

Article

Word-Final /s/-/z/ Omission in Vietnamese English

Stephen J. Disney ^{1,*}  and Le Nu Cam Le ² ¹ School of Education, Plymouth Marjon University, Plymouth PL6 8BH, UK² School of Foreign Languages, Hanoi University of Science and Technology, Đại Cồ Việt, Bách Khoa, Hai Ba Trung, Hà Nội 100000, Vietnam

* Correspondence: sdisney@marjon.ac.uk

Abstract: Southeast Asian learners of English, including those from Vietnam, frequently omit word-final consonants in their English speech. Previous work on Vietnamese learners of English is limited, and errors are typically usually attributed to first-language transfer effects. No large-scale empirical study on Vietnamese learners has been carried out to aid the development of an evidence-based pedagogy. This study uses authentic spoken data to compare lexical and morphological word-final /s/ and /z/ in the speech of sixteen Vietnamese adult learners of English. We discuss the relative impact of frequency of use, whether the instance of a target /s/ or /z/ is in a root or bound morpheme, and whether the preceding phoneme is a consonant or vowel. An overall omission rate of 28.4% of expected instances was found. Morphological {-s} when it is preceded by a consonant has the highest error rate (50.7%). A multilevel binary logistic regression was performed to ascertain the relative effects. Morphological words containing /s/ or /z/ were significantly more likely to be pronounced with the /s/ or /z/ absent than lexical words containing a /s/ or /z/, as were those in clusters compared to those with a preceding vowel. The results indicate that phonological effects and morphological effects are stacked and not multiplicative and that the observed omission rates are not solely attributable to L1 transfer effects. Frequency of use is also highly correlated with accuracy.

Keywords: L2 Phonology; second language acquisition; morphology; frequency; Vietnamese English; Southeast Asian Englishes



Citation: Disney, Stephen J., and Le Nu Cam Le. 2024. Word-Final /s/-/z/ Omission in Vietnamese English. *Languages* 9, 0. <https://doi.org/>

Academic Editors: Martha Young-Scholten, Alex Ho-Cheong Leung and Natalia Pavlovskaya

Received: 22 September 2023
Revised: 26 September 2024
Accepted: 29 September 2024
Published: 8 October 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

“[I]f to err and to speak are each uniquely human, then to err at speaking, or to commit language errors, must mark the very pinnacle of human uniqueness.” (James 1998, p. 1)

Simplification of linguistic elements is common in learner English, particularly where there is a mismatch in the distribution patterns between a learner’s first language (L1) and their second (L2) or possible third language (L3). How language learners deal with difficult sounds and structures includes processes of elision, omission, substitution, and epenthesis. Omission is a commonly used strategy for learners to mitigate difficult L2 features. As Slobin (1992, p. 191) points out, in L2 acquisition, language gets “blurry”, and speakers “smudge” their phonology as they “delete and contract surface forms”, a perspective that resonates strongly with the current study.

Southeast (SE) Asian L2 learner Englishes are widely reported as exhibiting omission of consonant clusters and in syllable codas, particularly in word-final (WF) position. This can result in significant and persistent errors, with frequent breakdowns in communication, yet there are no evidence-based teaching materials specifically designed to mitigate these issues. In Viet Nam, the issue has not been the subject of any large-scale study based on samples of naturalistic data and it is unclear what the relative influence is of the different variables that may have an impact on the general tendency for phonemes to be omitted. To fill this gap, this paper presents the findings of a study that explores the speech patterns of

Vietnamese learners of English (VLEs) carrying out a monologue task. It uses a collated video archive of sixteen student presentations as the primary data source to highlight the relative influences on the error rates of WF alveolar fricatives /s/ and /z/. around 21% of all the different words used in spoken English have /s/ or /z/ as the final sound. The variables were whether the instance of a target /s/ or /z/ is in a root or bound morpheme, whether the preceding phoneme is a consonant or vowel, and frequency of use in the data. While the tendency to omit fricatives in VLEs affects the class as a whole, the current paper focuses only on word-final /s/ and /z/ because only these can be both a bound morpheme, as in *eat-s*, and non-morphological in a root morpheme, as in *ice*. As the former, they can be a third-person singular verb, a plural noun, a clitic (*it's* and *that's* and so on), or a possessive. As a result of these multiple uses, The underlying assumption often used to explain this pattern of omission is that it results from L1 interference and that by contrasting languages to find mismatches in surface forms, likely errors can be predicted. The current paper adds to the large and growing volume of work that undermines such a “contractive analysis” approach to working with L2 errors.

The paper first summarises approaches to morphophonemics, Second Language Acquisition (SLA), and the roles of L1 transfer and frequency. It then describes the limited previous research on Vietnamese learners’ acquisition of the target sounds. In Section 2, we outline the methodology, and in Section 3, the findings on WF fricatives /s/ and /z/ are presented. Section 4 discusses the results, and Section 5 briefly concludes the paper and outlines potential avenues for future work. This paper concerns only Vietnamese learners of English; however, speakers with other L1s and their teachers from the wider SE Asia area will also find the results relevant.

1.1. Theoretical Background

Most adult second-language learners who are not acquiring two languages in a naturalistic bilingual setting undergo formal L2 instruction of some kind. Their learning is characterised by systematic “errors” (as opposed to performance-related “mistakes”, see Corder 1967 and James 1998). Their psychological representation of the target language can be seen to evolve from sporadically target-like towards a (supposed) goal of entirely target-like. Between these poles lies a learner’s “interlanguage”, an unstable, evolving state of language development (Corder 1967; Gass and Selinker 2010; Selinker 1972; Tarone and Han 2014). Accounting for learner errors in SLA as their interlanguage becomes more target-like or, in many cases, as it fails to do so, is central to the research agenda. The results have often informed practice in L2 pedagogy (see Gass et al. 2020).

