: none	0.38	0.04	8.54	< 0.001	***
Following: plosive	-0.09	0.04	-1.97	0.07	
Following: vowel	-0.08	0.04	-2.04	0.04	*
Voicing	-0.0	0	-5.22	< 0.001	***
Grapheme: ss	0.03	0.05	0.56	0.58	
Grapheme: sz	0.18	0.07	2.53	0.01	**
Speech rate	-0.21	0.04	-4.89	<0.001	***

Significance codes: "*"=<0.05, "**"=<0.01, "***"=<0.001

The longest word-final |s| durations were recorded preceding fricatives and pauses (no following segment), the shortest ones when vowels followed. First language users of German pronounced the grapheme <ß> with the longest duration (see Figure 9).

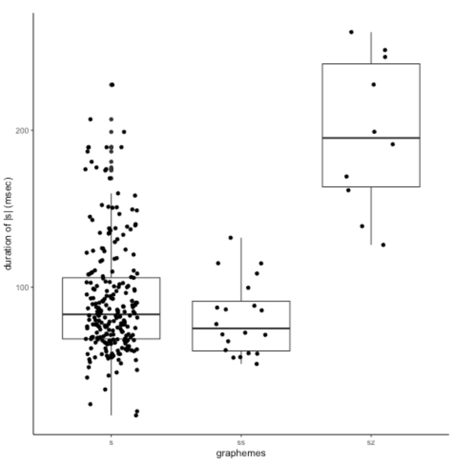

Figure 9. ***The grapheme <ß> ("sz") was associated with the longest |s| durations.***

Figure 10 shows that the overwhelming majority of word-final |s| were devoiced in German but the higher the voiced portions, the shorter the |s| became (Pearson's r: dependent variable and voicing =-0.36).

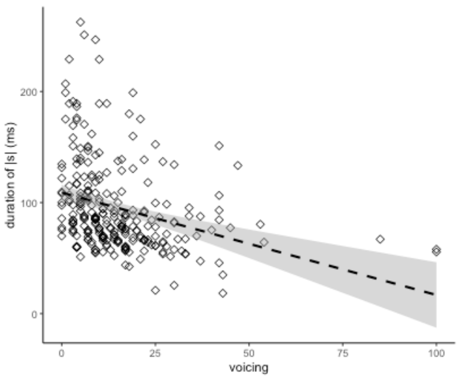

Figure 10. ***L1 German word-final |s| are generally unvoiced.***

The effect of speech rate on |s| durations was intuitive, with |s| becoming shorter in faster speech (Pearson's r: speech rate and dependent variable =-0.17).

3.2.2 L3 German

In L3 German the ANOVA results also yielded significant results for type of |s| (F=12.02, p<0.001; see Figure 11). The Tukey contrasts can be found in Table 17.

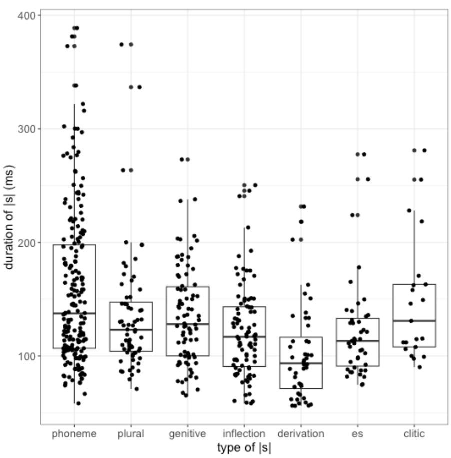

Figure 11. ***Absolute durations of different types of |s| in Korean speakers of German.***

Table 17
Multiple pair-wise comparisons (Tukey contrasts). L3 German speakers.

