Journal of Linguistics (2025), 1–33
doi:10.1017/S0022226725100777

Journal of Linguistics CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Acoustic disambiguation of homophonous morphs is exceptional

Ludger Paschen

Leibniz-Zentrum Allgemeine Sprachwissenschaft, Pariser Str. 1, 10719 Berlin, Germany
Email: paschen@leibniz-zas.de

Received: 14 July 2024; **Revised:** 18 April 2025; **Accepted:** 19 May 2025

Keywords: acoustics; endangered languages; homophones; morphology

Abstract
Homophonous morphs have been reported to show differences in acoustic duration in languages such as English and German. How common these differences are across languages, and what factors influence the extent of temporal differences, is still an open question, however. This paper investigates the role of morphological disambiguation in predicting the acoustic duration of homophones using data from a diverse sample of 37 languages. Results indicate a low overall contribution of morphological affiliation compared to other well-studied effects on duration such as speech rate and Final Lengthening. It is proposed that two factors increase the importance of homophony avoidance for the acoustic shape of morphs: crowdedness (i.e. the number of competing homophones) and segmental make-up, in particular the presence of an alveolar fricative. These findings offer an empirically broad perspective on the interplay between morphology and phonetics and align with the view of language as an adaptive and efficient system.

1. Introduction

The acoustic duration of segments is known to be influenced by a variety of factors, including speech rate, frequency, prosodic structure and phonological features (Klatt 1973, Gahl 2008, Fletcher 2010). The duration of segments in any given morph[1] can then be further influenced by morphological factors. A famous example is the case of segmentally homophonous morphs, which have been demonstrated to show subtle differences in acoustic duration.

One of the earliest reports of acoustic differences between homophones was the study by Walsh & Parker (1983), which compared the duration of plural /s/ and non-morphemic /s/ in

[1] Following the proposal in Haspelmath (2020), I will use the term *morph* to refer to roots, clitics and affixes as minimal linguistic forms. In the context of the language documentation data analyzed here, *morph* specifically refers to minimal segmentable units identified by field linguists and experts represented in the morphological annotations of recorded texts. I will reserve the term *morpheme* for abstract sets of forms with the same meaning or grammatical function.

three English homophone pairs (e.g. *lapse* vs. *lap-s*). Walsh & Parker (1983) found plural /s/ to be ca. 10ms longer than non-morphemic /s/. Lavoie (2002) observed differences in compressibility and range of realizations between the preposition *for* and the numeral *four* in that the preposition exhibits more extreme reduction possibilities than the numeral. Gahl (2008) reported a shorter duration for *time* compared to the homophonous noun *thyme*, which can be attributed to stark differences in lemma frequency. Drager (2011) noted systematic phonetic variation between three types of *like* in English (quotative, particle, lexical verb), and also differences between speakers and various social groups in the realization of these three lexemes. Plag et al. (2017) investigated the duration of various /s/ and /z/ morphs (corresponding to orthographic <s>) in English based on data from the Switchboard corpus (Godfrey & Edward 1997). The study compared a total of seven homophonous <s>: the third-person singular suffix, the plural suffix, the genitive singular enclitic, the genitive plural enclitic, the cliticized forms of *has* and *is*, and non-morphemic <s>. The authors observed that non-morphemic and plural <s> were substantially longer than the other <s> suffixes, while the cliticized variants of *has* and *is* had the shortest durations. In a related study, Seyfarth et al. (2017) found that morphologically complex English forms like *free-s* tend to have longer durations than homophonous freestanding forms like *freeze*, which the authors attribute to a paradigm uniformity effect matching the timing of articulatory gestures in *free-s* to that of the base form *free.*

For humans to communicate efficiently, there is an obvious tension between the benefit of the phonetic disambiguation of homophones and the increased effort of controlling the duration of individual morphs at a sub-phonemic level. It is an open question whether consistent acoustic differences between homophones exist across a large and diverse sample of languages while controlling for known effects such as frequency, pre-boundary lengthening, etc. The hypothesis I want to put forward here is that the case of the English <s> homophones examined by Plag et al. (2017) is extraordinary for a number of reasons, and that homophones that share certain properties with English <s> are more prone to showing acoustic differences compared to other homophones. In my view, there are three particular grounds on which English <s> homophones need to be considered exceptional. First, the <s> homophone set[2] in English is extremely crowded. Plag et al. (2017) discuss six different morphemes that all share the same phonological exponent or set of exponents, plus non-morphemic <s>. This extreme density is the result of phonetic convergence and morphological simplification over time, and cross-linguistically, it is rather uncommon for one morph to compete with five or more homophonous morphs. A sensible question to ask is whether acoustic differences between homophones are limited to, or more likely to occur in, highly crowded homophone sets, as compared to smaller groups with two or three homophones.

Second, the alveolar fricative /s/ is one of the most common and perceptually salient consonants in the world's languages (Maddieson & Disner 1984). What sets /s/ apart from most other fricatives is its high inter-speaker variability combined with a low within-speaker variability (Gordon et al. 2002, Smorenburg & Heeren 2020, Chodroff & Wilson 2022). /s/ requires fine motor control and tactile feedback and shows considerable coarticulatory resistance (Recasens 2018). Sociophonetic research suggests fine phonetic detail in /s/

[2] Groups of homophonous morphs with at least two distinct meanings are referred to as *homophone sets* throughout the paper.

https://doi.org/10.1017/S0022226725100777 Published online by Cambridge University Press

production and perception can be employed to mark social status (Stuart-Smith 2007), gender (Weirich & Simpson 2015) and sexual orientation (Linville 1998, Munson et al. 2006, Mack & Munson 2012, Nylén et al. 2024). Due to the sociophonetic versatility of /s/, it is plausible that morphs containing /s/ (or its voiced counterpart /z/) may also fulfill additional functions related to the acoustic disambiguation of homophones.

A third feature shared by all <s> morphs investigated by Plag et al. (2017) is their templatic homogeneity – that is, their morphological status as suffixes and enclitics occurring in the post-stem domain (or at the right edge of stems in the case of mono-morphemic <s>). This homogeneity exacerbates the confusability for the various <s> morphs, which raises the question whether homophonous morphs occupying different templatic positions (such as a prefix and a suffix) will show similar temporal differences. In a mixed group of homophonous morphs where one morph is a root and the other is an affix, one would also expect the root to be longer based on the fact that roots are phonologically stronger than affixes (Beckmann 1998). The present paper therefore seeks to answer two fundamental questions: First, what role does morphology play in accounting for durational differences between segmentally homophonous morphs across languages? And second, do homophonous morphs that are highly susceptible to acoustic disambiguation share certain properties with each other, and with the English <s> homophones?

The remainder of the paper is structured as follows. Section 2 provides an overview of previous literature on homophony and phonetic traces of morphology. Section 3 introduces the multilingual corpus used in this study and outlines the methods applied to assess the role of lexical identity in explaining temporal differences between homophonous morphs. Section 4 presents the results of the study and offers a discussion of potential factors affecting the likelihood by which homophone sets become susceptible to acoustic disambiguation. Section 5 discusses the results in light of grammatical theories and highlights potential avenues for follow-up research. Section 6 summarizes the main findings of the study.

2. Literature review

A standard definition of homophony requires that two phonologically identical forms have two *unrelated* meanings (1). Sometimes, the term *homophony* is used more loosely and extends to cases of polysemy, metonymy and syncretisms. All of these phenomena have in common that two phonologically identical forms differ in at least one semantic or grammatical detail. However, delineating these differences is not always straightforward, and the boundaries between these types of relations may depend on the particular framework or research question.

(1) Homophony

Two linguistic units have the same phonological form but two unrelated meanings.

Vicente & Falkum (2017) emphasize the relatedness criterion to distinguish homophony from polysemy, and critically evaluate the boundaries of reducing two meanings to one core meaning. Valera (2020) deals with processes resulting in the emergence of homophones, including sound change (phonological convergence), semantic change (semantic divergence), and borrowing. In typology, a key concept related to homophony is *colexification*,

https://doi.org/10.1017/S0022226725100777 Published online by Cambridge University Press

which refers to the phenomenon whereby a single lexical item expresses multiple distinct meanings across different contexts (Rzymski et al. 2020). Colexification is common across languages, and languages show certain preferences for expressing closely related meanings with the same linguistic form. Cross-linguistic tendencies in the co-expression of grammatical categories have similarly been analyzed in terms of semantic maps (Haspelmath 2003, Cysouw 2007). In the context of language documentation data, drawing a clear line between homophony and other semantic relations is particularly challenging when the amount and detail of available material is limited (Rice 2014). Moreover, the bulk of existing research on acoustic disambiguation focuses on homophones such as lexical homophones or homophonous grammatical markers, with few exceptions (Schlechtweg et al. 2020). For that reason, the present study will take a conservative stance and limit itself to homophonous elements that can with some degree of certainty be considered semantically unrelated.[3]

Even though homophony is a widespread phenomenon in human language, homophone sets are challenging for efficient communication because they introduce ambiguity that may lead to confusion. It is thus not surprising that languages show a certain dispreference towards homophones, which can manifest itself in various ways. Apart from subtle acoustic differences, homophony avoidance may operate on many levels, from affix selection in inflectional paradigms and conditions governing suppletive allomorphy to incomplete neutralization and inhibited sound change in phonology. Munteanu (2021) noted that Russian stressed case allomorphs -ú ‘-PREP.SG’ and -á ‘-NOM.PL’ tend to be selected by stems where a homophonous form with an unstressed case suffix occurs elsewhere in the inflectional paradigm. Goudswaard (2004) argued that homophony avoidance drives the choice between two suppletive infixes in Ida’an-Begak past tense paradigms. Goudswaard alleged that verbs starting with a consonant followed by /i/ select the *-ən-* past tense infix over the expected *-i-* infix because the latter would create an output that is homophonous with the base form. Within the framework of Optimality Theory, the empirical domain of homophony avoidance has seen proposals revolving around the REALIZEMORPHEME constraint (Kurisu 2001) or transderivational anti-faithfulness constraints (Benua 2000). Tomaschek & Berg (2023) reported that the contrast between German [ɐ] and [a] in unstressed syllables, which is generally assumed to be neutralized for Standard German, is in fact maintained in quasi-homophonous pairs such as *Opa* ‘grandpa’ and *Oper* ‘opera’ (but see Rathcke & Mooshammer 2022 for a similar study finding conflicting evidence). Blevins & Wedel (2009) critically examined cases of inhibition of sound change that would have resulted in the creation of homophony.

The acoustic duration of linguistic units is subject to various influences.[4] The *frequency* of words, and of linguistic units more generally, is known to affect their acoustic shape, with high-frequency items often being associated with phonetic reduction or phonological erosion (Zipf 1935, Bybee 2002, Pluymaekers et al. 2006, Strunk et al. 2020, Linders & Louwerse 2023). Frequency has been explicitly identified as a factor driving durational

[3] Aronoff (2017) discusses a curious argument put forward by Nida (1948) that claims a polysemous relationship between the third-person singular and plural *-s* in English. The argument goes that both affixes have a common abstract semantic demoninator ‘number distinctiveness’ and their concrete meaning can be inferred with reference to the grammatical features of the stems they attach to. This is perhaps one of the most extreme proposals aiming to stretch the boundaries of polysemy.

[4] Phonological length obviously also impacts the duration of segments, but this dimension is not relevant for the present study, as homophones by definition have segments of identical phonological quantity.

differences between lexical homophones (Gahl 2008). The *size* of words (in terms of phonological or morphological units) affects duration through two closely related effects known as *polysyllabic shortening* (Lehiste 1972) and *Menzerath's Law* (Stave et al. 2021), which both state that there is an inverse relation between the size of units and the duration of the respective sub-units. The *segmental context* within a word plays a key role for the duration of phones due to coarticulatory effects, most notably the fact that vowels tend to be longer before voiced than before voiceless consonants (Chen 1970, Beckman & Edwards 1980). The segmental context can be extended to word boundaries, which may also affect the duration of boundary-adjacent segments. Lengthening of word-initial consonants has been reported to be pervasive across languages (Blum et al. 2024), while lengthening of word-final vowels appears to be a language-specific process (White et al. 2020).