Researchers have long tried to predict errors by examining differences between a speaker’s L1 and their L2 and to explain errors as resulting from the “transfer” of features found in the L1 onto the evolving L2. Transfer is “the influence resulting from similarities and differences between the target language and any other language that has been previously (and perhaps imperfectly) acquired” (Odlin 1989, p. 27). Lado’s (1957) Contrastive Analysis Hypothesis explicitly argued that it is possible to contrast the system found in the L1 (e.g., the grammar, phonology and lexicon) with that of the L2 in order to predict the difficulties an L2 learner may experience. This later evolved to argue that where two languages are similar, positive transfer would occur, and where they are different, negative transfer, or interference, would result (Cummins 1979). Transfer can then be thought of as being both helpful and unhelpful to the learner. In particular, interference, i.e., where the L1 “interferes” with L2 acquisition, has attracted much attention. However, the contrastive approaches turned out to both predict errors that did not occur in authentic learner data and failed to predict errors that did occur.

In order to produce a sound in a language, one must first be able to perceive it, and transfer effects also apply to the processing of input. It has been observed that the perception of sounds may be inaccurate, and some illusion effects can occur, which serve to confound the consistent and accurate weighting of variants (Leung et al. 2023). In L2 acquisition, the perception of phonemes is filtered by L1 contrasts, and where a match

between L2 input and L1 is not clear, “listeners perceive the sound categorically” (Hwa-Froelich et al. 2002, p. 266). If the perception of a sound is thus filtered, then accurate production is clearly made more difficult. The extent to which input feedback via the phonological loop of one’s own production might negatively affect learning is also unclear (Baese-Berk and Samuel 2016).

While transfer effects can account for some sources of error, the field accepts in general that (a) it is difficult to determine what exactly that effect is, and (b) there are other sources of error that the L1 transfer effect cannot account for, especially where morphology and phonology interact. The 1980s saw an explosion of work in SLA seeking to downplay the role of the L1 (Larsen-Freeman 1991). The evidence from these studies suggests that there is no simple one-to-one correspondence between ease of learning and areas where the L1 and target language exhibit mismatches in features.

This is particularly so when considering how a learner’s phonology develops as it interacts with morphology. The phonemes /s/ and /z/ that are the subject of the study here can be root or bound morphemes, where the latter account for 42.5% of instances in our data. The field of morphophonemics has evolved over the last century from structuralist accounts (Bloomfield 1933) through generative work (Chomsky and Halle 1968; Aronoff 1976; Aronoff 1994) and beyond, including Lexical Phonology (Kiparsky 1982) and Paradigm Uniformity (Benua 2005). However, the fact that there is still significant discussion reveals the difficulties in attempting to conclusively explain how phonology and morphology interact and develop, particularly for an L2 learner with a highly advanced interlanguage, where errors are still highly variable.

One important example of an alternative to L1 transfer-based accounts is the markedness approach, which draws on implicational universals (Greenberg 1965). For an opposition such as +/– obstruent voicing, a markedness approach would consider the variant that is the more widely distributed to be simpler, or more basic; this is the unmarked member of the pair. The other member of the opposition is the marked member. The Markedness Differential Hypothesis (Eckman 1977), based on the Prague School phonological theory of markedness and the Similarity Differential Rate Hypothesis (Major and Kim 1996), built on this work to help account for issues that arise in L2 learning. It is argued that a marked phenomenon that does not exist in the L1 is usually acquired at a slower rate than an unmarked phenomenon (Major 2008). It is only with some considerable difficulty and time that such contrasts are learnt and used consistently. Major and Kim (1996) concluded that language teaching should focus on areas of difference in order to increase accuracy. Voiced obstruents, for instance, are marked compared to voiceless ones, so they are slower to be acquired in SLA (Gass et al. 2020, p. 224; see Flege et al. 1992). It has also been observed that an intermediate value may be produced by a learner (Flege and Eefting 1987), thus making the contrast between phonemes less distinct.

Markedness has also been shown to apply to cross-linguistic data and language change processes. For example, it appears that no languages natively have /s/ word-finally without also having it word-initially. A small number of languages appear to break this universal, but their word-final /s/ seems to have been borrowed from English. For example, Luo has borrowed the English plural morpheme, and Tongan has borrowed a few words with alveolar fricatives at the end. As a result, there is an implicational universal (after Greenberg 1965) that suggests that word-initial /s/ is ‘unmarked’ with respect to /s/ word-finally. Further, in language change during a process of grammaticalisation, lenition through devoicing is common, but the opposite is rare (Bybee and Hopper 2001). For example, the English BE *supposed to* construction, the /d/ in *supposed* underwent reduction to /t/ (Traugott 1989; Disney 2016). This also suggests that + voicing is the marked member of the pair.

In terms of how learners deal with issues arising from such differences between their L1 and L2, there is a range of mitigating strategies available to learners. For example, given an issue such as a WF voicing contrast, speakers of different L1s may develop a strategy which allows them to align more closely with their individual L1. For example, while both

Spanish and Mandarin do not allow voiced WF obstruents, Eckman (1981) showed how Spanish speakers devoiced them and that Chinese learners added a schwa. To relate this to the current study, both Vietnamese and English have initial fricatives, including /s/ and /z/, and these do not give rise to significant errors in learners' speech, except where /s/ appears in a cluster. In contrast, English has WF fricatives, but for VLEs, /s/ and /z/ are 'marked' word- finally, and absence is unmarked; it is here that the errors are found.

Moving on from L1 effects, the interaction between rule learning, constraints, lexical entrenchment of frequent forms, and the need to manage cognitive load during communication efficiently has resulted in a field that is increasingly integrative in its approach to SLA and learner error. For example, Optimality Theory (Hayes et al. 2003; McCarthy 2002; Prince and Smolensky 1993, 1997) argued that there are universal constraints that can differ in rank between languages, and a learner has to re-order the relevant constraints in order to become more target-like. Major (2001) developed the Ontogeny and Phylogeny approach, which links effects from the L1, the target language, markedness, and universals, as well as embedding task variation. This revealed how accuracy increases proportionately to stylistic formality, mirroring Labov's seminal work on style shift (Labov 1972, p. 208). Major (2004) further found that the influence of style was variable across language groups; while the influence is strong for native speakers of English, for Japanese speakers, style as a variable was not significant. The potential influence of style as a variable has informed the current study's use of the middle ground of relatively formal running, monologue speech, where we might expect higher accuracy than in casual conversation but also less attention to form than found in word lists. Other recent work has integrated many diverse approaches, such as computational and corpus linguistics (Albright and Hayes 2004), linguistic typology (Haspelmath and Sims 2010), and neural network models (Goldwater and Johnson 2006). Finally, wholesale transfer effects are being questioned by research into transfer in L3 students. For example, in the Linguistic Proximity Model (Archibald 2022; Westergaard 2021; Westergaard et al. 2017), it is argued that transfer occurs property-by-property, based mainly on structural linguistic similarities between previously learnt languages.