	Estimate	Std. Error	t value	Pr(>|t|)	Sign.
Phonemic – plural	-0.02	0.007	-2.47	0.16	
Phonemic –genitive	0.02	0.007	2.45	0.17	*
Phonemic – inflection	0.03	0.007	4.85	< 0.001	***
Phonemic – derivation	0.07	0.01	6.42	< 0.001	***
Phonemic – es	0.03	0.008	4.01	0.0012	**
Phonemic – clitic	0.004	0.012	0.35	0.99	
Plural – genitive	-0.001	0.008	-0.12	1.0	
Plural – inflection	0.02	0.008	2.01	0.39	
Plural – derivation	0.05	0.011	4.4	< 0.001	***
Plural – es	0.02	0.009	1.65	0.63	
Plural – clitic	-0.014	0.012	-1.12	0.92	
Genitive – inflection	-0.018	0.008	-2.22	0.27	
Genitive – derivation	0.05	0.01	4.58	< 0.001	***
Genitive – es	0.02	0.01	1.81	0.52	
Genitive – clitic	-0.013	0.01	-1.1	0.94	
Inflection – derivation	0.03	0.01	2.98	0.044	*
Inflection – es	-0.001	0.01	-0.14	1.0	
Inflection – clitic	-0.03	0.02	-2.48	0.16	
Derivation – es	0.04	0.01	2.89	0.05	*
Derivation – clitic	-0.06	0.01	-4.39	< 0.001	***
Es – clitic	-0.03	0.01	-2.24	0.26	

Significance codes: “*”=<0.05, “**”=<0.01, “***”=<0.001

Table 18 shows the result of the L3 German model. Pseudo R-squared for the fixed effects equaled 0.56; pseudo R-squared for the total model equaled 0.62.

Table 18
Results of the L3 German mixed effects model.

Fixed effect	Estimate	St. Error	t-value	Pr(>|t|)	Signif.
Intercept	2.32	0.04	63.62	< 0.001	
Type of |s|: phoneme	-0.02	0.01	-1.3	0.2	
|s|: plural	-0.01	0.01	-0.42	0.68	
|s|: genitive	-0.01	0.01	-1.1	0.28	
|s|: inflection	-0.02	0.01	-1.4	0.17	
|s|: derivation	0.01	0.02	0.31	0.75	
|s|: es	0.0	0.01	0.12	0.91	
Voicing	-0.0	0.0	-8.0	< 0.001	***
Following sound: none	0.06	0.01	6.37	< 0.001	***
Following: plosive	-0.02	0.01	-3.36	< 0.001	***
Following: vowel	-0.0	0.01	-0.67	0.51	
Grapheme: ss	0.02	0.01	2.33	0.02	*
Grapheme: ß, sz	0.01	0.01	0.76	0.45	
Syllables	-0.01	0.01	-1.32	0.19	
Lexical frequency	-0.01	0.0	-4.16	< 0.001	***
Speech rate	-0.03	0.01	-3.45	< 0.001	***

Significance codes: “*”=<0.05, “**”=<0.01, “***”=<0.001

Word-final |s| showed longest durations when no segments followed. In the case of preceding a plosive consonant, |s| contracts the most. Sibilants transcribed with the grapheme <ß> showed the longest durations also in speakers of German as a foreign language. Voicing generally shortened word-final |s| in Korean speakers of German, and the proportion of voiced |s| was larger than in the German first language speakers (see Figure 12). Pearson’s r showed a negative correlation between voicing and the dependent variable (-0.41).

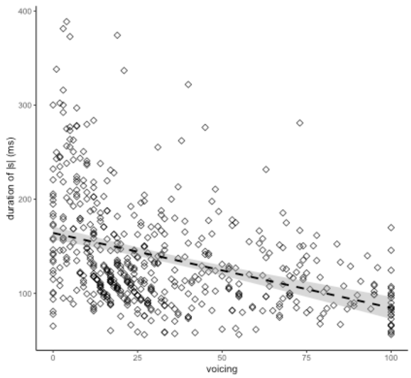

Figure 12. ***Korean learners of L3 German frequently voice word-final |s|.***

Words of higher lexical frequency rates displayed reduced word-final |s| (Pearson's r dependent variable with lexical frequency=-0.12), and faster speech rates also shortened |s| (Pearson's r: dependent variable and speech rate=-0.14).