As far as larger prosodic domains are concerned, the *position* of words also plays a key role in determining acoustic duration. Most prominently, *Final Lengthening* increases segmental duration in up to two syllables before a major prosodic boundary (Turk & Shattuck-Hufnagel 2007, Fletcher 2010). Final Lengthening is attested in a wide variety of languages, and the lack of known counter-examples suggests it may be a (quasi-)universal process in human speech (Paschen et al. 2022). In addition, local *speech rate* is a natural source of variation in segmental duration: in stretches of speech when a speaker talks fast, the duration of phones, and by extension the duration of morphs, is shorter than when a speaker talks more slowly. Speech rate varies extensively across languages and individuals, and also between utterances from the same speaker (Coupé et al. 2019, Strunk et al. 2020, Tilsen & Tiede 2023). In general, variation across and within *speakers* is a strong influence on acoustic duration, and this variable also interacts with some of the other variables (for example, see Esposito 2020 on gender-related differences in the degree of Final Lengthening).

In addition to these factors, morphological structure has been observed to manifest itself acoustically in various ways. Zuraw et al. (2021) noted that the duration of onset voice onset time (VOT) depends on morphological segmentability in English prefixed and pseudo-prefixed words such as *dis-claimer* vs. *disco*, with long VOT indicating the presence of a strong boundary and transparent semantics. In a study on German syncretic singular and plural nouns, Schlechtweg et al. (2020) found plural forms to be longer than singular forms. Schmitz & Baer-Henney (2024) found subtle differences in duration between non-morphemic and suffixal word-final /s/ in German pseudowords. Regarding the perception of morpho-phonetics, Kemps et al. (2005) reported Dutch speakers to be sensitive to prosodic differences between singular nouns and their homophonous stem forms in plural nouns. Similarly, Schmitz (2022) showed that speakers of English are sensitive to durational mismatches in word-final /s/.

However, internal morphological structure and lexical identity are not always detectable in the acoustic signal. In a reading experiment with speakers of British English, Schlechtweg & Corbett (2021) did not find significant differences in the duration of plural *-s* in regular plural nouns compared to pluralia tantum nouns. Wu (2017) investigated various factors influencing word duration in Chinese and reports no significant effect of the internal structure of compounds on duration. Morrison (2021), in a production experiment measuring nasal airflow, found no support for traces of underlying nasal features in Scottish Gaelic mutated consonants. For most of the world's 7000 languages, no data on the interplay between morphology and phonetics exist to date. It is thus still an open question just how common phonetic reflections of morphological structure truly are across languages, and

https://doi.org/10.1017/S0022226725100777 Published online by Cambridge University Press

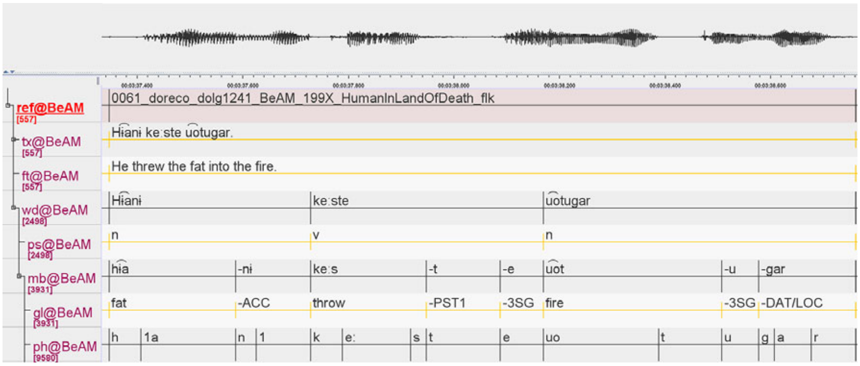

Figure 1. Time alignment at the utterance, word, morph and phone levels in a recording from the DoReCo Dolgan dataset (Däbritz et al. 2024, ELAN editor view). The ‘tx’, ‘wd’, ‘mb’, ‘gl’ and ‘ph’ tiers contain the time-aligned utterances, words, morphs, glosses and phones, respectively.

whether acoustic disambiguation of homophones is the norm or an exception found only in specific languages or only with a certain class of homophones.

3. Methods

3.1. Data

Data for this study were taken from the DoReCo corpus version 2.0 (Seifart et al. 2024).[5] DoReCo 2.0 contains time-aligned data from 53 languages, of which 39 contain morphological tokenization (also referred to as *morpheme breaks*) and interlinear glossing (Figure 1), of which 37 are included in this study.[6] Table 1 lists these 37 languages, and Figure 2 shows the approximate geographic locations where the languages are spoken. The 37 datasets combined had a size of 340,000 word tokens.[7]

DoReCo was created with the goal of mobilizing language documentation data from public archives, and in some cases, private collections, for cross-linguistic research and to provide a resource to the humanities for carrying out empirical work beyond the scope of WEIRD (Western, Educated, Industrialized, Rich, Democratic) languages and speakers (Henrich et al. 2010, Blasi et al. 2022). Data in DoReCo consist of transcribed – and in many cases, also

[5] All scripts used in this study are available at https://github.com/LuPaschen/homophonous_morphs.

[6] Two datasets were excluded from this study. First, Komnzo (komn1238) features extensive multiple exponence that makes it impossible to define a fixed meaning or function for most affixes (Döhler 2018). Second, the English dataset represents a variety from Southern England (Kent, sout3282) and is written according to standard British English orthographic conventions. This is in stark contrast to the other datasets in DoReCo, which follow practical orthographies developed by field workers closely matching the (surface) phonological representation of words. This means that segmentally but not graphematically homophonous words such as *four* and *for* would not be correctly classified as homophones with the methods used in this study.

[7] As with other curated corpora of endangered languages, datasets in DoReCo have a lower word token count than most well-established corpora of English and other major languages. This should be taken into account when interpreting the results of quantitative studies based on language documentation data.

https://doi.org/10.1017/S0022226725100777 Published online by Cambridge University Press

Table 1. DoReCo datasets with word-level and morph-level time alignment included in this study. Words = word tokens in the DoReCo core datasets. Homophones = distinct homophone sets after filtering included in the final models

	Language	Glottocode	Family	Words	Speakers	Homophones	Source
1.	Arapaho	arap1274	Algic	4740	4	1	Cowell (2024)
2.	Baïnounk Gubëëher	bain1259	Atlantic-Congo	11432	9	12	Cobbinah (2024)
3.	Beja	beja1238	Afro-Asiatic	15430	5	8	Vanhove (2024)
4.	Bora	bora1263	Boran	8384	6	11	Seifart (2024)
5.	Cabécar	cabe1245	Chibchan	10616	10	6	Quesada et al. (2024)
6.	Cashinahua	cash1254	Pano-Tacanan	9666	3	6	Reiter (2024)
7.	Daakie	port1286	Austronesian	11883	10	8	Krifka (2024)
8.	Dalabon	ngal1292	Gunwinyguan	3415	4	2	Ponsonnet (2024)
9.	Dolgan	dolg1241	Turkic	8781	6	6	Däbritz et al. (2024)
10.	Evenki	even1259	Tungusic	8320	23	4	Kazakevich & Klyachko (2024)
11.	Fanbyak	orko1234	Austronesian	10167	10	10	Franjieh (2024)
12.	Goemai	goem1240	Afro-Asiatic	9340	5	13	Hellwig (2024)
13.	Gorwaa	goro1270	Afro-Asiatic	10670	8	8	Harvey (2024)
14.	Hoocąk	hoch1243	Siouan	7376	11	6	Hartmann (2024)
15.	Jahai	jeha1242	Austroasiatic	7631	3	2	Burenhult (2024)
16.	Jejuan	jeju1234	Koreanic	7276	5	9	Kim (2024)
17.	Kakabe	kaka1265	Mande	9529	4	8	Vydrina (2024)
18.	Kamas	kama1351	Uralic	10058	1	6	Gusev et al. (2024)
19.	Mojeño Trinitario	trin1278	Arawakan	7749	6	12	Rose (2024)
20.	Movima	movi1243	(isolate)	10186	5	2	Haude (2024)
21.	Nafsan	sout2856	Austronesian	10243	12	8	Thieberger (2024)
22.	Nisvai	nisv1234	Austronesian	11898	8	16	Aznar (2024)
23.	N‖ng	nngg1234	Tuu	10282	6	11	Güldemann et al. (2024)
24.	Northern Alta	nort2875	Austronesian	8558	2	9	Garcia-Laguia (2024)
25.	Northern Kurdish	nort2641	Indo-European	9657	2	11	Haig et al. (2024)

Table 1. *Continued*

	Language	Glottocode	Family	Words	Speakers	Homophones	Source
26.	Pnar	pnar1238	Austroasiatic	8545	6	7	Ring (2024)
27.	Ruuli	ruul1235	Atlantic-Congo	8250	7	7	Witzlack-Makarevich et al. (2024)
28.	Sanzhi Dargwa	sanz1248	Nakh-Daghestanian	5142	5	8	Forker Schiborr (2024)
29.	Savosavo	savo1255	(isolate)	11381	7	12	Wegener (2024)
30.	Sümi	sumi1235	Sino-Tibetan	7759	23	10	Teo (2024)
31.	Tabasaran	taba1259	Nakh-Daghestanian	5060	2	13	Bogomolova et al. (2024)
32.	Teop	teop1238	Austronesian	12136	11	16	Mosel (2024)
33.	Texistepec Popoluca	texi1237	Mixe-Zoque	8468	1	6	Wichmann (2024)
34.	Totoli	toto1304	Austronesian	9756	12	13	Bardají et al. (2024)
35.	Urum	urum1249	Turkic	10039	30	7	Skopeteas et al. (2024)
36.	Vera'a	vera1241	Austronesian	12383	7	12	Schnell (2024)
37.	Yali	apah1238	Nuclear TNG	7477	10	13	Riesberg (2024)
	Total			339,683	287	319	

morphologically annotated – recordings of spontaneous speech, predominantly personal and traditional narratives. Most datasets contain around 10,000 word tokens. The DoReCo workflow involves a number of manual and automatic processing steps, including manual labeling of disfluencies and computer-assisted creation of word and phone alignments using the WebMAUS service (Kisler et al. 2012). The morphs and glosses are then aligned with the word and phone units in a separate step using the 'reinjection' pipeline built on the Needleman-Wunsch algorithm (Stave et al. 2019).[8] For more details on the DoReCo workflow, see Paschen et al. (2020).

3.2. Data processing

Tabular CSV files from 37 DoReCo core datasets were processed in R (R Core Team 2023) using RStudio (RStudio Team 2023). Several additional columns were added for the analysis in Section 3.3: `morph_duration` was defined as the sum of the duration of the phones contained in any given morph. `word_size` was calculated as the number of phones per word minus the number of phones of the respective morph. Values for `speech_rate` were calculated as the number of phones per inter-pausal unit (IPU) divided by the duration of the IPU. Following the suggestion in Tilsen & Tiede (2023), the number of phones in the unit of interest (here, the morph) and their durations were subtracted before calculating speech rate to avoid the auto-correlation bias. As speech rate cannot be reliably calculated for IPU's consisting of a single monomorphemic word, these tokens were excluded from analysis (2.6%). `word_frequency` was defined as the number of occurrences of any given word +gloss(es) combination per language. The variable `segmental_context` encodes whether the morph was word-initial or word-final, and whether the following segment (if any) was a vowel, a voiced consonant or a voiceless consonant. Lastly, each morph was tagged for whether it occurred in IPU-final `position` to control for Final Lengthening.

Three additional factors were extracted for post-hoc analyses. The factor *presence of an alveolar fricative* codes whether a morph contains an alveolar fricative (<s> or <z>). The factor *crowdedness* contains information on the number of competing meanings in a given set of homophones. The factor *homogeneity* encodes whether all morphs in a homophone set belong to the same morph type (root, prefix, suffix, proclitic, enclitic). These three factors are not included in the statistical models because each homophone set is evaluated by a separate model, and these factors are uniform within each homophone set.