1.2. Frequency Effects

In modern usage-based approaches, frequency of input and use is the focus of attention in both L1 and L2 research, and we are "exquisitely sensitive" to frequency (Taylor 2008, p. 46). In L1 acquisition, type and token frequency are critical (see Lieven and Tomasello 2008). In L2 acquisition, frequency of input and repetition by the learner are both key factors in successful L2 acquisition (Bod et al. 2003; Bybee and Hopper 2001). Frequency has consistently been shown to strongly correlate with the order of acquisition of morphemes and lower error rates (Bybee and Hopper 2001; Ellis 2002; Ellis and Collins 2009; Ellis and Schmidt 1997; Larsen-Freeman 1975, 1991, 2002). Recent studies have argued that frequency of use has a key role to play in overcoming L1 transfer effects and might help explain why the models fail to accurately predict errors (Piske et al. 2001; Flege 2009).

As Ellis (2013, p. 196) points out, "Frequency is a key determinant of acquisition because 'rules' of language, at all levels of analysis from phonology, through syntax, to discourse, are structural regularities which emerge from learners' lifetime analysis of the distributional characteristics of the language input. Statistical and connectionist approaches to language learning theory have been developed that show how this analysis by the learner might work (Bates and MacWhinney 1987; Bybee and Hopper 2001; Colantoni et al. 2015; Ellis 2013; Trofimovich et al. 2012; Yang and Bod 2008). Learners appear to make constant comparisons between the input and their output and store statistically based weightings for many slightly different variants of a given variable. These are then available for selection during speech output. Learners make subconscious micro-adjustments to their speech output so that it becomes progressively closer to the input, or at least close enough that communication is relatively smooth. As with most language learning, this requires communicatively meaningful contexts.

As the learner is exposed to more target-like input, the higher frequency variants should have an increasing influence on their output, and less target-like output should slowly decrease. Under such a statistical learning model, development is predicted to be piecemeal and follow general learning patterns. Such an approach implies that their output can vary because diverse ways of producing a variant are always available for selection, even if they are weakly weighted and do not closely match the input. Within this process, learners may develop a strategy such as ignoring minor differences in form that do not seem to make a difference to the system, regardless of L1/L2 contrasts, probably to reduce cognitive load (Sweller et al. 2011). This process then may not actually result in a consistently more target-like output (Fullana and Mora 2009), and it is this process of monitoring the comparisons and making micro-adjustments that L1 interference interacts with. Indeed, these interacting processes may lead to the entrenchment of non-target-like forms. For example, this process could account for the persistent lack of voicing contrasts in our data, even by speakers with otherwise high accuracy. It is also becoming clearer how target features can be acquired at different rates, such as the feature-by-feature approach to L2 learning (Archibald 2022; Westergaard 2021).

The final consideration is to point out that a learners input is also likely to be highly variable. For adult learners of English in the internet age, there are vast amounts of mass media input freely available, all of which can have an effect on language learning (Bohn and Bundgaard-Nielsen 2009; Piske and Young-Scholten 2008; Leung 2014).

The standard English morphological system that forms the bulk of the input is not complex, and there is little difference between standard UK and US Englishes. While it is consistent, some measurements typically do not have plural marking (as in *three pound twenty* for £3.20), and some dialects may differ in distribution. The morphemes that result in word final /s/ and /z/ are the remnants of a more complex inflectional system in Old English that gradually disappeared over time. Typically, language change does not happen at the same rate in all dialects, and some will likely retain older forms or may develop in a divergent way from the standard. Some dialects of English do have nonstandard verb and noun forms that have no morphological {-s}, or where the distribution is different to Standard Englishes (Trudgill 1974, 1992). While it is not possible to determine with any certainty what the input consists of for a given learner, strong regional dialects and accents are unlikely to form any significant part of the input to the students in a place like Vietnam. Where a native speaker does not have a {-s} that is required in the standard variety, it is likely due to this. Note that such a speaker would not omit non-morphological instances as found in VLEs, where the /s/ in *ice* is susceptible to omission.

We therefore assume that the teachers' speech and UK/US-informed mainstream media English form most of the input and are the main influences on the speech forms of the learners in the study. We acknowledge, however, that the teachers of these participants are mostly VLEs themselves, and the extent to which they exhibit similar omission patterns is unknown. Further, the learners may be accessing much non-native speaker content from elsewhere, such as China, adding more uncertainty to assumptions concerning their input.

1.3. Elision and Omission

The omission of phonemes is a complex area of study, not aided by some significant variation in the application of terminology, particularly where the terms omission, deletion, and elision are concerned. In sum, elision is a type of omission of sounds that usually refers to the phonological process of simplification during running speech, e.g., /d/ elision in /nd/ clusters in strings like *brand new*. In contrast, omission is a broader term used to capture cases beyond those described as elision, for instance, where a child says /bɪ/ for *big*. The term *deletion* appears to be used in the literature where the distinction in a given case is less clear.