4. Discussion

This study investigated the link between morphemic status and phonetic implementation of homophonous word-final |s| in Korean learners of English and German. In English and German as first languages, word-final |s| durations can be related to morphological boundary strengths, with phonemes and morphemes differing in overall durations. As Korean phonology is characterized by a merger of all word-final obstruents as unreleased [t̚], no acoustic surface form exists for word-final |s| in the language. Tracing how Korean learners of English and German fill this acoustic gap can yield insights into phonetic implementation of subphonemic detail in foreign languages.

The findings on word-final |s| durations in L1 English speakers are in accordance with previous literature that describes phonemic |s| as longer than morphemic |s| types in various varieties of English (e.g., Plag et al., 2017; Zimmermann, 2016). Those studies investigated spontaneous speech, and the present study demonstrates the same effect in laboratory-elicited speech. In terms of Korean learners' approximation of the L1 English phonetic patterns, it was clearly shown that the same phonetic length differentiations across the |s| types were also found in L2 English. Phonemic |s| was also the longest |s| type in Korean L2 English, with all morphemic types significantly reduced in length. This suggests that Korean learners of English are able to incorporate English subphonemic detail into their L2 English speech. It is known from loan phonology (see, e.g., Iverson & Lee, 2006) that perception of phonetic subtleties can be rather sophisticated in Korean listeners/ speakers of English words. The phonetic experience that Korean learners of English have with L1 English likely shapes their phonetic spaces associated with the various |s| types, and this effect may be mediated by occurrence probability of the |s| types (see Tomaschek et al., 2019). The sequence of |s| durations was identical in L1 and L2 speakers, with the exception of *is*, which was prolonged in the Korean learners. L1 English users contract this verb more. Confounding variables impacting on |s| durations in English were also quite similar between the L1 and the L2 speakers. Unvoiced |s| was of longer duration (see, e.g., Ernestus et al., 2006; Plag et al., 2017); the findings on phonological environments and base durations are also in agreement with what previous studies have shown.

Similar to English, German |s| types are also distinguished by durational cues, but important differences emerged. German first language users produced genitive |s| with the longest overall durations, followed by phonemic |s|. The other morphemic |s| types were more reduced. Given the fact that genitive-s suffixes are rare in Austrian German (Bülow et al., 2022), their lengthening could result from a low-frequency effect. While it is unclear whether this prolonged |s| type reflects a regional phonetic variation, unpublished findings on Standard German spoken in Berlin and the surrounding province suggest that morphemic word-final |s| of that variety is also longer than phonemic |s| [link to data set in an open repository: http://hdl.handle.net/11234/1-5044] More research is needed to gain clarity on the phonetic realization of word-final |s| in German varieties.

Concerning L3 German, Korean learners produced phonemic |s| and clitics with the longest durations and showed phonetic reduction in the other morphemic types. As opposed to L1 German speakers, L3 speakers of German prolonged clitics and markedly reduced derivations. In terms of overall occurrence probability, clitics are rare in spoken German (see Figure 4) and this may influence Korean learners' phonetic experience with the |s| in these constructions. Koreans' L3 German speech resembled L1 German in the fact that graphemes influenced |s| durations, with <ß> being longer than <ss>, and <s> showing the shortest duration. The German grapheme <ß> was historically linked to long vowels and diphthongs in German that also lead to lengthening of succeeding |s| (Walder, 2020). Even though vowel length has in recent decades become independent of <ß>, allowing, for instance, *Spaß* ('fun') to be pronounced with long or short vowel in speakers of central and northern German (Tröster-Mutz, 2004), Austrian <ß> is still strongly associated with long preceding vowels and lengthened |s|. While L3 German speech was influenced by similar variables than L1 German (e.g., type of |s| variation, following segment, voicing, grapheme), there were some additional variables that influenced Koreans' but not Austrians' speech (e.g., number of syllables, lexical frequency rate of target words). Less phonetic experience with German could result in less clearly defined (or less stable) phonetic spaces for German sounds and potentially more numerous factors that influence phonetics.