The data were further processed to unify common notational variants for glosses on a per-language basis. This step was necessary to correctly group together morphs for which annotators had used different notational variants, and to ensure homophone sets contain morphs with unrelated meanings. For example, in the Kamas dataset (Gusev et al. 2024), the adverbial locative marker *-n* is variously glossed as '-ADV.LOC', '-LOC.ADV' or '-LOCADV', which all obviously encode the same grammatical category. In the Yali (Apahapsili) dataset (Riesberg 2024), the lexical item *tag* is sometimes glossed as 'bad' and sometimes as 'not. good'. Variation in glossing due to polysemy or metonymy was also accounted for, and pairs of glosses forming those semantic relations were considered equivalent. For example, in the Cabécar dataset (Quesada et al. 2024), both the glosses 'water' and 'river' are used for *diglö* 'water, body of water'; the glosses were chosen by the annotators depending on which

[8] The code is available at https://github.com/DoReCo/reinjection.

https://doi.org/10.1017/S0022226725100777 Published online by Cambridge University Press

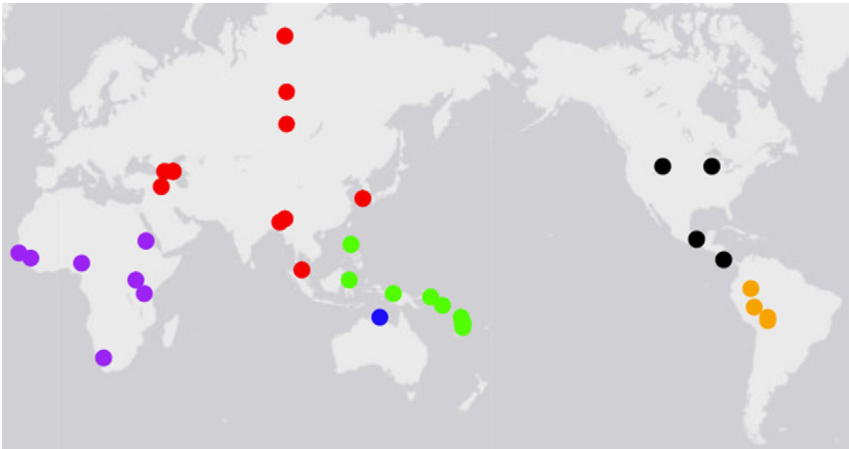

Figure 2. Map showing the locations of the 37 languages included in the sample. Color coding indicates macro-area. Image created with the `lingtypology` package for R (Moroz 2017).

meaning was suitable in a given context. Further harmonization was applied to glosses that were overspecified for grammatical categories, such as the second-person absolutive agreement prefix *dja-* in the Dalabon dataset (Ponsonnet 2024), which is glossed '2SG-' in intransitive but '1SG>SG-' or '3SG>2SG-' in transitive sentences. Lastly, glosses for lexical roots with flexible parts of speech were unified if necessary, such as the glosses 'female' and 'woman' for *hinųk* in the Hoocąk dataset (Hartmann 2024), or the glosses 'beauty' and 'beautiful' for *d'ong* in the Goemai dataset (Hellwig 2024).[9]

Before running the inferential models, all words labeled as disfluencies or code-switching, and other unusual speech events were excluded (2.3%), as well as morphs with generic filler glosses indicating unknown content (8.2%). Morphs with less than 10 observations were also excluded (28.6%). Lastly, only morphs with at least two distinct meanings (i.e. homophones) were included in the models (removing 83.4% of the remaining data). The final dataset on which the statistical models were run contained a total of 164,162 morphs.

3.3. Model fitting

For each homophone set in every language, a Random Forest model was run using the `randomForest` package in R (Liaw & Wiener 2002). The goal was to compare the relative importance of seven variables (2) for predicting `morph_duration`.[10] Random Forests are well-suited for assessing the relative contribution of variables in predicting a target variable, especially when dealing with smaller datasets and limited amounts of data, which was the

[9] These equivalence relations are not to be interpreted as general statements about polysemy or colexification in the respective languages. Rather, the justification for these groupings is that there is a non-empty set of common morphs shared by two or more glosses in a given dataset and that these meanings or grammatical functions are not considered sufficiently independent for the purpose of the present study.

[10] The models do not include variables controlling for linguistic family or area because homophone sets are created on a per-language basis, and each homophone set is evaluated by a separate model.

https://doi.org/10.1017/S0022226725100777 Published online by Cambridge University Press

case for some homophone sets (Levshina 2021).[11] Random Forests have been successfully applied to linguistic data to assess the relative contribution of variables (Levshina 2016, Winter & Perlman 2021). Each model was run with 500 individual decision trees and three randomly selected features (mtry = 3). The models tested the importance of seven variables for predicting `morph_duration`, with the main variable of interest being `morph`, representing the morphological identity of a morph, identified by the glosses annotated in the corpus. A high importance value for the `morph` variable implies acoustic differences between homophonous morphs are primarily due to them being homophones, rather than due to other factors such as position or frequency. Alternative models with a different arrangement of variables are presented in an Appendix at the end of the paper.

(2) (a) `morph`
(b) `word_frequency`
(c) `segmental_context`
(d) `word_size`
(e) `speech_rate`
(f) `position`
(g) `speaker`

For each Random Forest model, variable increase in node purity was calculated. Models below a performance threshold $R^2 < 0.25$ were not included in the analysis in Section 4, removing 60.4% of the remaining data. It is important to clarify that the analysis is based on the roots, clitics and affixes annotated on the morpheme break tier in the DoReCo corpus, which are usually oriented towards surface-level forms. The analysis does not conflate allomorphs into groups of more abstract morphemes. To give a concrete example from Sanzhi Dargwa, one of the languages included in the sample, the oblique and imperative suffixes *-a* would be part of the same homophone set, and the oblique and habitual past suffixes *-i* would be part of a different homophone set, but there would be no abstract set combining the *-a* and *-i* allomorphs of the oblique morpheme. The present analysis is concerned with 1:1 comparisons of morphs and does not compare complex words with homophonous freestanding words (such as English *free-s* and *freeze*), or non-morphemic root parts with homophonous affixes (such as English plural *-s* with the final /s/ in *lapse*).

4. Results

4.1. Variable importance

In this section, a global view of the importance of the `morph` variable compared to other variables for predicting morph duration is presented. Then, variable importance is compared across various data subsets to assess whether homophone sets that share certain properties with English <s> homophones are more prone to acoustic disambiguation.

Figure 3 shows the average relative contributions of the seven factors across all languages and homophone sets. By far, the most important factor was local `speech_rate` at 34.3%.

[11] The smallest homophone set in the study consisted of two members with a total of 20 observations, the largest contained four homophones with a total of 2294 observations. The average homophone set contained 179 morphs.

https://doi.org/10.1017/S0022226725100777 Published online by Cambridge University Press

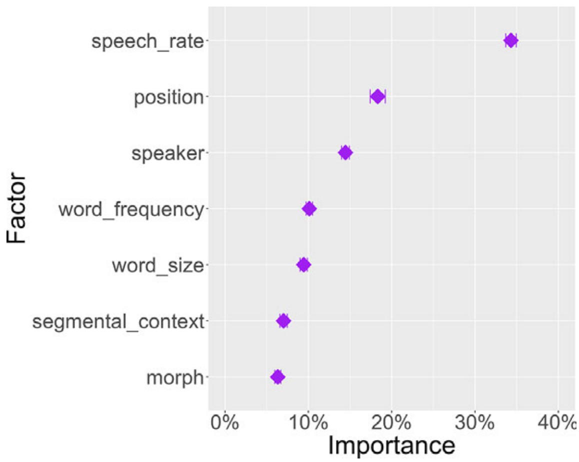

Figure 3. Average contribution of seven variables for predicting morph duration.

The second most important factor was `position` at 18.3%, reflecting the well-known process of Final Lengthening. The third most important variable was `speaker` at 14.5%. Note that the importance of this variable may be underappreciated because some DoReCo datasets contain data from 20 or more speakers, while others contain recordings from only a single speaker. In an ideal scenario with a perfectly balanced corpus, the relative contribution of `speaker` would possibly be higher. The fourth most important variable was `word_frequency` at 10.1%. The factor `word_size` ranked fifth with a contribution of 9.4%.

The `segmental_context` variable had a rather low overall influence on morph duration (7.0%). The low importance of this variable is likely linked to the fact that coarticulation and word boundary effects apply locally, meaning that the significance of segmental context decreases with larger morph sizes. Furthermore, the extent of segmental effects depends heavily on the phonetic and phonological properties of the individual languages. The morphological affiliation of a morph (i.e. the `morph` variable) had the lowest overall variable importance, with an average contribution of 6.3%. On a general population level, the morphological identity thus only plays a minor role for the acoustic duration of morphs compared to other factors. However, the relative importance of `morph` differs considerably across languages (Figure 4) and also between individual homophone sets (see Section 4.2). The three languages with the highest average scores in Figure 4 (Arapaho, Kamas, Texistepec Popoluca) are all highly synthetic. Texistepec Popoluca also has the homophone set with the highest overall `morph` contribution score, which will be discussed in more detail in Section 4.2.

Turning now to the question if homophone sets that resemble the English <s> homophones imply a stronger importance of `morph`, three independent factors were explored: crowdedness, presence of an alveolar fricative, and homogeneity. The first, *crowdedness*, indeed affected the importance of the `morph` variable. Figure 5 illustrates the relation between the size of homophone sets and the weight of the `morph` variable across all languages in the sample. The overlaid regression line reveals a positive and significant correlation between these two dimensions ($R = 0.12$, $p < 0.05$). This can be interpreted such that the more competing homophonous morphs exist in a language, the more their durations are clustered around individual morphs. Note, however, that the number of homophone sets

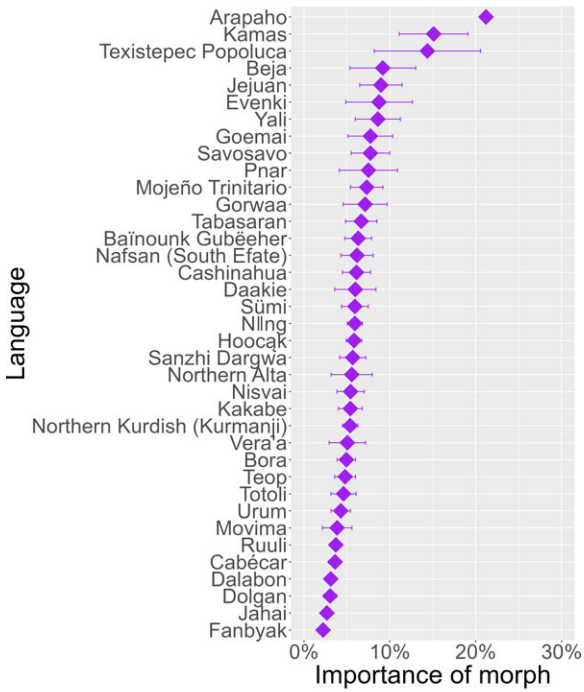

Figure 4. Average contribution of the `morph` variable across 37 languages.

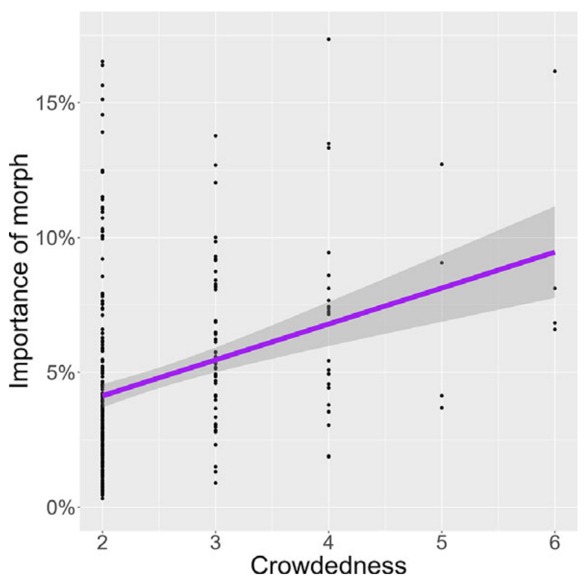

Figure 5. Correlation between crowdedness (homophone set size) and average importance of the `morph` variable.

https://doi.org/10.1017/S0022226725100777 Published online by Cambridge University Press

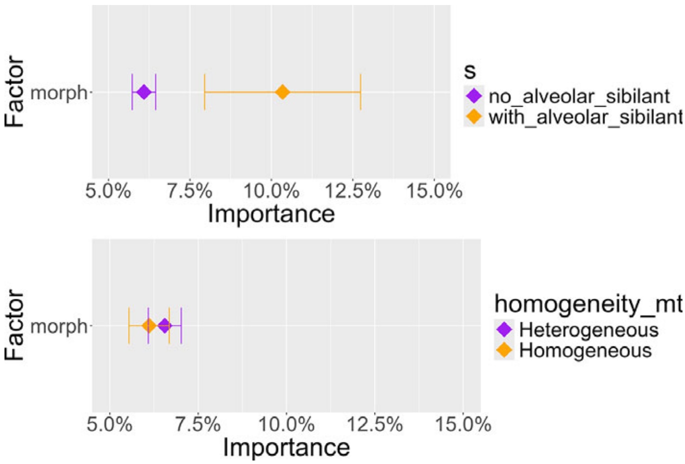

Figure 6. Variable importance depending on presence of an alveolar sibilant (top) and homogeneity of morph type (bottom).

is not equally distributed over the size categories. The majority of homophone sets had two (227/319) members, and sets with three (54/319), four (27/319) or more (11/319) members were increasingly less frequent.