Elision itself is a common feature cross-linguistically and occurs in rapid speech, in unstressed syllables, and often across word boundaries (Collins and Mees 2013; Cruttenden 2014). It is also a typical feature of children's language (McLeod et al. 2001; Torkildsen

and Horst 2019). Elision of phonemes in clusters occurs word-initially (Schreier 2005a), but is more common in word-final clusters, occurring in native varieties (Schreier 2005b; Zsiga 2013) and non-native varieties (Eckman 1987; Edwards 2011; Gut 2007), including Vietnamese English (Osburne 1996; Nu 2009). Clusters of three consonants are highly prone to elision, where the second consonant is routinely elided, both within words (e.g., *acts*) and between words (e.g., *act six*) and causes problems for L2 learners (Temperley 1983). In terms of specific phonemes, elision of /t/ and /d/ is a widely found feature of American speech, especially following /n/, as in /tweni/ for *twenty*. It also increasingly occurs in other native speaker English varieties (Amos et al. 2020; Baranowski and Turton 2020; Colantoni et al. 2015; Moran 1993; Raymond et al. 2016). Word final /s/ and /z/ elision in native varieties is not common, but may occur as part of the three consonant cluster pattern across word boundaries or in certain dialects (Trudgill 1997). The omission of consonants occurs in Vietnamese learners' speech even when words are spoken slowly, are fully stressed, and are said in isolation. It is, therefore, the preferred term for this process in the current study. Typical cases one can hear in everyday usage are [ɾaɪ] for *rice*, [sɪ] for *six*, and [masa] for *massage*.

1.4. The Vietnamese Context

Vietnamese is a Vietic language within the Mon-Khmer branch of the Austroasiatic family (Kirby 2011) and is closely related to other languages in the geographic area, such as Lao, Thai, and Cambodian (see Alves 2006). Due to extensive historical population movements and fluid, sometimes rather notional, territorial borders, the languages within the Sino-Tibetan and Austroasiatic families are likely to either contain elements of a common protolanguage or have significant areal influence in their development (Blench 2015). As such, they have converged somewhat over time, and there are similarities throughout the region in syntax and phonology where, for example, they all have similar phonological tones (Blench 2015; Enfield and Comrie 2015). It may therefore be unsurprising to find similar patterns in the English of people in this area, and word-final omission of some members of the English consonant inventory is a good example of this. One of the major existing resources for teachers on learner English is Swan and Smith (1987, 2001), which describes common patterns found in the English of learners from different L1 backgrounds. It is telling that they provided a description of Vietnamese learners in a dedicated chapter in the first edition (Swan and Smith 1987), but in the following edition (Swan and Smith 2001), Vietnamese was conflated within a single SE Asia section. This clearly reflects the fact that similar issues exist across the region, with WF omission being a core example.

Vietnamese has a smaller consonant inventory than English, and therefore, there is scope to contrast the two languages to try and predict likely errors. Vietnamese “licenses eight segments in coda position: three unreleased voiceless obstruents /p t k/ ([p̚ t̚ k̚]), three nasals /m n ŋ/ and two approximants /j w/” (or semi-vowels, Kirby 2011, p. 383). The fricative inventory is highly uniform across the country and contains initial fricatives /f v s z x ɣ h/ (Kirby 2011, p. 382). Fricatives are not permitted word-finally. English allows more clusters in onsets and codas than Vietnamese, so again, one would expect any transfer effect to apply in these contexts (Tang 2007). Honey (1987, p. 240) describes the “interconsonantal” omission of /s/ and states that final /s/ “when following a consonant is frequently omitted. In terms of codas, English allows a nasal + /z/ or /s/ as well as obstruent clusters like /ts/ and /dz/. Note that consonant clusters are also well known to cause issues for many L2 learners (Altenberg 2005). As a final point, Vietnamese is not a single homogenous language, and there are various regional dialects and accents (Honey 1987; Hwa-Froelich et al. 2002; Kirby 2011; Nguyễn-Đặng-Liêm 1970). The Saigon variety does not have initial /v/, for example. This regional variation has no impact on the current study because there are no final fricatives in any Vietnamese dialect, although people from the North, and Hanoi in particular, have been shown to be less likely to elide word-final consonants (Hwa-Froelich et al. 2002, p. 270).

There are differences between previous studies in terms of what sounds this omission affects. For example, [Hwa-Froelich et al. \(2002, p. 265\)](#) claim that VLEs do have final voiceless plosives /p/ /t/ and /k/ but that there is a tendency towards “deletion” (*ibid.*, p. 271). In contrast, [Honey \(1987\)](#) says they are present but are “unexploded”. It is unclear if this means that they undergo lenition to [ʔ] or are unreleased, as [Kirby \(2011\)](#) argues. Observationally, I see little lip closure when there is a target of /p/ and /b/ finally, but the airflow is stopped. What occurs between words, where a stopped airflow must resume, for example, is less clear. In English, the voiceless stops are articulated with a glottal stop, at least in aspirated positions, but this, in turn, may also not occur in Vietnamese ([Kirby 2011](#)), which adds to the analytic challenges. This point is not inconsequential because previous studies have reported fricative stopping as a simplification strategy, albeit in small-scale studies using word lists with low token frequencies. It has been argued that final /f/ and /v/ are stopped to /p/ ([Hwa-Froelich et al. 2002, p. 268](#)), while /θ/ and /ð/ ([Bui 2016, p. 126](#)) and the affricates ([Loi 2018](#)) were stopped to /t/. Note the voicing neutralisations here, supporting our position to conflate targets /s/ and /z/, as described below. Also, given these conclusions, it might be argued that /s/ and /z/ are also stopped to /t/, which then undergoes lenition to [ʔ]. The naturally occurring data samples used in the current study are not clear enough to more accurately determine omission from lenition to [ʔ].

The two mismatches in distribution between English and Vietnamese, consonant clusters and word-final fricatives, between Vietnamese as an L1 and English as an L2 are potential areas for negative L1 transfer to affect pronunciation. Indeed, [Honey \(1987, p. 240\)](#) overtly states that this contrast “gives rise to mistakes”, which are “very difficult to eliminate”. [Hwa-Froelich et al. \(2002, p. 271\)](#) state that because of the lack of word-final consonants in Vietnamese, it is “usually difficult” for Vietnamese learners to pronounce them. [Nu](#) states that final consonants in Vietnamese are “never pronounced” and that this is “the reason why Vietnamese learners of English often omit final consonants of words in English.” ([Nu 2009, p. 44](#)). This perspective firmly frames dissimilarity between the phonological systems as the root cause of the issues. However, supposed transfer effects are highly nuanced, particularly in the more advanced examples of interlanguages found in our participants. Articulation is assumed to be easier when moving from a vowel to one of these coronal fricatives, as opposed to moving from another consonant. so for instance, an observed effect may be the result of unfamiliar articulatory complexity and not transfer *per se*.