Phonetic variability in second languages can be higher as compared to first languages (Baese-Berk & Morrill, 2015; Jongman & Wade, 2007; Wade et al., 2007), but that also depends on the specific phonetic feature in question (Baker et al., 2011; Morrill et al., 2016; Vaughn et al., 2019). In addition, the specific L1s and L2s under investigation as well as the proficiency learners have achieved in an L2 play crucial roles for individual and group-level phonetic variability (Jongman & Wade, 2007; Vaughn et al., 2019). Learners of foreign languages typically map foreign sounds to known ones from the first language (Escudero, 2009), but when a sound has no correspondence in L1, learners may be better able to more accurately chart a phonetic space for it that more closely resembles L1 (Bohn & Flege, 1992; Flege, 1995). In such cases, the learners have to rely more on the acoustic input they receive from L1 users of the foreign language, with less interference from their own L1 (but see Escudero et al., 2014). Subsequently, phonetic experience starts to play a larger role. Some of the phonetic variability in foreign languages is inevitably linked to language proficiency, and less proficiency may invite more phonetic variation due to a larger degree of phonetic instability and/ or phonetic insecurity on the part of the learners (be it due to perceptual or production-related processes). Koreans study English from an early age and the (written) English language is ubiquitous in Korea (at least in the Seoul region where the participants for the present study were recruited, see Luef, Ghebru & Ilon, 2018), thus Koreans are regularly exposed to English. German is a rare third language in Korea nowadays and the German proficiency level of Korean students is generally below that of English (Luef, Ghebru, & Ilon, 2019). Phonetic learning in general is certainly more progressed in Korean learners of English, and honing in on particular L1 English subphonemic features, which are influential in the language, could be more developed.

One crucial question concerns the acquisition mechanisms of the acoustic details of the sibilant durations in the second and third languages. Are the durations learned through exposure, or do the productions reflect the same psycholinguistic mechanism (be they listener-oriented or speaker-oriented) that cause the durational differences within the first-language speakers? Likewise, one could ask whether the first language speakers learn the durational differences from their input data or simply produce them because of psycholinguistic mechanisms. The present data cannot attempt to find answers to these questions but they may be probed by future studies focussing on the developmental trajectories of word-final sibilant acoustics. The present study has shown that Korean L2 English learners closely mirror L1 English users' phonetic patterns, while Korean L3 German learners show less convergence with L1 German speakers. If purely psycholinguistic mechanisms were the cause of the durational differences between the |s|-types, one would not expect Korean L3 German speakers to adapt the acoustics of their word-final |s| as their German evolves over time. A development towards more similarity with L1 German |s| durations when learners attain higher linguistic (and phonetic) proficiency in German, might indicate a role of perception-based phonetic learning. Additionally, future studies should focus on the early L2 English acquisition phase in Korean learners to understand whether English-like word-final |s| acoustics change as learning progresses.

5. Conclusion

Korean learners of L2 English and L3 German show sensitivity to the systematic morpho-phonetic variation that can be in word-final |s| in L1 English and L1 German. Korean learners of foreign languages seem to implicitly gather subphonemic details from their experience with the target languages and incorporate them into their productive abilities.

References

Abramson, A. S., & Whalen, D. H. (2017). Voice onset time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. *Journal of Phonetics, 63*, 75-86.

Ahn, K. (2011). Conceptualization of American English native speaker norms: A case study of an English language classroom in South Korea. *Asia Pacific Education Review, 12*, 691-702.

Ahn , S.-C., & Iverson, G. K. (2004). Dimensions in Korean laryngeal phonology. *Journal of East Asian Linguistics, 13*, 345-379.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subject and items. *Journal of Memory & Language, 59*(4), 390-412.

Baese-Berk, M., & Morrill, T. H. (2015). Speaking rate consistency in native and non-native speakers of English. *The Journal of the Acoustical Society of America, 138*, EL223-228.