To investigate whether morphs containing an alveolar sibilant are better candidates for the acoustic disambiguation of homophone sets, the data were split into two subsets, one including morphs with at least one alveolar sibilant (in any position), and one including the other morphs. To test whether homophone sets containing only a certain type of morph (roots, suffixes, prefixes, enclitics, proclitics, infixes) are more prone to acoustic differentiation than mixed sets, the data were again split into two subsets, one containing homogeneous homophone sets and the other containing heterogeneous sets. Figure 6 shows the importance of the `morph` variable across these four subsets. First, there is a clear difference in the average variable importance between homophones with and without an alveolar fricative. Homophones with an alveolar sibilant (orange) had notably lower average importance for `morph` than homophones without such a fricative (purple), with average scores being 10.3% and 6.0%, respectively. These results should be treated with some caution, though, as the two sets were not equal in size: homophone sets containing an alveolar fricative only made up 5.6% of homophone sets included in the final models. Second, the average `morph` importance scores for homogeneous and heterogeneous sets were very similar, with values of 6.1% and 6.5%, respectively. The two sets were of almost identical size, with 50.7% being homogeneous.

The data thus support the initial hypothesis regarding homophones with an alveolar sibilant bearing higher importance for `morph`, but results are inconclusive regarding the homogeneity dimension.

4.2. Homophone sets with high `morph` *importance*

The quantitative analysis has revealed an overall low importance of morphological identity for the acoustic duration of homophones but has also identified factors related to crowdedness and

https://doi.org/10.1017/S0022226725100777 Published online by Cambridge University Press

segmental properties that appear to increase the likelihood of morphological disambiguation. As shown in Figure 4, there are considerable differences in the contribution of the `morph` variable across languages. These differences ultimately boil down to differences between individual homophone sets, some of which display exceptionally high scores for the `morph` variable. Table 2 lists the 25 homophone sets with the highest `morph` contributions, ranging from 19% to 41%. A crucial observation is that there is no single language, linguistic family, or area dominating this list. Likewise, the size and shape of the morphs are highly varied, ranging in size from a single segment (such as Northern Alta *o*) to polysyllabic forms (such as Nisvai *likanim*). Consequently, there do not appear to be any particular macro-level features interacting with the `morph` variable.

In the following, five homophone sets will be discussed in more detail, highlighting additional factors that may contribute to the relatively high variable importance for those homophone sets.

Table 2. Top 25 homophone sets with the highest relative importance of the `morph` variable. Meanings = distinct meanings (glosses) after filtering. Tokens = number of morphs in homophone sets evaluated by the statistical models

	Language	Morph	Meanings	Tokens	Importance
1.	Texistepec Popoluca	wää	2	63	41%
2.	Beja	ji	4	62	33%
3.	Yali	su	2	188	32%
4.	Kamas	ku	2	76	31%
5.	Savosavo	sua	2	248	30%
6.	Goemai	d'a	2	140	30%
7.	Nisvai	likanim	2	27	27%
8.	Goemai	hok	2	171	26%
9.	Tabasaran	š	2	66	26%
10.	Vera'a	su	2	88	26%
11.	Yali	tu	2	275	24%
12.	Mojeño Trinitario	o'i	2	112	23%
13.	Texistepec Popoluca	'	6	155	22%
14.	Jejuan	t^ha	2	39	22%
15.	Northern Alta	o	2	53	22%
16.	Daakie	te	2	58	22%
17.	Arapaho	nihii	2	45	21%
18.	Pnar	kam	3	59	20%
19.	Jejuan	tɕi	3	64	20%
20.	Gorwaa	ee	3	304	20%
21.	Kamas	kuʔ	2	24	20%
22.	Totoli	s	2	33	20%
23.	Evenki	a	2	47	19%
24.	Pnar	nong	2	20	19%
25.	Baïnounk Gubëeher	d	3	99	19%

https://doi.org/10.1017/S0022226725100777 Published online by Cambridge University Press

4.2.1. Texistepec Popoluca wää

The homophone set *wää* in Texistepec Popoluca consists of the adjective/adverb ‘good, well’ and the auxiliary verb ‘can(AUX)’.[12] Both the adjective/adverb and the auxiliary *wää* usually occur on their own, but may occasionally also appear within complex words: the auxiliary may be accompanied by the plural marker *-be* or the perfective clitic =*am*, while the adjective is attested in DoReCo within certain compound-like constructions such as *wää-dä’* ‘(good) water’. The examples in (3) and (4) show two typical contexts in which these forms are used.

(3) maas **wää** y<y>oomä’-baa
more[Sp] **good** <3:A>woman-DIM
‘His daughter is better.’
[Wichmann 2024, doreco_texi1237_JuanFlojo_part4]

(4) yoomä’ ’eñch **wää** ny-daks d<y>äm
woman no **can(AUX)** 2/3-hit <3:A>say
‘You can’t hit the woman, he says.’
[Wichmann 2024, doreco_texi1237_LaMujerFloja]

For *wää* ‘good, well’, 38 tokens with an average duration of 261ms are included in the models, while *wää* ‘can(AUX)’ is attested 25 times with a mean duration of 161ms (Table 3). A plausible explanation for the differences in duration is that content words tend to be generally more phonetically robust than function words. Moreover, evaluative adjectives such as *wää* ‘good, well’ can be prosodically prominent in emphatic contexts (Selting 1994). As the corpus data used here contain spontaneous speech, it is likely that emphatic prosody is used widely, and the marked temporal differences are due to a combination of word class and prosodic factors.

4.2.2. Beja ji

The CV string *ji* in Beja can be a genitive suffix ‘-GEN’, a relative enclitic ‘=REL’, a relative plural masculine proclitic ‘REL.PL.M=’, a first-person singular nominative/accusative possessive marker ‘=POSS.1SG.NOM/ACC’, an aorist tense suffix ‘-AOR.3SG.M’, a copula enclitic ‘=COP.3SG’, and an emphatic clitic ‘=EMPH’ (Table 4).[13] Only the case suffix, the possessive

Table 3. Texistepec Popoluca *wää* homophone set with average morph durations

wää	‘good, well’	261ms	wää	‘can(AUX)’	161ms

[12] The original annotations for this dataset are in Spanish and are translated into English here.
[13] There is also one instance of an aorist stem *ji* glossed ‘say\AOR.3SG.M’ in the corpus, but this gloss combination is not attested in Vanhove (2017), so its status remains spurious and it was not included in the table. The gloss POSS.1SG.NOM/ACC is the unification of the glosses POSS.1SG.NOM, POSS.1SG.ACC and POSS.1SG. All glosses and translations in the examples shown here have been translated from French into English.

Table 4. Beja *ji* homophone set with average morph durations. Grey shading indicates morphs not included in the models due to low token frequency

-ji	-GEN	76ms	=ji	=POSS.1SG.NOM/ACC	71ms
ji=	REL.PL.M=	236ms	=ji	=REL	107ms
-ji	-AOR.3SG.M		=ji	=COP.3SG	
			=ji	=EMPH	

marker, and the two relative clitics are included in the models, as the other three were not frequent enough in the corpus. (5)–(7) show examples of the genitive suffix, the relative enclitic and the copula marker.

The Beja *ji* homophone set is an exceptionally crowded set, similar to English <s>. In addition, there are a number of near-homophonous morphs in Beja, including the two case clitics =*jiː* 'ABL.SG' and =*jiːb* 'LOC.SG'. The situation is further complicated by the fact that all *ji* morphemes except the relative proclitic attest a shorter *i* allomorph that lacks the onglide. As [i], [ji] and [jiː] are perceptually close, the perceptual space in the post-stem domain can be assumed to be even more crowded than the list of morphs in Table 4 suggests. Beja *ji* thus presents a case of an extremely crowded 'extended' homophone set, with an exceptionally high number of acoustically similar competing morphs, likely representing a typological rarity.

(5) hinin toː=tji t=asarama girma-**ji**
1PL.NOM DEF.SG.F.ACC=serpent DEF.F=seven head-**GEN**
dir-a=b ni-kati=eːb
kill-CVB.MNR=INDF.M.ACC 1PL-become\INAC=REL.M
'(We swear) that we killed the seven-headed serpent!' [Vanhove 2017: 181]

(6) baːbija dibi-a=b tiː-kti=**ji**
PTCL keep-CVB.MNR=INDF.M.ACC 3SG.F-become\AOR=**REL**
'If only you had kept it!' [Vanhove 2017: 181]

(7) haːl=oːkna=**ji**
state=POSS.2PL.ACC=**COP.3SG**
'What happened to you? (lit. it is your state)' [Vanhove 2017: 206]

4.2.3. Vera'a su

In the Vera'a DoReCo dataset, the morph *su* occurs in two contexts: as a reduplicative prefix *su~* and as a lexical root *su* 'to/a paddle' (Table 5). The latter is a common allomorph of *suō*; the exact conditions governing the *su/suō* variation are not entirely clear. The reduplicant

Table 5. Vera'a *su* homophone set

su(ō)	'paddle'	197ms	su~	RED~	114ms

https://doi.org/10.1017/S0022226725100777 Published online by Cambridge University Press

su~ mostly appears before *su(ō)* in the corpus, but cases of *su~* occurring before other bases such as *sur* 'knit' are also mentioned in the literature (Schnell 2011: 202).

Examples (8) through (10) show typical usage examples for *su* and *su~* found in the Vera'a DoReCo dataset. All three examples illustrate reduplicative *su~* before *su(ō)* 'paddle' in the same word form. Example (9) showcases *su~* occurring in the frequent construction *nak su~suō* 'canoe RED~paddle' = 'canoe (for paddling)' (cf. Schnell 2011: 71). (10) illustrates an instance of multiple reduplication involving two consecutive *su~*.

(8) dir=k **su~su** sar ma so=k 'ēn ~'ēn enteg sir
3PL=TAM2 **RED~paddle** inland hither NC=TAM2 RED~see well for
'and as they were paddling up in order to have a good look at (him) […]'
[Schnell 2024, doreco_vera1241_veraa_jjq]

(9) dir'ōl=k rēv sur ēn nak **su~suō**
3TL=TAM2 drag down ART canoe **RED~paddle**
'The three dragged down a canoe.'
[Schnell 2024, doreco_vera1241_veraa_palaa]

(10) duru=m **su~su~suō** **suō** di sar ēn ēre'ie
3DL=TAM1 **RED~RED~paddle** **paddle** reach bushwards ART coast
'The two were paddling, paddled ashore […]'
[Schnell 2024, doreco_vera1241_veraa_hhak]

Reduplication inevitably leads to (partial) homophony between base and reduplicant. What is special about this type of homophony is that the two homophones are bound to the same word form. Differences in duration between such homophones likely result from a prosodic asymmetry between base and reduplicant. It should be noted, though, that while about half of the 37 languages in the present sample attest productive reduplication, and reduplicants occur in a total of 16 homophone sets (after filtering), only two homophone sets (Vera'a *su* and Totoli *s*) out of the 25 homophone sets with the highest morph importance (Table 2) contain a reduplicant. Phonetic differences between base and reduplicant thus seem to only marginally affect the overall importance of `morph`.

4.2.4. Arapaho nihii

Arapaho *nihii* is ambiguous between a verbal root 'say' and a discourse particle glossed as 'well…' (Table 6). Both morphs have a low tone in the first and a high tone in the second syllable. Verbal *nihíí* is an intransitive stem that takes an animate subject. There are two formally similar and semantically related applicative verb stems, although they do not seem to be derived from *nihíí* through synchronically productive means: *nihii*θ 'say/tell to s.o.' and *nihiit* 'say sth.'. The discourse marker *nihíí* belongs to a closed set of 'substitutionary or pausal particle[s]' and is found in contexts of doubt or uncertainty (Cowell & Moss Sr. 2008: 449).