1.5. Concluding Remarks

The focus of the current paper was restricted to WF /s/ and /z/ in order to focus on the lexical/morphological distinction more closely. The interaction between English word-final morphology and deletion of consonants by VLEs is not discussed in previous papers, except for a brief note in [Hwa-Froelich et al. \(2002, p. 271\)](#), who cite [Sato \(1990\)](#). However, this is perhaps unsurprising because these morphemes are mostly /s/, /z/, /t/, and /d/, and as such, they are already covered by the claimed transfer effects (i.e., word-final deletion as a phonological process and the lack of word-final morphology in Vietnamese). We excluded the plosives /t/ and /d/ from the study despite their similar status as being WF morphemes and appearing word-finally in lexical words. This was partly due to issues in perceptual salience in the authentically recorded data and partly due to indeterminacy as to whether they are actually omitted or if they are replaced with a glottal stop, which is lenition, not omission. Superlatives with {-est} morphemes were very rare in the data and were not considered for analysis. The only other consonant morpheme in English is /ɹ̥/, which is not routinely deleted in the speech of VLEs. In fact, word-final nasals /n/ /m/ and /ŋ/ all occur finally in both languages and none are typically omitted in VLEs. For the sake of completeness and to exclude the rest of the English consonant inventory from the study, the approximants /j/ and /w/ do not occur word-finally in either language, although their status as consonant or vowels is debatable ([Kirby 2011](#)). WF /l/, where there is a target [ɫ], tends to be vocalised, not omitted. Finally, some accents are rhotic and

postvocalic rhoticity where it occurred before WF target /s/ and /z/ was classified as an r-flavoured vowel and not as a consonant in our analysis.

This section has described the background and the theoretical context for the current study. Given the paucity of existing research and the dearth of targeted materials available to aid teachers in their attempts to mitigate the issues, there was clearly a need for evidence-based research [Derwing and Munro \(2015\)](#), which this study provides. We believe that the results from this study may be reliably extended to capture the likely patterns found throughout the country and, to an extent, in the wider SE Asia area.

2. Materials and Methods

This section outlines the methodology used in the collation and analysis of the data and states the hypotheses. The specific aims of the research were to describe the variation in error patterns of /s/ and /z/ within the population sample and to quantify the relative effects of phonological context, morphology and frequency of use on observed error rates. The overarching goal of the research was to add empirical evidence to these classroom-based observations with the intention of informing pedagogy. The null hypothesis H0 against which to test the data is that there are no significant patterns in the errors of /s/ and /z/ in word-final position. Three further hypotheses, H1-H3 below, were addressed:

- H1: Error rates are sensitive to morphological context; we predicted that morphological {s} realised as /s/ and /z/ would have more errors than vs. lexical /s/ and /z/.
- H2: Error rates are sensitive to the preceding phoneme; we predicted that error rates would be higher with a preceding consonant than with a preceding vowel.
- H3: Error rates vary in direct proportion to the relative frequency of occurrence in the data, i.e., we predicted that higher-frequency words would have fewer errors.

We predicted that we would confirm H1, H2, and H3 and that context, morphology, and frequency of use would all have a measurable effect on accuracy. We, therefore, predicted that we would reject the null hypothesis.

Participants

The participants had just completed a four-year dual award BA undergraduate programme in Hanoi validated by a UK university. They were studying to be English language professionals such as teachers, translators, and interpreters. They have all achieved a minimum equivalent of C1 for their English language in terms of the Common European Framework of Reference for Languages (CEFR). Ethical approval was awarded by the lead researcher's university ethics panel.

In the course of their degree, the students conducted a fifteen-minute individual presentation on their final graduation paper, which was video recorded for quality assurance purposes. These form the data source for this paper. Note that the recordings were not collected for the purposes of the research, so they are technically "naturalistic" data. Due to the exacting standards expected from the UK and Vietnamese degree-awarding universities, it is reasonable to suppose that these students have some of the most advanced EFL English skills achievable by the time they complete their degree, and this is the main reason we chose a sample from this demographic group. Further, due to the formality and importance of the event itself, and the fact they were graded on their accuracy of speech, it is safe to assume that the students were aligning as closely as they could with high-prestige norms and that their language is, therefore, high on [Labov's \(1972\)](#) "attention to form" scale. If even these people exhibit extensive elision of WF /s/ and /z/, then for less advanced learners, the vast majority, the situation must be even more challenging.

Consent was requested from the entire cohort ($n = 38$), and those who completed the form were considered eligible for inclusion ($n = 32$). We selected for final inclusion all of the samples that were of usable audio quality ($n = 16$), which is an essentially random sample because the quality of the recording was random. The individual recordings were on different topics but are highly comparable because the task type, the purpose and the audience are almost identical in each instance. This level of homogeneity in the data is

important when drawing meaningful conclusions and generalisations. An online tool (Otter.ai) was used to create text from the audio files, and the outputs were anonymised and corrected. It was important to preserve their anonymity because the study explicitly identifies errors in the English of people who are budding English language professionals.

Once names, fillers, unintelligible words, and Vietnamese words were removed, there were 25,000 usable words (range: 944–2662) representing 240 min of authentic spoken data. A total of 3858 words had a target of /s/ or /z/ (mean 241 = 15.4%: range 184 = 12.4% to 337 = 21.7%). Annotations for the variables (lexical/morphological and preceding consonant/vowel) and the accuracy of /s/ and /z/ pronunciation were conducted by the lead researcher, a native speaker of standard southern British English. The coding of the speech was made based on judgements on the presence of WF /s/ and /z/ perceptually, with the assistance of speech analysis software (PRAAT) in a few instances, to see if there was evidence in the spectrogram of a short fricative, as opposed to nonrelevant background noise for example. The same coder worked on the entire data set to maintain consistency and reliability. Target words with final /s/ and /z/ were compared to actual pronunciations and annotated for accuracy (where /s/ and /z/ are \pm absent), preceding sound (consonant or vowel) and morphology (where /s/ and /z/ are \pm morphological). Possessive clitics were counted as morphemes, but these were rare. Samples were double-checked, and the error rate after the first analysis run was < 1%.