Baese-Berk, M., & Samuel, A. G. (2016). Listeners beware: Speech production may be bad for learning speech sounds. *Journal of Memory & Language, 89*, 23-36.

Baker, R. E., Baese-Berk, M., Bonnasse-Gahot, L., Kim, M., van Engen, K. J., & Bradlow, A. (2011). Word durations in non-native English. *Journal of Phonetics, 39*, 1-17.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255-278.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). {lme4}: Linear mixed-effects models using Eigen and S4. *R Package version 1.1-7*.

Berkovits, R. (1993). Utterance-final lengthening and the duration of final-stop closures. *Journal of Phonetics, 21*(4), 479-489.

Boersma, P., & Weenink, D. (2019). Praat, http://www.Praat.org (Version 6.0.46).

Bohn, O. S., & Flege, J. E. (1992). The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition, 14*, 131-158.

Box, G. E. P., & Cox, D. (1964). An analysis of transformations. *Journal of the Royal Statistical Society, Series B, 26*(2), 211-252.

Brown, E. K. (2009). The relative importance of lexical frequency in syllable- and word-final /s/ reduction in Cali, Colombia. In J. Collentine (Ed.), *Selected Proeedings of the 11th Hispanic Linguistics Symposium* (pp. 165-178). Cascadilla Proceedings Project.

Bülow, L., Vergeiner, P. C., & Elspaß, S. (2022). Structures of adnominal possession in Austria's traditional dialects: Variation and change. *Journal of Linguistic Geography, 9*(2), 69-85.

Cheon, S. Y., & Anderson, V. B. (2008). Acoustic and perceptual similarities between English and Korean sibilants. *Korean Linguistics, 14*, 41-64.

Cho, S., & Whitman, J. (2019). *Korean: A linguistic introduction*. Cambridge University Press.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. Harper and Row.

Council of Europe. (2018). *Common European framework of reference for languages: Learning, teaching, assessment. Companion volume with new descriptors*. https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2018/1680787989

Delattre, P. (1966). A comparison of syllable length conditioning among languages. *International Review of Applied Linguistics, 43*, 183.

Ernestus, M., Lahey, M., Verhees, F., & Baayen, R. H. (2006). Lexical frequency and voice assimilation. *Journal of the Acoustical Society of America, 120*(2), 1040-1051.

Escudero, P. (2009). Linguistic perception of "similar" L2 sounds. In P. Boersma & S. Hamann (Eds.), *Phonology in perception* (pp. 151-190). Mouton de Gruyter.

Escudero, P., Sisinni, B., & Grimaldi, M. (2014). The effect of vowel inventory and acoustic properties in Salento Italian learners of Southern British English vowels. *The Journal of the Acoustical Society of America, 135*, 1577-1584.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). York Press.

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression*. Sage.

Gahl, S. (2008). 'Time' and 'thyme' are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language, 84*(3), 474-496.

Haggard, M. (1972). Abbreviation of consonants in English pre-post-vocalic clusters. *Journal of Phonetics, 3*, 7-24.

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*(3), 1644-1655.

Herberich, E., Sikorski, J., & Hothorn, T. (2010). A robust procedure for comparing multiple means under heteroscedasticity in unbalanced designs. *PLoS ONE, 5*(3), e9788. https://doi.org/https://doi.org/10.1371/journal.pone.0009788

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal, 50*(3), 346-363.

Inkelas, S. (2014). *The interplay of morphology and phonology*. Oxford University Press.

Iverson, G. K., & Lee, A. (2006). Perception of contrast in Korean loanword adaptation. *Korean Linguistics, 13*, 49-87.

Jeon, H.-S., & Nolan, F. (2013). The role of pitch and timing cues in the perception of phrasal grouping in Seoul Korean. *Journal of the Acoustical Society of America, 133*(3039). https://doi.org/https://doi.org/10.1121/1.4798663

Jongman, A., & Wade, T. (2007). Acoustic variability and perceptual learning. In O. S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In Honor of James Emil Flege* (pp. 135-150). John Benjamins Publishing Company.