Table 6. Arapaho *nihii* homophone set

nihii	'say'	288ms	nihii	'well…'	484ms

https://doi.org/10.1017/S0022226725100777 Published online by Cambridge University Press

Table 7. Banounk Gubëeher *d* homophone set

-d	-NEG:PERF	38ms	-d	-BEN	32ms
d-	NEG:FUT-	65ms			

(11) *kóóxwoow heetkóókohtowúnihíí*
koox=woow e-eti-koon-kohtowu-**nihii**
again=now 2S-FUT-REDUP-funny/inappropriate-**say(AI)**
'I suppose you're going to say funny things now.'
[Cowell & Moss Sr. 2008: 199]

(12) *'oh howóó niiyóúno nihíí núhu' nihíí yóókoxuu.*
'oh owoo niiyou-no **nihii** nuhu' **nihii** yookox-ii
and also here.is-PL **well** this **well** willow-PL
'And also there were, well, these, well, willows.'
[Cowell & Moss Sr. 2008: 316]

The difference in duration between the two *nihíí* homophones (ca. 200ms on average) is substantial, and can at least be partially attributed to word size: verbal *nihíí* tends to occur within complex polymorphemic word forms whereas the particle *nihíí* only appears in isolation, cf. (11) and (12). As `word_size` is included as a factor in the models, the fact that morphological identity still substantially contributes to predict the duration of these morphs remains somewhat mysterious. The sheer magnitude of durational differences may have boosted the importance of the `morph` variable. Another potential explanation could be the floating H tone of verbal *nihíí*, which usually attaches to the preceding tone-bearing unit and thus influences the tonal properties of the whole stem, potentially leading to systematic differences in acoustic duration.[14]

4.2.5. Banounk Gubëeher d

The Baïnounk Gubëeher *d* homophone set has three members: the negative perfective/irrealis suffix *-d* '-NEG:PERF', the negative future/habitual prefix *d-* 'NEG:FUT-', and the benefactive suffix *-d* '-BEN'. These are also the only three *d* morphs attested in the Baïnounk Gubëeher DoReCo dataset (Table 7). Both the benefactive suffix and the negative perfective suffix occur 20 times in the data sample, while the negative future prefix is attested 59 times.

(13) and (14) provide examples of the two negative markers, and (15) illustrates the use of the benefactive suffix *-d.*

(13) me g-u-cil-ex indan bi-na'-hen **d-**ë-bun
1SG FOC.OBJ–2-laugh-HAB go:IMP CL.bi-way-2SGPOSS **NEG:FUT-**3-good
'You are laughing about me? Go, your journey will not be successful.'
[Cobbinah 2013: 244]

[14] Thank you to Nicholas Rolle for bringing the tonal properties of verbal *nihíí* to my attention.

https://doi.org/10.1017/S0022226725100777 Published online by Cambridge University Press

(14) me wul-aa-**d**-i (< wul-aar-**r**-i)
1SG see-ALREADY-**NEG:PERF**-1SGSUBJ
'I have never seen that!' [Cobbinah 2013: 245]

(15) a-yen-eneŋ a-n-dú-**d**-ët-om
3-say-3PL.OBJ 3-PL-draw.water-**BEN**-VEN–1SG.OBJ
'They told them to fetch them some.'
[Cobbinah 2024, doreco_bain1259_DJI070211AC]

A crucial observation about the two negative affixes is that the future prefix is underlyingly /d/ but the negative perfective suffix is underlyingly /r/, with a predictable allophone [d] after stems ending in /n/ and /r/ (Cobbinah 2013: 231). This allophony is instantiated by two regular phonological processes in the language: fortition of /r/ to /d/ after /n/ and /r/, and subsequent deletion of the first /r/ (Cobbinah 2013: 173). It is plausible that the negative perfective *-d* shows a slightly longer duration because it still retains traces of the two underlying segments it is derived from, even though a difference of 7ms is unlikely to be perceptually relevant. The longer duration of the negative future prefix *d-*, which often occurs word-initially in Baïnounk Gubëeher (13), may be linked to the process of word-initial lengthening (Blum et al. 2024).

5. Discussion

5.1. General discussion

Homophony is a common phenomenon in the world's languages, and in the present study, this is reflected by the fact that even after applying several layers of filters, languages in the 37-language sample had an average of eight to nine distinct homophone sets (Table 1). Once a broad cross-linguistic perspective is adopted, the role of morphological identity in explaining acoustic differences between homophones turns out to be far less impactful compared to previous reports based on languages such as English or German. It appears that most homophone sets in most languages do not pose any communicative challenge and therefore do not prompt any sub-phonemic durational cues to their lexical identity. And indeed, it is difficult to construe a scenario in which an interlocutor is unable to infer from the communicative context whether a speaker intended to say *time* or *thyme*, for instance.

Previous studies have pointed out that acoustic differences between certain homophones can readily be explained by frequency or prosodic factors, prompting the question how meaningful morphological identity is in shaping the acoustic form of homophones. While prosodic factors (speech rate and position) and word frequency had a consistently high impact on the duration of morphs in the present study, the strength of the relative contribution of morphological identity varied considerably across homophone sets. Three properties of homophone sets were tested specifically as to whether they made homophonous morphs more susceptible to acoustic disambiguation.[15]

[15] The perceptual threshold for durational differences in fricatives has been reported to lie at around 30ms (Plag et al. 2017), but the perception of temporal differences also depends on individual factors such as age (Price & Simon 1984). Any differences in segmental duration above 30ms can be considered to be perceptible and meaningful to listeners, while no definitive statement can be made regarding smaller differences.

First, the *crowdedness* of the homophone space was found to correlate with the relative importance of morphological identity. Homophone sets with more members showed a tendency to display higher morphological contribution scores than homophone sets with fewer members. The majority of homophone sets had two or three members, while sets with four or more members were uncommon. A possible interpretation of these facts is that languages tend to avoid crowded homophone sets, but when they occur, languages may selectively employ subtle durational differences to disambiguate segmentally homophonous morphs.

Second, the *presence of an alveolar sibilant* increased the average importance of the morphological identity variable. The fricative /s/ is one of the perceptually most salient sounds, and is notorious in sociophonetic research for being the carrier of various non-lexical meanings. It stands to reason that certain segments are more prone than others to express communicative nuances by sub-phonemic detail. The acoustic salience of /s/ appears to contribute to its importance in differentiating homophones by means of acoustic duration.

Third, *homogeneity* of morph type did not show any notable effect on the importance of lexical identity. It did not matter whether all members in a homophone set were of the same morph type (e.g. all suffixes) or whether sets were mixed. This suggests that there is no *a priori* acoustic advantage of root material over affix or clitic material, as would be expected from the phonological literature, but also from previous studies such as Plag et al. (2017), who found that root segments were longer than most affix segments, which in turn were longer than clitic segments in English. This negative result further suggests that positional effects (e.g. the fact that suffixes undergo Final Lengthening more often than prefixes) may nullify the effect of morphological identity in mixed homophone sets.

Lastly, there may be additional features within individual homophone sets that further influence to what degree acoustic differences between homophones are driven by morphology. These features include word class, as in the case of Texistepec Popoluca *wää*; the distinction between underlying vs. derived segments, as in the case of Baïnounk Gubëeher *d*; and morphological interdependence, as in the case of Vera'a *su*, where the two homophones often occur as base and reduplicant within the same word form.

5.2. *Implications for theories of grammar*

The results align with functional theories of grammar that view languages as adaptive systems shaped primarily by the goal of efficient communication (Nichols 1984, Croft 2010, Levshina 2023). The general low importance of morphological identity in explaining acoustic differences between homophonous morphs can be seen as avoiding additional articulatory and cognitive effort, as in most cases, homophones can easily be disambiguated through context. This has also been observed in language acquisition, where homophones that are distributed across syntactic and semantic categories did not pose a learning challenge for children (Dautriche et al. 2018). In highly crowded homophone sets, the risk of confusion increases, making phonetic differentiation more beneficial. As certain segments such as /s/ are particularly well-suited to differentiate meanings, homophones with such segments are also more prone to show systematic acoustic differences. Against this background, the results reported for English <s> can be understood as an adaptation to phonetically disambiguate otherwise homophonous morphs catalyzed by independent factors related to the shape and structure of this particular homophone set. This situation is paralleled by certain other homophone sets such as Beja *ji* but is not the default scenario in the sample studied here.

https://doi.org/10.1017/S0022226725100777 Published online by Cambridge University Press

The role of crowdedness as a crucial factor in predicting temporal differences between homophones appears to be in line with the idea of discriminative learning and the discriminative lexicon (Chuang & Baayen 2021). When a form corresponds to multiple meanings, competition arises in the mental lexicon as acoustic cues are used to resolve ambiguity and predict the correct meaning. Durational differences, by making homophones acoustically distinct, reduce this competition and facilitate more accurate form-to-meaning mapping during comprehension. This resonates with results reported in Tomaschek et al. (2021), who demonstrated that naïve discriminative learning effectively predicts the durational variation of English <s> morphs based on contextual and morphological features. Similarly, Schmitz et al. (2021) showed that linear discriminative learning models capture morpho-phonetic effects even in pseudowords.

As far as formal theories of grammar are concerned, Plag et al. (2017) make the argument that their findings regarding English <s> homophones challenge modular feed-forward theories of grammar. In these frameworks (Kiparsky 1985, Bermúdez-Otero 2012), morphological information is assumed to be inaccessible to phonology and phonetics due to an operation called *bracket erasure*, or *phonetization* in Colored Containment (Trommer 2011). Consequently, sub-phonemic fine-tuning based on information stored in the lexicon has no straightforward explanation in those frameworks. While the acoustic disambiguation of homophones appears to be a rather weak factor for most languages investigated in the present study, some homophones, including English <s>, showed a considerable effect of morphological identity on acoustic duration. This raises the question of how those sub-phonemic differences can possibly be accounted for in formal theories of grammar.

In theories where information on morphological affiliation is not accessible after spell-out, it is possible to encode subphonemic detail directly in the lexicon.[16] Whether this information includes timing gestures, articulatory targets, or simply richer phonological representations depends largely on the chosen framework. In Q theory (Garvin et al. 2018, Shih & Inkelas 2019), otherwise homophonous members of a homophone set could be lexically specified as differing in their subsegmental composition. While segments are usually made up of three subsegments, Q Theory has a provision allowing temporal flexibility by assigning certain segments a lower or higher number of subsegments. Another strand of research, the theory of Gradient Symbolic Representations, suggests that phonological elements can vary in their levels of presence within an underlying representation (Zimmermann 2019). The activity of segments or prosodic nodes can be quantified numerically. For example, two homophones could have slight differences in the underlying activity of certain segments or prosodic nodes associated with those segments, resulting in subtle acoustic differences. Thus, models that allow rich phonological representations obviate the need for morphological look-up in accounting for sub-phonemic acoustic differences.

Lastly, differences in duration could also be the result of different prosodic structures in the lexicon. Plag et al. (2017) do acknowledge this possibility but dismiss this proposal as not applicable to the various <s> homophones in English (see also Schmitz et al. 2021). While a convincing analysis of the English data in terms of prosodic structure is yet to be presented, there is no reason why a prosodic account could not work for other homophone sets in other languages. As has been demonstrated in the literature, differences in prosodic specifications can explain otherwise unexpected phonological and morphological patterns (McCarthy &

[16] Some strands of theory allow the phonology to directly access the morphological identity of morphs – for example, through morphological indexation (Pater 2009). This obviously captures various cases of apparent morpheme-specific phonology, but it is not clear how it would extend to fine phonetic detail.

Prince 1986, Saba Kirchner 2013, Zimmermann 2017). This avenue appears especially promising for heterogenous homophone sets under the condition that independent evidence exists for differences in the amount of prosodic structure between roots, clitics, and affixes.

5.3. Limitations

While the present study encompasses a wide range of typologically diverse languages and considers various factors known to affect morph duration, there are certain limitations inherent to the study design and data that need to be addressed.

Apart from the variables included in this study, there are other factors known to influence segmental duration that could not be considered here for various reasons. One such factor is *prominence*, be it in the form of word-level stress or as phrasal accent under focus. Unfortunately, annotations of prominence are not part of the DoReCo datasets, and the details of word- and phrase-level prosody are still not fully understood for many languages in the sample. Another potentially meaningful variable is *predictability*, which is a more fine-grained measure to predict phonetic reduction than token frequency, which was employed in this study. Predictability takes into consideration the local context in which a word appears and requires large corpora, as well as a deeper understanding of the grammar of a given language (Kliegl et al. 2004, Lee et al. 2012, Tang & Bennett 2018).