As previous studies have found, while Vietnamese speakers of English tend to omit WF /s/ and /z/, where they do produce them, voicing contrasts are neutralised. This may partly be because they are so rarely contrastive in that position in the input and, therefore, are not significant enough to warrant sufficient attention for more accurate realisation. Where a word-final alveolar fricative was salient in our data, it was not possible to aurally determine whether a target of [z] was voiced, devoiced or voiceless. There are three points here. For native English speakers, there is a voicing harmony effect with WF /s/ /z/ /t/ /d/ when they are morphological, which is a next-level skill for VLEs. US and UK English speakers also typically shorten the vowel before voiceless codas (pre-fortis clipping; see Wells 1990), and voiceless fricatives are usually longer than their voiced counterparts, although there is much variation in style-shifting contexts (Maniwa et al. 2009). These are important cues for children learning an English L1, but none of these factors were observationally salient or measurable in these data to assist in the identification of voicing contrasts. Ascertaining voicing in the data samples was, therefore, impossible, and because the error rate in our data for targets [s] and [z] are so similar (27% and 29%, respectively), it appears that the voicing status of the target is not a significant factor in accuracy. Therefore, for the purposes of the current study, the results are conflated and the use of [s] where [z] is the target is not considered an error. For the sake of brevity, the combined phonemes in the data may be referred to as /s~z/ below.

Finally, to establish how important this feature is and why describing and mitigating errors is necessary for L2 learners, we analysed the British National Corpus (BNC Consortium 2007) spoken component to use as a frequency benchmark. Following the removal of non-word annotations (e.g., # and discourse markers *oh hmm*), there are 10,035,525 words. Of these, 1,061,520 (10.5%) tokens and 13,185 different words (21.5%) end in /s~z/. This clearly shows that focussing on errors in these phonemes is an essential task. We observed that just over a third of the tokens in the BNC data are accounted for by 's and *is* and a quarter more by *was*, *this*, *as*, *yes*, *has* and *does*. Together, these eight forms account for over half of WF /s/ and /z/. As the results below reveal, these words exhibit low error rates when they are frequent in the VLEs data, which is why it is worth noting the high frequency of use in native speaker speech at the outset.

3. Results

All of the participants had more present than absent WF /s~z/ compared to the expected (target-like) form. Table 1 shows that the overall error rate was 28.4% and that, across the participants, the range of errors was 11.5% to 48.8% (see Appendix A for the full

results). This highlights both the size of the problem and the extent of the variation at the level of the individual.

Table 1. Overall accuracy results of all 3858 instances.

| +/- Present | All /s/ and /z/ | Lowest | Highest | Lowest % | Highest % |
|-------------|-----------------|--------|---------|----------|-----------|
| Present | 2771 (71.6%) | 105 | 261 | 51.2% | 88.5% |
| Absent | 1087 (28.4%) | 24 | 115 | 11.5% | 48.8% |

Table 2 shows the contrast between the error rates for morphological and lexical /s~z/ as a percentage of the frequency of their respective target. The error rate is more than double for morphological /s~z/ than for lexical tokens, with a 28.3 percentage point difference.

Table 2. Error rates for morphological and lexical /s~z/.

| | Total /s~z/ | Lexical /s~z/ | Morphological /s~z/ |
|-----------------------------|-------------|---------------|---------------------|
| Total Instances | 3858 | 2219 | 1639 |
| Instances as % of all /s~z/ | 100% | 57.5% | 42.5% |
| Mean error rate | 28.4% | 16.8% | 45.1% |
| Std. Dev | 10.8 | 12.4 | 20.8 |

Turning now to the issue of phonological context, Table 3 shows the contrast in error rates for /s~z/ with either a preceding consonant or vowel, also as a percentage of their respective target frequencies. All participants except one had more errors when the target was a consonant cluster than when there was a preceding vowel.¹ The error rate is also doubled when /s~z/ is preceded by a consonant compared to when it is preceded by a vowel, with a 27-percentage point difference in accuracy.

Table 3. Error rates for preceding phoneme type (consonant or vowel).

| | Total /s~z/ | Preceding Vowel | Preceding Consonant |
|-----------------------------|-------------|-----------------|---------------------|
| Total Instances | 3858 | 2650 | 1208 |
| Instances as % of all /s~z/ | 100% | 68.7% | 31.3% |
| Mean error rates | 28.4% | 19.7% | 46.7% |
| Std. Dev | 10.8 | 11.1 | 18.1 |

A multilevel binary logistic regression was performed to ascertain the effects of word type (lexical/morphological) and preceding phoneme type (consonant/vowel) on the likelihood that words containing a /s/ or /z/ would be pronounced with the /s~z/ absent rather than present. The regression model included word type and preceding phoneme type as fixed effects, an interaction effect of word type by preceding phoneme type, and a random intercept to account for data clustered by participants. An overall F-test indicated this model was a significantly better fit than a null model ($F(3,3744) = 126, p < 0.001$), and it correctly classified 75.4% of cases. In terms of specific results, morphological /s~z/ cases were significantly more likely to be pronounced with the target /s~z/ absent than lexical words containing a /s~z/ ($OR = 3.59, 95\% CI(2.55, 5.05), p < 0.001$). Words with a consonant immediately preceding a /s~z/ were significantly more likely to be pronounced with the /s~z/ absent than words with a vowel immediately preceding a /s~z/ ($OR = 2.14, 95\% CI(1.70, 2.68), p < 0.001$).² These results support confirmation of H1 and H2.

To summarise thus far, as one would expect from simply observing an L2 learner from this demographic, accuracy is higher when WF /s~z/ is preceded by a vowel than when it is preceded by a consonant, and accuracy is higher when WF /s~z/ is lexical than when it is morphological. However, we are now able to quantify this effect; there is an apparent “accuracy penalty” of 27–28 percentage points for each.

The next level of detail to discuss is the combination of these features. In other words, what are the accuracy rates for lexical /s~z/ compared to morphological /s~z/ when they are preceded by a vowel or by a consonant? Table 4 shows the error rates as a percentage of the target for the given combination, e.g., it shows the tokens that were lexical and preceded by a vowel as a percentage of all lexical/vowel tokens.