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America, 108*, 1252-1263.

Kang, Y. (2003). Perceptual similarity in loanword adaptation: Adaptation of English post-vocalic word-final stops in Korean. *Phonology, 20*, 219-273.

Kim, H., & Jongman, A. (1996). Acoustic and perceptual evidence for complete neutralization of manner of articulation in Korean. *Journal of Phonetics, 24*, 295-312.

Kim, S. (1999). *Sub-phonemic duration difference in English /s/ and few-to-many borrowing from English to Korean* University of Washington].

Kim, S., & Curtis, E. (2002). Phonetic duration of English /s/ and its borrowing into Korean. *Japanese/ Korean Linguistics, 10*, 406-419.

Kinoshita, S., Gayed, M., & Norris, D. (2018). Orthographic and phonological priming effects in the same-different task. *Journal of Experimental Psychology: Human perception and performance, 44*(11), 1661-1671.

Klatt, D. H. (1974). On predicting the duration of the phonetic segment [s] in English. *Journal of Speech and Hearing Research, 17*, 51-63.

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *The Journal of the Acoustical Society of America, 59*(5), 1208-1221.

Lee, S., & Katz, J. (2016). Perceptual integration of acoustic cues to laryngeal contrasts in Korean fricatives. *The Journal of the Acoustical Society of America, 139*(605). https://doi.org/http://dx.doi.org/10.1121/1.4926435

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.

Long, J. A. (2022). jtools: An analysis and presentation of social scientific data. *R package version 2.2.0*. https://cran.r-project.org/package=jtools

Losiewicz, B. L. (1995). Word frequency effects on the acoustic duration of morphemes. *Journal of the Acoustical Society of America, 97*, 3243.

Luef, E. M., Ghebru, B., & Ilon, L. (2018). Apps for language learning: Their use across different languages in a Korean context. *Interactive Learning Environments, 28/8*, 1036-1047.

Luef, E. M., Ghebru, B., & Ilon, L. (2019). Language proficiency and smartphone-aided second language learning: A look at English, German, Swahili, Hausa, and Zulu. *Electronic Journal of E-Learning, 17/1*, 25-37.

Mangiafico, S. (2019). rcompanion: Functions to support extension education program evaluation. *R package version*(10).

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-linguistic easy access resource for phonological and orthographic neighborhood densities. *PLoS ONE, 7*(8), e43230.

Mathieu, L. (2015). The influence of foreign scripts on the acquisition of a second language phonological contrast. *Second Language Research, 32*(2). https://doi.org/https://doi.org/10.1177/0267658315601882

Morrill, T. H., Baese-Berk, M., & Bradlow, A. (2016). Speaking rate consistency and variability in spontaneous speech by native and non-native speakers of English. *Proc. Speech Prosody, 8*, 1119-1123.

Pamies Bertrán, A. (1999). Prosodic typology: On the dichotomy between stress-timed and syllable-timed languages. *Language Design, 2*, 103-130.

Plag, I., Homann, J., & Kunter, G. (2017). Homophony and morphology: The acoustics of word-final S in English. *Journal of Linguistics, 53*(1), 181-216.

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

Schmidt, A. M. (1996). Cross-language identification of consonants Part 1: Korean perception of English. *The Journal of the Acoustical Society of America, 99*(5), 3201-3211.

Schmidt, A. M. (2007). Cross-language consonant identification: English and Korean. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 185-200). John Benjamins.

Sheather, S. (2009). *A modern approach to regression with R*. Springer Science & Business Media.

Snow, D. (1994). Phrase-final syllable lengthening and intonation in early child speech. *Journal of Speech, Language, and Hearing Research, 37*(4), 831-840.

Song, J. Y., Demuth, K., Evans, K., & Shattuck-Hufnagel, S. (2013). Durational cues to fricative codas in 2-year-olds' American English: Voicing and morphemic factors. *Journal of the Acoustical Society of America, 133*(5), 2931-2946.