The present study focused on segmental homophones, based on the annotations on the morpheme break tier in the DoReCo corpus. However, as mentioned in Section 4.2, some of the languages in DoReCo are tone languages, which means that segmentally homophonous morphs may differ in their tonal make-up, which, in a strict sense, would deprive them of their status as homophones. In addition, some languages are lexical stress languages, which means that two segmentally homophonous morphs may consistently differ in stress placement. In the 37-language sample, seven languages can with some certainty be classified as tone languages and 16 languages as lexical stress languages. At first glance, tone languages and languages with lexical stress seem to be neither over-represented nor under-represented in the list of homophones with high morph importance scores in Table 2. Nevertheless, taking into account more information on prosodic features in cross-linguistic studies of homophones would be desirable in future research.

A methodological difference between this study and studies such as Plag et al. (2017), Seyfarth et al. (2017) or Zuraw et al. (2021) is that the present study is concerned with homophonous morphs but not parts of stems or complex full words that are homophonous with freestanding word forms. While it would be in theory possible to automatically compare homophonous strings smaller or larger than a single morph, interpreting the results over a large sample of typologically diverse languages could prove to be challenging. It appears a more sensible approach would start with an informed pre-selection on an individual language basis and then incrementally build up a reasonably sized sample for comparison. This will hopefully be possible in the future as more annotated and time-aligned data for endangered and understudied languages become available.

6. Conclusion

Do speakers systematically use fine phonetic details to distinguish segmentally homophonous morphs? Based on annotated spontaneous speech data from 37 languages, the answer appears to be ‘no’, at least on a broad scale. While homophones often exhibit differences in acoustic duration, those differences are only minimally influenced by their homophonic nature. Instead, factors such as speech rate and frequency are significantly stronger predictors. As most homophones do not present communicative challenges, it is generally unnecessary to enhance their acoustic distinction. However, in exceptional cases where there is a crowded homophone space with multiple competing meanings and the segmental composition allows for phonetic fine-tuning, speakers may selectively differentiate homophones based on their temporal duration. In these instances, disambiguating homophones through additional articulatory effort seems to benefit efficient communication. These findings provide a framework for understanding the well-known case of English <s> and contextualize it in a wider cross-linguistic perspective.

Acknowledgements. This research was supported by a grant from the German Research Foundation to Ludger Paschen (DFG-PA2368/1-1). I want to thank the audiences of the ‘Mismatches and Disambiguation’ group at ZAS Berlin and the members of the Potsdam Morpho-Syntax Lab for helpful comments and feedback. I am grateful to Aleksandr Schamberger for his support at various stages of this study. I thank three anonymous reviewers for their suggestions and constructive criticism. I am especially grateful to Nicholas Rolle, Matt Stave and Frank Seifart for previous discussions about homophony in the context of the DoReCo corpus, which partially inspired this study. All errors remain my own.

References

Aronoff, Mark. 2017. A fox knows many things but a hedgehog one big thing. In Andrew Hippisley & Gregory Stump (eds.), *The Cambridge handbook of morphology*, 186–205. Cambridge: Cambridge University Press.

Aznar, Jocelyn. 2024. Nisvai DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*, Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.2801565f. https://doreco.huma-num.fr/languages/nisv1234.

Bardají, Maria, Christoph Bracks, Claudia Leto, Datra Hasan, Sonja Riesberg, Winarno S. Alamudi & Nikolaus P. Himmelmann. 2024. Totoli DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0.* Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.c8b6ei29 https://doreco.huma-num.fr/languages/toto1304.

Beckman, Mary E. & Jan Edwards. 1980. The phonological domains of final lengthening. *The Journal of the Acoustical Society of America.* 67(5). 1764–1771. doi:10.1121/1.384304.

Beckmann, Jill N. 1998. *Positional faithfulness*. University of Massachusetts, Amherst dissertation.

Benua, Laura. 2000. *Phonological relations between words: Outstanding dissertations in linguistics*. New York & London: Garland.

Bermúdez-Otero, Ricardo. 2012. The architecture of grammar and the division of labour in exponence. In Jochen Trommer (ed.), *The morphology and phonology of exponence*, 8–83. Oxford: Oxford University Press.

Blasi, Damián E., Joseph Henrich, Evangelia Adamou & David Kemmerer. 2022. Over-reliance on English hinders cognitive science. *Trends in Cognitive Sciences.* 26(12). 1153–1170. doi:10.1016/j.tics.2022.09.015.

Blevins, Juliette & Andrew Wedel. 2009. Inhibited sound change: An evolutionary approach to lexical competition. *Diachronica.* 26(2). 143–183. doi:10.1075/dia.26.2.01ble.

Blum, Frederic, Ludger Paschen, Robert Forkel, Susanne Fuchs & Frank Seifart. 2024. Consonant lengthening marks the beginning of words across a diverse sample of languages. *Nature Human Behaviour*. doi:10.1038/s41562-024-01988-4.

Bogomolova, Natalia, Dmitry Ganenkov & Nils Norman Schiborr. 2024. Tabasaran DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.ad7f97xr. https://doreco.huma-num.fr/languages/taba1259.

Burenhult, Niclas. 2024. Jahai DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.6a71xp0p. https://doreco.huma-num.fr/languages/jeha1242.

Bybee, Joan. 2002. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change.* 14(3). 261–290. doi:10.1017/S0954394502143018.

Chen, Matthew. 1970. Vowel length variation as a function of the voicing of the consonant environment. *Phonetica.* 22(3). 129–159. doi:10.1159/000259312.

Chodroff, Elanor & Colin Wilson. 2022. Uniformity in phonetic realization: Evidence from sibilant place of articulation in American English. *Language.* 98(2). 250–289. doi:10.1353/lan.2022.0007.

Chuang, Yu-Yin & R. Harald Baayen. 2021. Discriminative learning and the lexicon: NDL and LDL. In *Oxford Research Encyclopedia of Linguistics*. Oxford: Oxford University Press. doi:10.1093/acrefore/9780199384655.013.375.

Cobbinah, Alexander. 2013. *Nominal classification and verbal nouns in Baïnounk Gubëher*. SOAS, University of London dissertation.

Cobbinah, Alexander Yao. 2024. Baïnounk Gubëeher DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.a332abw8. https://doreco.huma-num.fr/languages/bain1259.

Coupé, Christophe, Yoon Mi Oh, Dan Dediu & François Pellegrino. 2019. Different languages, similar encoding efficiency: Comparable information rates across the human communicative niche. *Science Advances.* 5(9). doi: 10.1126/sciadv.aaw2594.

Cowell, Andrew. 2024. Arapaho DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.36f5r1b6. https://doreco.huma-num.fr/languages/arap1274.

Cowell, Andrew & Alonzo Moss Sr. 2008. *The Arapaho language*. Boulder, CO: University Press of Colorado.

Croft, William. 2010. Construction Grammar. In Dirk Geeraerts & Hubert Cuyckens (eds.), *The oxford handbook of cognitive linguistics*, 463–508. Oxford: Oxford University Press. doi:10.1093/oxfordhb/9780199738632.013.0018.

Cysouw, Michael. 2007. Building semantic maps: The case of person marking. In Matti Miestamo & Bernhard Wälchli (eds.), *New challenges in typology: Broadening the horizons and redefining the foundations*, 225–248. Berlin: De Gruyter Mouton. doi:10.1515/9783110198904.4.225.

Dautriche, Isabelle, Laia Fibla, Anne-Caroline Fievet & Anne Christophe. 2018. Learning homophones in context: Easy cases are favored in the lexicon of natural languages. *Cognitive Psychology.* 104. 83–105. doi:10.1016/j.cogpsych.2018.04.001.

Drager, Katie. 2011. Sociophonetic variation and the lemma. *Journal of Phonetics.* 39(4). 694–707. doi:10.1016/j.wocn.2011.08.005.

Däbritz, Chris Lasse, Nina Kudryakova, Eugénie Stapert & Alexandre Arkhipov. 2024. Dolgan DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.f09eikq3. https://doreco.huma-num.fr/languages/dolg1241.

Döhler, Christian. 2018. *A grammar of Komnzo* (Studies in Diversity Linguistics 22). Berlin: Language Science Press. doi:10.5281/zenodo.1477799.

Esposito, Lewis. 2020. Linking gender, sexuality, and affect: The linguistic and social patterning of phrase-final posttonic lengthening. *Language Variation and Change.* 32(2). 191–216. doi:10.1017/S0954394520000095.

Fletcher, Janet. 2010. The prosody of speech: Timing and rhythm. In William J. Hardcastle, John Laver & Fiona E. Gibbon (eds.), *The handbook of phonetic sciences* (2nd edn), 521–602. Chichester: Blackwell Publishing. doi: 10.1002/9781444317251.ch15.

Forker, Diana & Nils Norman Schiborr. 2024. Sanzhi Dargwa DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference Corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.81934177. https://doreco.huma-num.fr/languages/sanz1248.

Franjieh, Michael. 2024. Fanbyak DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.02084446. https://doreco.huma-num.fr/languages/orko1234.

Gahl, Susanne. 2008. Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language.* 84(3). 474–496. doi:10.1353/lan.0.0035.

Garcia-Laguia, Alexandro. 2024. Northern Alta DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du

Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.efea0b36. https://doreco.huma-num.fr/languages/nort2875.

Garvin, Karee, Myriam Lapierre & Sharon Inkelas. 2018. A Q-Theoretic approach to distinctive subsegmental timing. *Proceedings of the Linguistic Society of America.* 3(1). 9:1–13. doi:10.3765/plsa.v3i1.4330.

Godfrey, John J. & Holliman Edward. 1997. *Switchboard-1 release 2*. Philadelphia: Linguistic Data Consortium. https://catalog.ldc.upenn.edu/ldc97s62.

Gordon, Matthew, Paul Barthmaier & Kathy Sands. 2002. A cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association.* 32. 141–174.

Goudswaard, Nelleke. 2004. Infix allomorphy in Ida'an-Begak. *ZAS Papers in Linguistics.* 34. 89–102. doi:10.21248/zaspil.34.2004.205.

Gusev, Valentin, Tiina Klooster, Beáta Wagner-Nagy & Alexandre Arkhipov. 2024. Kamas DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.cdd8177b. https://doreco.huma-num.fr/languages/kama1351.

Güldemann, Tom, Martina Ernszt, Sven Siegmund & Alena Witzlack-Makarevich. 2024. N||ng DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.f6c37fi0. https://doreco.huma-num.fr/languages/nngg1234.

Haig, Geoff, Maria Vollmer & Hanna Thiele. 2024. Northern Kurdish (Kurmanji) DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.ca10ez5t. https://doreco.huma-num.fr/languages/nort2641.

Hartmann, Iren. 2024. Hoocąk DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.b57f5065. https://doreco.huma-num.fr/languages/hoch1243.

Harvey, Andrew. 2024. Gorwaa DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.a4b4ijj2. https://doreco.huma-num.fr/languages/goro1270.

Haspelmath, Martin. 2003. The geometry of grammatical meaning: semantic maps and cross-linguistic comparison. In M. Tomasello (ed.), *The new psychology of language*, 211–243. Mahwah, NJ: Erlbaum.

Haspelmath, Martin. 2020. The morph as a minimal linguistic form. *Morphology.* 30. 117–134. doi:10.1007/s11525-020-09355-5.

Haude, Katharina. 2024. Movima DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.da42xf67. https://doreco.huma-num.fr/languages/movi1243.

Hellwig, Birgit. 2024. Goemai DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.b93664ml. https://doreco.huma-num.fr/languages/goem1240.

Henrich, Joseph, Steven J. Heine & Ara Norenzayan. 2010. The weirdest people in the world? *Behavioral and Brain Sciences.* 33(2–3). 61–83. doi:10.1017/S0140525X0999152X.

Kazakevich, Olga & Elena Klyachko. 2024. Evenki DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.5e0d27cu. https://doreco.huma-num.fr/languages/even1259.

Kemps, Rachèl J.J.K., Mirjam Ernestus, Robert Schreuder & Harald R. Baayen. 2005. Prosodic cues for morphological complexity: The case of Dutch plural nouns. *Memory & Cognition.* 33. 430–446. doi:10.3758/BF03193061.

Kim, Soung-U. 2024. Jejuan DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.06ebrk38. https://doreco.huma-num.fr/languages/jeju1234.

Kiparsky, Paul. 1985. Some consequences of lexical phonology. *Phonology Yearbook.* 2. 85–138.

Kisler, Thomas, Florian Schiel & Han Sloetjes. 2012. Signal processing via web services: The use case WebMAUS. *Proceedings of Digital Humanities* 2012. 30–34.

Klatt, Dennis H. 1973. Interaction between two factors that influence vowel duration. *The Journal of the Acoustical Society of America.* 54. 1102–1104.