Table 4. Error rates for combinations of the variables as % of the total for the combination.

| Variable Combinations | Mean | Std. Dev |
|--|-------|----------|
| Lexical /s~z/ preceding vowel | 15.3% | 13.1 |
| Lexical /s~z/ with preceding consonant | 32.5% | 19.3 |
| Morphological /s~z/ with preceding vowel | 36.4% | 22.7 |
| Morphological /s~z/ with preceding consonant | 50.7% | 21.7 |

Based on the results thus far, we predicted that accuracy would be lowest when morphological /s~z/ is preceded by a consonant and highest when a lexical /s~z/ is preceded by a vowel, which was confirmed; an average of 50.7% of possible /s~z/ instances were not pronounced where a morpheme was preceded by a consonant. This contrasts with the error rate of 15.3% for the opposite condition, where a lexical instance was preceded by a vowel. Note that all but one participant had more errors on the morphological + consonant measure than any other.³

Recall that we observed a 45.1% error rate for morphology and a 46.7% error rate with a preceding consonant. These results strongly suggest that both morphological status and preceding phoneme are significant factors in the accuracy of /s~z/ realisation in these data. However, when the variables are combined (i.e., the error rate for tokens that are morphological with a preceding consonant), they exhibit only a slight increase in error rate to 50.7, suggesting that they are not interacting strongly. A multilevel binary logistic regression was performed on the whole data set to ascertain the potential interaction between the variables. No significant interaction was observed between word type and preceding phoneme type on the likelihood that words containing a /s~z/ would be pronounced with the /s~z/ absent or present (OR = 0.88, 95% CI (0.59, 1.31), $p = 0.519$). We conclude that performance on the lexical/morphological measure does not predict performance of the vowel/consonant measure.

We examined the data for more detail, despite some low number of instances for the specific consonants and found little difference between preceding nasals (44.3% error rate) and non-nasals (43.3% error rate). Place of articulation is a little more interesting. The most accurate cluster is /ts/, with an error rate just under the overall mean of 24.3%. This might be thought predictable because the former is homorganic with /s~z/, and the others can involve a change in place of articulation. However, such logic cannot account for the preceding target /d/ error rate of 49.5% or the error rate with preceding /p/ of 29.5%.⁴ Finally, a preceding homorganic alveolar /t/ /d/ or /n/ in a consonant cluster with /s~z/ has a lower error rate (41.2%) than a preceding bilabial or velar (51.6%). This is presumably due to the increase in articulatory effort required to shift the place of articulation, but it is not a great difference.

Frequency and Accuracy

The relationship between frequency of use and error rates is not as clear. Firstly, it is only possible to measure output and to draw inferences concerning the input based on corpora like the BNC, as we showed above. However, given that high-frequency function words in the output are likely to be high-frequency in the input, it is worth investigating this issue here. The error rates across the participants weakly correlate with the overall frequency of /s~z/ in each sample ($r = 0.34$). In other words, there is a weak correlation between higher overall frequency of use of /s~z/ and higher accuracy. However, correlations are much stronger when only higher-frequency words and accuracy are compared; for example, words with a token frequency of ≥ 8 returned $r = 0.996$.

The data reflect the general tendency for Zipfian distributions to occur in natural language. There were 593 different words in the data with a WF target /s/ or /z/, 27 (7%) of which account for 50.1% of the total instances, with an error rate of 20%. For low-frequency words, the situation is more variable. The majority were single or dual examples, so correlations between frequency and accuracy at word level were not feasible. However, we saw that those words with ≤ 3 instances accounted for 33.3% of the total errors and only 17.2% of the total target-like examples, suggesting that low-frequency words are more prone to error.

The relationship between accuracy and frequency is most clearly revealed in function words ending in a vowel + /s~z/, where those with relatively high frequency are also high in accuracy. Two words alone, *is* and *this*, account for 24.8% of all /s~z/ examples in these data, with an error rate of only 12%. Other examples include auxiliary *has* (29; 10% error rate), *as* (83 tokens; 14%), auxiliary and copular *was* (75; 12%) and *because* (78; 7%), a total of 262, with an overall 11% error rate. When combined, these six words account for 31% of all WF /s~z/ in these data and 38% of all target-like examples. Note that all six have a preceding vowel, which is also indicative of higher accuracy. The effect of these high-frequency examples is most starkly revealed by excluding them. If *is* and *this* are excluded, the total error rate rises from 28.4% to 34.4%, the preceding vowel error rate rises to 24.7%, and two participants would have more errors than target-like examples overall. When all six words are excluded, the overall error rate rises to 41%.

While these results are clear, the numerical data do not do justice to the variability in the data at the individual level. The topics the participants chose were all different and contained much repetition, and as a result, high-frequency major/lexical class words with /s~z/ endings are more variable. For example, the fifty examples of the word *use* had a 10% error rate (noun and verb examples were conflated), of sixty-nine *US* (said as /ju: es/) tokens, 14.5% were errors, and of forty-seven *strategies*, 17% were errors. In contrast, 36.5% of seventy-four *students* have no WF /s/, which reflects the trend for omission where morpheme {-s} is in a cluster. In addition, some words had high use by single participants. For example, *Vietnamese* was used sixty-two times with an error rate of 27.4%, but this was skewed by P1's high error rate. There were thirty-eight examples of *words*, but a single participant accounts for twenty-six of these, with twenty-two errors. Finally, the word *questions* occurred forty-seven times with an error rate of 72.4%. This usually occurred in a phrase similar to (presumed target) "If you have any questions (...)" and was consistently categorised as a target of *questions*. We accept that this could be a result of the speaker failing to correctly distinguish between the determiners *a* and *any*, so the target may have been singular and not plural and, therefore, not an error.