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

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

Tröster-Mutz, S. (2004). Die Realisierung von Vokallängen: erlaubt ist, was Sp[a(:)]ß macht? . *SKY Journal of Linguistics, 17*, 249-265.

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

Umeda, N. (1977). Consonant duration in American English. *The Journal of the Acoustical Society of America, 61*(3), 846-858.

Vaughn, C., Baese-Berk, M., & Idemaru, K. (2019). Re-examining phonetic variability in native and non-native speech. *Phonetica, 76*(5), 327-358. https://doi.org/DOI: 10.1159/000 487269

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S*. Springer. https://www.stats.ox.ac.uk/pub/MASS4/

Wade, T., Jongman, A., & Sereno, J. (2007). Effects of aacoustic variability in the perceptual learning of non-native accented speech sounds. *Phonetica, 64*, 122-144.

Walder, A. (2020). Das versale Eszett: Ein neuer Buchstabe im deutschen Alphabet. *Zeitschrift f. Germanistische Linguistik, 48*(2), 211-237.

Walsh, T., & Parker, F. (1983). The duration of morphemic and non-morphemic /s/ in English. *Journal of Phonetics, 11*, 201-206.

Wang, Y., Jongman, A., & Sereno, J. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *The Journal of the Acoustical Society of America, 113*(2), 1033-1043.

Wedel, A., Kaplan, A., & Jackson, S. (2013). High functional load inhibits phonological contrast loss: A corpus study. *Cognition, 128*(2), 179-186.

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer.

Yuan, J., & Liberman, M. (2009). *Penn Phonetics Lab Forced Aligner for English.* In (Version 1.002)

Zeileis, A. (2006). Object-oriented computation of sandwich estimators. *Journal of Statistical Software, 16*(9), 1016.

Zimmermann, J. (2016). Morphological status and acoustic realization: Findings from New Zealand English. 16th Australasian International Conference on Speech Science and Technology, Paramatta, Australia.

https://doi.org/10.59960/12.1.a4

Supplementary Table S1

Carrier sentences containing different types of English and German word-final /s/. Longer compound nouns were separated by hyphen in order to facilitate reading for second language learners.

Language	Sentence with target words in bold	Type of /s/
English	**Buffalos** are large **animals**.	Plural, plural
	Tellers have to work long **hours**.	Plural, plural
	Puff adders are very **dangerous**.	Adjective
	Beavers live in lakes and **rivers**.	Plural, plural
	Gum **ruins** your teeth.	3rd person
	Deans of **colleges** have to work long **hours**.	Plural, plural, plural
	Tuxedos are the dress-code for **this** wedding.	Plural, phoneme
	Gills of fish can look different **ways**.	Plural, plural
	Customs is an agency responsible for collecting tariffs at the airport.	Plural, is
	Pucks are the balls of **ice** hockey.	Phoneme
	Pictures of Tom can be found everywhere in **this house**.	Plural, phoneme, phoneme
	Text-writing **is** a central feature of **this class**.	Is, phoneme, phoneme
	Garry **is his** first name.	Is, phoneme
	Bathrooms are green nowadays.	Plural
	Battles of World War 2 included the one at Normandy.	Plural
	Passion for sports **runs** in my family.	3rd person
	Kitties are little cats.	Plural
	Kerosine is fuel for jet **engines** and lamps.	Plural
	Geese can swim.	Phoneme
	Bees make honey.	Plural
	Dust **gathers** easily in the **corners** of apartments.	3rd person, plural
	Bats live in hollow **trees**.	Plural
	Dance **balls** are old-fashioned.	Plural
	Custard **recipes** are typically milk-based.	Plural
	Deals in the **business** world are hard to make.	Plural, phoneme
	Gifts are given for **Christmas**.	Phoneme
	Guesswork is the **process** of making a **guess** when you do not know all the facts.	Phoneme, phoneme

https://doi.org/10.59960/12.1.a4

Supplementary Table S1 continued...