Kliegl, Reinhold, Ellen Grabner, Martin Rolfs & Ralf Engbert. 2004. Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology.* 16(1–2). 262–284. doi: 10.1080/09541440340000213.

Krifka, Manfred. 2024. Daakie DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*, Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.efeav5l9. https://doreco.huma-num.fr/languages/port1286.

Kurisu, Kazutaka. 2001. *The phonology of morpheme realization*. University of California, Santa Cruz dissertation.

Lavoie, Lisa. 2002. Some influences on the realization of for and four in American English. *Journal of the International Phonetic Association.* 32(2). 175–202. doi:10.1017/S0025100302001032.

Lee, Chia-Ying, Yu-Ning Liu & Jie-Li Tsai. 2012. The time course of contextual effects on visual word recognition. *Frontiers in Psychology.* 3. doi:10.3389/fpsyg.2012.00285.

Lehiste, Ilse. 1972. The timing of utterances and linguistic boundaries. *Journal of the Acoustical Society of America.* 51. 2018–2024.

Levshina, Natalia. 2016. Why we need a token-based typology: A case study of analytic and lexical causatives in fifteen European languages. *Folia Linguistica.* 50(2). 507–542. doi:10.1515/flin-2016-0019.

Levshina, Natalia. 2021. Conditional inference trees and random forests. In Magali Paquot & Stephan Th. Gries (eds.), *A practical handbook of corpus linguistics*, chap. 25, 611–643. New York: Springer. doi:10.1007/978-3-030-46216-1_25.

Levshina, Natalia. 2023. *Communicative efficiency: Language structure and use*. Cambridge: Cambridge University Press.

Liaw, Andy & Matthew Wiener. 2002. Classification and regression by randomForest. *R News.* 2(3). 18–22. https://CRAN.R-project.org/doc/Rnews/.

Linders, Guido M. & Max M. Louwerse. 2023. Zipf's law revisited: Spoken dialog, linguistic units, parameters, and the principle of least effort. *Psychonomic Bulletin & Review.* 30. 77–101. doi:10.3758/s13423-022-02142-9.

Linville, Sue Ellen. 1998. Acoustic correlates of perceived versus actual sexual orientation in men's speech. *Folia Phoniatr Logop.* 50. 35–48.

Mack, Sara & Benjamin Munson. 2012. The influence of /s/ quality on ratings of men's sexual orientation: Explicit and implicit measures of the 'gay lisp' stereotype. *Journal of Phonetics.* 40(1). 198–212.

Maddieson, Ian & Sandra Ferrari Disner. 1984. *Patterns of sounds*. Cambridge: Cambridge University Press.

McCarthy, John J. & Alan S. Prince. 1986. Prosodic Morphology. Unpublished Ms., University of Massachusetts, Amherst, and Brandeis University, Waltham, MA.

Moroz, George. 2017. *lingtypology: Easy mapping for linguistic typology.* https://CRAN.R-project.org/package=lingtypology.

Morrison, Donald Alasdair. 2021. Vowel nasalisation in Scottish Gaelic: No evidence for incomplete neutralisation in initial mutation. *Morphology.* 31. 121–146. doi:10.1007/s11525-020-09347-5.

Mosel, Ulrike. 2024. Teop DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.9322sdf2. https://doreco.huma-num.fr/languages/teop1238.

Munson, Benjamin, Elizabeth C McDonald, Nancy L DeBoe & Aubrey R White. 2006. The acoustic and perceptual bases of judgments of women and men's sexual orientation from read speech. *Journal of Phonetics.* 34(2). 202–240.

Munteanu, Andrei. 2021. Homophony avoidance in the grammar: Russian nominal allomorphy. *Phonology.* 38(3). 401–435. doi:10.1017/S0952675721000257.

Nichols, Johanna. 1984. Functional theories of grammar. *Annual Review of Anthropology.* 13. 97–117.

Nida, Eugene A. 1948. The identification of morphemes. *Language.* 24. 414–441.

Nylén, Fredrik, Jenny Holmberg & Maria Södersten. 2024. Acoustic cues to femininity and masculinity in spontaneous speech. *The Journal of the Acoustic Society of America.* 155(5). 3090–3100. doi:10.1121/10.0025932.

Paschen, Ludger, François Delafontaine, Christoph Draxler, Susanne Fuchs, Matthew Stave & Frank Seifart. 2020. Building a time-aligned cross-linguistic reference corpus from language documentation data (DoReCo). In *Proc. LREC 2020*, Marseille, 2657–2666. https://www.aclweb.org/anthology/2020.lrec-1.324.

Paschen, Ludger, Susanne Fuchs & Frank Seifart. 2022. Final lengthening and vowel length in 25 languages. *Journal of Phonetics.* 94. doi:10.1016/j.wocn.2022.101179.

Pater, Joe. 2009. Morpheme-specific phonology. In S. Parker (ed.), *Phonological argumentation: Essays on evidence and motivation*, 123–154. London: Equinox.

Plag, Ingo, Julia Homann & Gero Kunter. 2017. Homophony and morphology: The acoustics of word-final S in English. *Journal of Linguistics.* 53(1). 181–216. doi:10.1017/S0022226715000183.

Pluymaekers, Mark, Mirjam Ernestus & R. Harald Baayen. 2006. Effects of word frequency on the acoustic durations of affixes. In *Proc Interspeech* 2006, 953–956.

Ponsonnet, Maïa. 2024. Dalabon DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.fae299ug. https://doreco.huma-num.fr/languages/ngal1292.

Price, P. J. & H. J. Simon. 1984. Perception of temporal differences in speech by "normal-hearing" adults: Effects of age and intensity. *The Journal of the Acoustical Society of America.* 76(2). 405–410. doi:10.1121/1.391581.

Quesada, Juan Diego, Stavros Skopeteas, Carolina Pasamonik, Carolin Brokmann & Florian Fischer. 2024. Cabécar DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.ebc4ra22. https://doreco.huma-num.fr/languages/cabe1245.

R Core Team. 2023. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna, Austria. https://www.R-project.org/.

Rathcke, Tamara & Christine Mooshammer. 2022. 'Grandpa' or 'opera'? Production and perception of unstressed /a/ and /əʁ/ in German. *Journal of the International Phonetic Association.* 52(1). 33–58. doi:10.1017/S0025100320000110.

Recasens, Daniel. 2018. Coarticulation. In *Oxford research encyclopedia of linguistics*. Oxford: Oxford University Press. doi:10.1093/acrefore/9780199384655.013.416.

Reiter, Sabine. 2024. Cashinahua DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.a8f9q2f1. https://doreco.huma-num.fr/languages/cash1254.

Rice, Keren. 2014. Athabaskan. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of derivational morphology*, 669–701. Oxford: Oxford University Press.

Riesberg, Sonja. 2024. Yali (Apahapsili) DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.9d91nkq2. https://doreco.huma-num.fr/languages/apah1238.

Ring, Hiram. 2024. Pnar DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.5ba1062k. https://doreco.huma-num.fr/languages/pnar1238.

Rose, Françoise. 2024. Mojeño Trinitario DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.cbc3b4xr. https://doreco.huma-num.fr/languages/trin1278.

RStudio Team. 2023. *Rstudio build 446*. RStudio, PBC Boston, MA. http://www.rstudio.com/.

Rzymski, Christoph, Tiago Tresoldi, Simon J. Greenhill, Mei-Shin Wu, Nathanael E. Schweikhard, Maria Koptjevskaja-Tamm, Volker Gast, Timotheus A. Bodt, Abbie Hantgan, Gereon A. Kaiping, Sophie Chang, Yunfan Lai, Natalia Morozova, Heini Arjava, Nataliia Hübler, Ezequiel Koile, Steve Pepper, Mariann Proos, Briana Van Epps, Ingrid Blanco, Carolin Hundt, Sergei Monakhov, Kristina Pianykh, Sallona Ramesh, Russell D. Gray, Robert Forkel & Johann-Mattis List. 2020. The database of cross-linguistic colexifications, reproducible analysis of cross-linguistic polysemies. *Scientific Data*. 7(13). doi:10.1038/s41597-019-0341-x.

Saba Kirchner, Jesse. 2013. Minimal reduplication and reduplicative exponence. *Morphology.* 23. 227–243.

Schlechtweg, Marcel & Greville G. Corbett. 2021. The duration of word-final s in English: A comparison of regular-plural and pluralia-tantum nouns. *Morphology* doi:10.1007/s11525-021-09381-x.

Schlechtweg, Marcel, Melina Heinrichs & Marcel Linnenkohl. 2020. Differences in acoustic detail: The realization of syncretic nouns in German. In Marcel Schlechtweg (ed.), *The learnability of complex constructions: A cross-linguistic perspective*, Trends in Linguistics: Studies and Monographs, 39–62 (vol. 345). Berlin: De Gruyter. doi:10.1016/0378-2166(94)90116-3.

Schmitz, Dominic. 2022. *Production, perception, and comprehension of subphonemic detail: Word-final /s/ in English*, Studies in Laboratory Phonology (vol. 11). Berlin: Language Science Press. doi:10.5281/zenodo.7267830.

Schmitz, Dominic & Dinah Baer-Henney. 2024. Morphology renders homophonous segments phonetically different: Word-final /s/ in German. In *Proceedings of Speech Prosody 2024 (SP2024)*, Universiteit Leiden, Netherlands, doi:10.21437/SpeechProsody.2024-119.

Schmitz, Dominic, Ingo Plag, Dinah Baer-Henney & Simon David Stein. 2021. Durational differences of word-final /s/ emerge from the lexicon: Modelling morpho-phonetic effects in pseudowords with linear discriminative learning. *Frontiers in Psychology.* 12. doi:10.3389/fpsyg.2021.680889.

Schnell, Stefan. 2011. A grammar of Vera'a, an Oceanic language of North Vanuatu. Manuscript, Universität Kiel.

Schnell, Stefan. 2024. Vera'a DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.3e2cu8c4. https://doreco.huma-num.fr/languages/vera1241.

Seifart, Frank. 2024. Bora DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.6eaf5laq. https://doreco.huma-num.fr/languages/bora1263.

Seifart, Frank, Ludger Paschen & Matthew Stave. 2024. Language Documentation Reference Corpus (DoReCo) 2.0. doi:10.34847/NKL.7CBFQ779.

Selting, Margret. 1994. Emphatic speech style – with special focus on the prosodic signalling of heightened emotive involvement in conversation. *Journal of Pragmatics.* 22(3). 375–408. doi:10.1016/0378-2166(94)90116-3.

Seyfarth, Scott, Marc Garellek, Gwendolyn Gillingham, Ackerman Farrell & Robert Malouf. 2017. Acoustic differences in morphologically-distinct homophones. *Language, Cognition and Neuroscience*. doi:10.1080/23273798.2017.1359634.

Shih, Stephanie S. & Sharon Inkelas. 2019. Subsegments and the emergence of segments. *Proceedings of the Linguistic Society of America.* 4(1). 37:1–8. doi:10.3765/plsa.v4i1.4541.

Skopeteas, Stavros, Violeta Moisidi, Nutsa Tsetereli, Johanna Lorenz & Stefanie Schröter. 2024. Urum DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.ac166n10. https://doreco.huma-num.fr/languages/urum1249.

Smorenburg, Laura & Willemijn Heeren. 2020. The distribution of speaker information in Dutch fricatives /s/ and /x/ from telephone dialogues. *The Journal of the Acoustical Society of America.* 147(2). 949–960.

Stave, Matthew, François Delafontaine, Frank Seifart & Ludger Paschen. 2019. A bioinformatics solution to inter-rater agreement for forced time-alignment of data from underresourced languages. In *Proc. LIFT, Orléans*, 8–12.

Stave, Matthew, Ludger Paschen, François Delafontaine & Frank Seifart. 2021. Optimization of morpheme length: a cross-linguistic assessent of Zipf's and Menzerath's laws. *Linguistics Vanguard.* 7(s3). 20190076. doi:10.1515/lingvan-2019-0076.

Strunk, Jan, Frank Seifart, Swintha Danielsen, Iren Hartmann, Brigitte Pakendorf, Søren Wichmann, Alena Witzlack-Makarevich & Balthasar Bickel. 2020. Determinants of phonetic word duration in ten language documentation corpora: Word frequency, complexity, position, and part of speech. *Language Documentation & Conservation.* 14. 423–461. http://hdl.handle.net/10125/24926.

Stuart-Smith, Jane. 2007. Empirical evidence for gendered speech production: /s/ in Glaswegian. In J. Cole & J.I. Hualde (eds.), *Laboratory phonology* 9, 65–86. New York: Mouton de Gruyter.

Tang, Kevin & Ryan Bennett. 2018. Contextual predictability influences word and morpheme duration in a morphologically complex language (Kaqchikel Mayan). *The Journal of the Acoustical Society of America.* 144(2). 997–1017. doi:10.1121/1.5046095.

Teo, Amos. 2024. Sümi DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.5ad4t01p. https://doreco.huma-num.fr/languages/sumi1235.

Thieberger, Nick. 2024. Nafsan (South Efate) DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.ba4f760l. https://doreco.huma-num.fr/languages/sout2856.

Tilsen, Sam & Mark Tiede. 2023. Parameters of unit-based measures of speech rate. *Speech Communication.* 150. 73–97. doi:10.1016/j.specom.2023.05.006.

Tomaschek, Fabian & Kristian Berg. 2023. Phonetic characteristics of German unstressed vowels: The case of Opa vs Oper. *JASA Express Letters.* 3(9). 095201. doi:10.1121/10.0020813.

Tomaschek, Fabian, Ingo Plag, Mirjam Ernestus & R. Harald Baayen. 2021. Phonetic effects of morphology and context: Modeling the duration of word-final S in English with naïve discriminative learning. *Journal of Linguistics.* 57(1). 123–161. doi:10.1017/S0022226719000203.

Trommer, Jochen. 2011. *Phonological aspects of Western Nilotic mutation morphology.* Leipzig University Habilitation.

Turk, Alice E & Stefanie Shattuck-Hufnagel. 2007. Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics.* 35(4). 445–472. doi:10.1016/j.wocn.2006.12.001.

Valera, Salvador. 2020. Polysemy versus homonymy. In *Oxford research encyclopedia of linguistics*. Oxford: Oxford University Press. doi:10.1093/acrefore/9780199384655.013.617.

Vanhove, Martine. 2017. *Le Bedja*. Leuven & Paris: Peeters.

Vanhove, Martine. 2024. Beja DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.edd011t1. https://doreco.huma-num.fr/languages/beja1238.

Vicente, Agustín & Ingrid L. Falkum. 2017. Polysemy. In *Oxford research encyclopedia of linguistics*. Oxford: Oxford University Press. doi:10.1093/acrefore/9780199384655.013.325.

Vydrina, Alexandra. 2024. Kakabe DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.d5aeu9t6. https://doreco.huma-num.fr/languages/kaka1265.

Walsh, Thomas & Frank Parker. 1983. The duration of morphemic and non-morphemic /s/ in english. *Journal of Phonetics.* 11(2). 201–206. doi:10.1016/s0095-4470(19)30816-2.

Wegener, Claudia. 2024. Savosavo DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.b74d1b33. https://doreco.huma-num.fr/languages/savo1255.

Weirich, Melanie & Adrian P. Simpson. 2015. Gender-specific differences in sibilant contrast realizations in English and German. In *Proc. ICPhS 2015*, Paper 0261.

White, Laurence, Silvia Benavides-Varela & Katalin Mády. 2020. Are initial-consonant lengthening and final-vowel lengthening both universal word segmentation cues? *Journal of Phonetics.* 81. 100982. doi:10.1016/j.wocn.2020.100982.

Wichmann, Søren. 2024. Texistepec Popoluca DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi:10.34847/nkl.c50ck58f. https://doreco.huma-num.fr/languages/texi1237.

Winter, Bodo & Marcus Perlman. 2021. Size sound symbolism in the English lexicon. *Glossa: A Journal of General Linguistics.* 6(1). 1–13. doi:10.5334/gjgl.1646.

Witzlack-Makarevich, Alena, Saudah Namyalo, Anatol Kiriggwajjo & Zarina Molochieva. 2024. Ruuli DoReCo dataset. In Frank Seifart, Ludger Paschen & Matthew Stave (eds.), *Language documentation reference corpus (DoReCo) 2.0*. Lyon: Laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). doi: 10.34847/nkl.fde4pp1u. https://doreco.huma-num.fr/languages/ruul1235.

Wu, Di. 2017. *Cross-regional word duration patterns in Mandarin.* University of Illinois at Urbana-Champaign dissertation. http://hdl.handle.net/2142/97356.

Zimmermann, Eva. 2017. *Morphological length and prosodically defective morphemes*, Oxford Studies in Phonology and Phonetics (vol. 1). Oxford: Oxford University Press.

Zimmermann, Eva. 2019. Gradient symbolic representations and the typology of ghost segments. In *Proceedings of AMP* 2018, 1–12. doi:10.3765/amp.v7i0.4576.

Zipf, George Kingsley. 1935. *The psycho-biology of language: An introduction to dynamic philology.* Houghton, Mifflin.

Zuraw, Kie, Isabelle Lin, Meng Yang & Sharon Peperkamp. 2021. Competition between whole-word and decomposed representations of English prefixed words. *Morphology.* 31. 201–237. doi:10.1007/s11525-020-09354-6.

Appendix

This appendix includes information on four alternative models that use a slightly different array of variables or model structures. It also discusses whether dataset size may influence the importance of the morph variable.

Reversed model

It is in principle possible to construct a reversed model that predicts morphological identity from morph duration. The variable importance scores of such a 'flipped' model is given in Figure 7. The figure shows that morph_duration is slightly better at predicting morphological identity than vice versa, with an importance of ca. 11%. The order of the other factors is almost reversed compared to the original model, with word_frequency and segmental_context having high and position and speaker having low contribution values. These results are not surprising given that this is a reversed model.

Model without the `speaker` variable

As noted in Section 4.1, DoReCo is not balanced when it comes to the speaker variable, and datasets differ dramatically in how many speakers they contain. The variable importance scores of a model that does not include the factor speaker is given in Figure 8. The ranking of the factors is identical to the one in the model described in Section 4 (Figure 3), indicating that the unbalanced nature of the speaker variable does not skew the model in any direction.

Model with extended segmental context information

Figure 9 shows a model in which segmental_context is split into left and right context and encodes the individual segments rather than broad groups of segments. In such a model with maximum information on segmental context, the two granular context factors are virtually identical to the broad segmental_context factor from the model described in Section 4 in terms of their relatively low contribution scores and their ranking barely above morph.

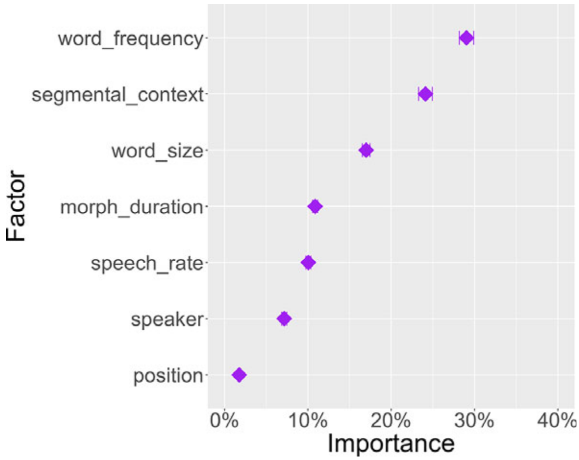

Figure 7. Average contribution of seven variables for predicting morphological identity in a 'flipped' model.

https://doi.org/10.1017/S0022226725100777 Published online by Cambridge University Press

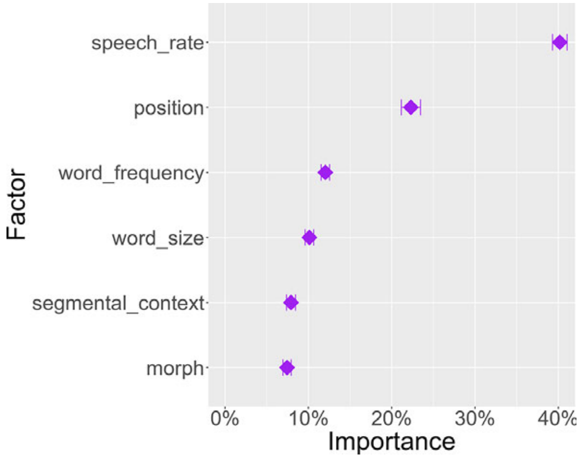

Figure 8. Average contribution of six variables (without `speaker`) for predicting morph duration.

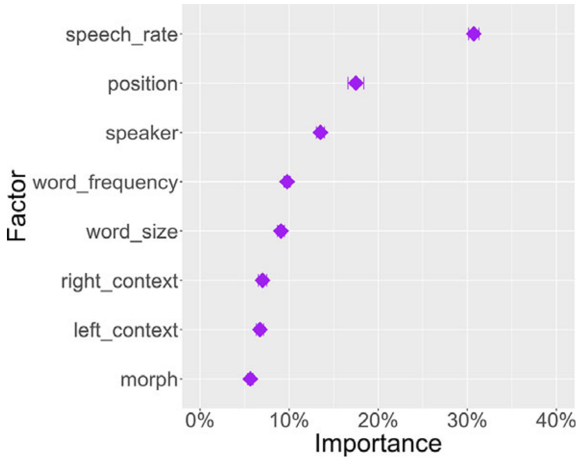

Figure 9. Average contribution of eight variables for predicting morph duration.

Model with an additional `morph_type` factor

Figure 10 shows a model with an additional `morph_type` variable. Intuitively, one might expect this variable to be highly informative, as phonological and phonetic differences between roots, clitics and affixes affecting their duration have been described in the literature (Beckmann 1998, Plag et al. 2017). However, `morph_type` has a variable importance of close to 0%. The reason is that there is an implicational relation between `morph` and `morph_type` whereby each `morph` has exactly one value for `morph_type`, and in any given homophone set, the maximum number of categories for `morph_type` depends on `morph` (but not vice versa).

The relation between dataset size and the `morph` variable

Figure 11 shows the relation between dataset size (in word tokens) and the importance of `morph`. No meaningful interaction between these two parameters is visible, and linear regression reveals a lack of a significant correlation (R = -0.03, p = 0.52).

https://doi.org/10.1017/S0022226725100777 Published online by Cambridge University Press

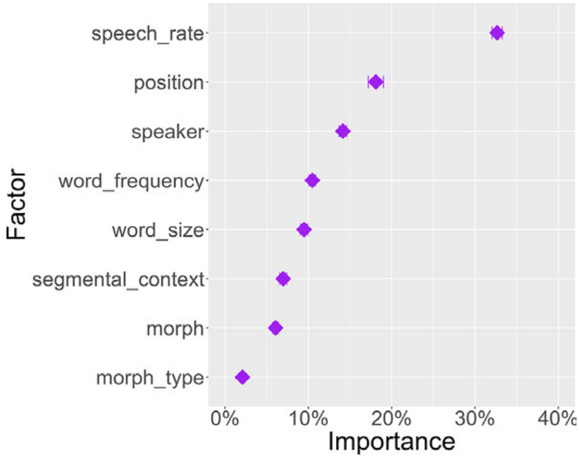

Figure 10. Average contribution of eight variables for predicting morph duration.

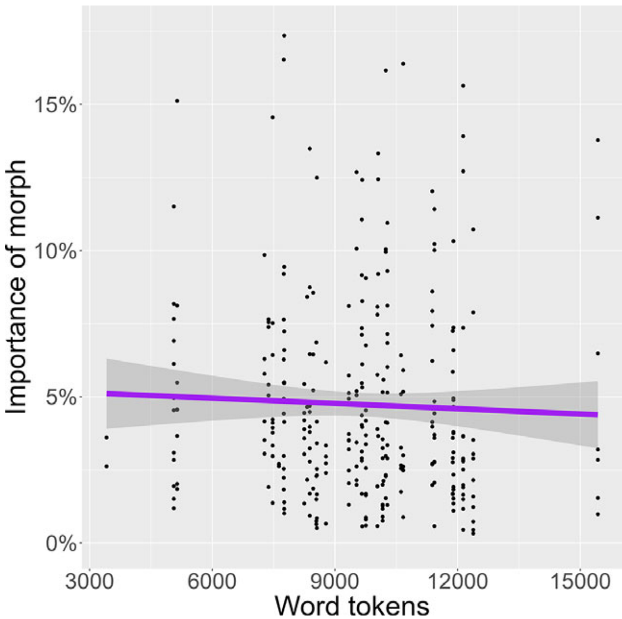

Figure 11. The relation between word tokens (per dataset) and the importance of `morph`.

Cite this article: Paschen, Ludger. 2025. "Acoustic disambiguation of homophonous morphs is exceptional." *Journal of Linguistics*. doi:10.1017/S0022226725100777