Language	Sentence with target words in bold	Type of /s/
English	**Pumpkins** are my favorite vegetable.	Plural
	Dusk is the time before the sun **rises**.	3rd person
	Buddy **systems** for language learning are a great invention.	Plural
	Gut **microbes** are important for your health.	Plural
	Pieces of the cake are in **his** hair.	Plural, phoneme
	Tins have to be recycled.	Plural
	Pillows can be expensive in this store.	Plural
	Gekkos are little reptiles.	Plural, plural
	Deserts are defined as dry lands.	Phoneme, plural
	Kings of England.	Plural
	Buns for **burgers** can be very soft.	Plural, plural
	Guns are used for killing people.	Plural
	Punch **contains** a lot of sugar.	3rd person
	Teak wood **comes** from the rainforest.	3rd person
	Gigs of **musicians** will take **place** all over the country.	Plural, plural, phoneme
	Bundles of joy.	Plural
	Pills are generally prescribed by your doctor.	Plural
	Differences in opinion should not be expressed.	Plural
	Telegrams are not used any more nowadays.	Plural
	Chemicals in your clothes are bad for your skin.	Plural
	Cans have to be recycled.	Plural
	Bills just keep piling up.	Plural
	Gutters can be found on the street.	Plural
	Ducks live in lakes and **ponds**.	Plural
	Tennis **players** need to have strong **muscles** in their **arms**.	Plural, plural, plural
	Ticks carry lots of **diseases**.	Plural
	Kids have to go to school.	Plural
	Beach **houses** were affected by the hurricane.	Plural
	Gears in the car are for shifting.	Plural
	Kiss for you, kiss for me.	Phoneme

https://doi.org/10.59960/12.1.a4

Supplementary Table S1 continued...

Language	Sentence with target words in bold	Type of /s/
German	Badezimmer sind heute **meistens** blau.	Derivation
	Tafeln gibt **es** in grün oder **weiß**.	Es, phoneme
	Bilder von dir hängen überall im **Haus**.	Phoneme
	Textschreiben ist eine zentrale Übung **des** DaF-**Unterrichts**.	Phoneme, genitive
	Dämm-Material wird **fürs** Haus benötigt.	Clitic, phoneme
	Bahamas heißen die Inseln in der Karibik.	Phoneme
	Kartenspielen ist **was** für alte Leute.	Phoneme
	Dienstautos gibt **es** in vielen Firmen.	Plural, es
	Gipfelsteigen ist der letzte Teil **des Wanderns**.	Phoneme, genitive
	Packeis ist gefährlich.	Phoneme
	Kabel für **das** Internet findest du in Regal 4.	Phoneme
	Bäcker müssen **morgens** früh zu arbeiten beginnen.	Derivation
	Tests schreiben wir **dieses** Jahr keine.	Inflection
	Kameras in **Handys** werden immer besser.	Plural, plural
	Pass befindet sich in Tasche.	Phoneme
	Pilger wandern **jedes** Jahr von Frankreich nach Spanien am sogenannten Jakobsweg.	Inflection
	Tag für Tag **muss** ich das gleiche machen!	Phoneme
	Bäche gibt **es** viele in den Bergen.	Es
	Katzen essen manchmal **Gras**.	Phoneme
	Kellner müssen **bis** in die Nacht arbeiten.	Phoneme
	“Gammelfleisch” ist **altes** und **schlechtes** Fleisch.	Inflection, inflection

Eva Maria Luef, Ph.D. is an Associate Professor of English Linguistics. Her main research areas are psycholinguistics and phonology. She has researched multilingual language learning involving English and German as first, second, and third languages. Her research uses network analysis to study cognitive structures, such as the mental lexicon and phonological and semantic processing in lexical memory. She uses a combination of experimental methods from cognitive psychology and psycholinguistics, computational modeling and mathematical methods from network science, and large-scale analyses of linguistic databases and corpora.