When we looked at multiple uses of a given word, while one assumes the participants were monitoring their speech for error, they did not simply utter an incorrect version of the target and then self-correct, which would have been evidence of an active and finely tuned self-monitoring system. By way of example, we closely analysed one participant, P4, who is highly representative of the sample as a whole. They are lower than the mean for a preceding vowel (6.2%, but closest to the mean of all the participants for the morpheme examples with a preceding consonant (52%). They are second closest both overall (25.1%) and when /s~z/ is preceded by a consonant (48.7%). They are consistently close on all other counts and also align well with the means for the high-frequency minor classes described above. It was interesting to find that there were no obvious patterns in multi-use lexical examples, with errors seeming to be rather random. For example, P4 used the target word *questions* six times, three without /s~z/ () and three with (+), in the order -+---. With their twelve tokens of *students*, seven were errors in the pattern +++++---. By coincidence, P4's topic was grammaticality in the speech of some L2 learners, and, ironically, the word *mistakes* was used sixteen times, all without /s/ at the end. Note that *questions* and *students* have preceding alveolar plosives, while in *mistakes*, there is a preceding velar plosive, which requires a change in place of articulation and is more prone to error. It is also intriguing that while they, like all the participants, were able to use their presentation

slides as prompts, mistakes were still copious even when an on-screen instance of a target /s/ or /z/ was present.

4. Discussion

In L2 data, one overriding theme is usually the sheer amount of individual variation that exists and which confounds attempts to use transfer effects to account reliably for observed errors. Some speakers align with predictions, while others rather stubbornly do not! The previous research has a small and largely anecdotal evidence base for the claims made on WF /s~z/ elision for VLEs. Simply stating that these word-final phonemes are “deleted” is inadequate, and such a general point is not particularly useful for ELT pedagogy, at least when it comes to helping high-level learners mitigate their problems. Our study has added evidence that supports many of the previous claims by using 3858 examples of WF /s~z/ taken from the running monologue speech of sixteen Vietnamese learners of English. The overall picture drawn by these results is not exactly surprising, but we were surprised by the extent of the variation and substantial error rates we found in these high-level learners.

These data have allowed some interesting evidence-based generalisations to be made. We found an average error rate of 28.4% and demonstrated a range of results across the different variables and their combinations and we believe that H1, H2, and H3 are all supported by the results. Thus, we can reject the null hypothesis. The results included how /s~z/ in more marked contexts, i.e., the morphological and consonant cluster examples, have more errors than those in less marked contexts. When these were combined, there was a 50.7% error rate, which is just 4–5 percentage points higher than the individual rates. This suggests that the issues are stacked rather than interacting and that the combination of cluster + morpheme is not creating additional problems for learners. The results from the multilevel binary logistic regression also support H1 and H2. This serves to highlight the difference between the general tendency for /s~z/ elision as a phonological process, in contrast to the omission of a morpheme as a grammatical construction-forming process. It also follows that it would be impossible to ascertain a phonological or morphological cause for a given case of C+ morphological {-s} elision.

We found a strong correlation between frequency of use and accuracy; high-frequency words are more accurately realised than less high-frequency words, which supports H3. For example, when six high-frequency function words were removed, we found an overall 41% error rate, which is almost four times higher than these high-frequency words taken on their own. that

However, there are methodological issues that may have affected the data. Firstly, the frequencies of forms in spoken *output*, as measured here, may not be directly proportional to their frequency in *input*, the primary locus of learning. This is an important consideration given the usage-based approaches described above and the connectionist view that people make constant comparisons between the input and output during learning. This is why, for example, merely hearing language from, e.g., watching TV, does not result in successful language learning for either L1 children or L2 learners, as speech in meaningful communicative contexts is essential (see [Lieven and Tomasello 2008](#)). In the current study, this input may have been from a native speaker teacher, a non-native speaker teacher who may or may not also exhibit elision, and/or from native speaker mass media ([Leung 2014](#)).

The second point on frequency effects, as opposed to other influences, arose in the accuracy comparison between nouns and verbs with bound morpheme {-s}. The morphological examples consist of 1511 plural nouns and only 112 verbs (excluding copular and auxiliary verbs) with respective error rates of 41.7% and 58%. This particular difference in error rates is not wholly accounted for by either grammatical or phonological transfer effects. The consistency of application, a form’s relative uniqueness of use and its salience also appear to have an effect on learning, with cue validity and contingency effects forming a significant part of the discussion on the variability found in SLA (see [De Villiers and Johnson 2007](#); [Ellis 2008](#)). Clearly, nominal morpheme {-s} has a more consistent application

than the verbal morpheme {-s} because it applies in almost every case of a plural countable noun, while only one out of the six subject types (third-person singular) triggers verbal {-s}. Further, there are three forms for this morpheme [s], [z] and [ɪz], and several different uses, plural, verbal, possessive, and clitic, which together also serve to confound predictability and hence increase non-target-like variability. Ellis argues that “a contingency analysis of these cue interpretation associations suggests that they will not be readily learnable.” (Ellis 2008, p. 376). It can also cause problems during child language acquisition (Kueser et al. 2018). This difference also potentially highlights a frequency effect, but the limited sample size of verbs prohibits further investigation.

The results with respect to frequency of use strongly imply that targeted mitigation work should be employed in the EFL classroom to help students focus on this tricky area of speech, if only to raise the frequencies in their input (VanPatten 1996 cited by Ellis 2008, p. 389) and attempt to raise their salience to improve accuracy (Doughty and Williams 1998, p. 220, cited Ellis 2008, p. 389). As Ellis points out (Ellis 2008, p. 394), “[a] sad irony for an L2 speaker under such circumstances of transfer is that more input simply compounds their error; they dig themselves ever deeper into the hole begun and subsequently entrenched by their L1”. The implication from these results is that if the frequency of exposure and use is critical, then the input needs to be consistently target-like in order for the output to become more target-like, assuming this is the goal of the learner. Clearly, high levels of input that exhibit a natural tendency towards omission of WF /s/ and /z/, as well as other persistent errors, will affect the students’ statistical modelling and, therefore, their output. We do not at all agree that this implies that only native speakers should be teaching and modelling the input to language learners, and we also do not agree that native speaker norms are the holy grail of all L2/3 language learners. To do so would run the risk of accusations of “nativespeakerism” (Holliday 2006). On the other hand, because omission of /s~z/ is so frequent in speech and can have a significant effect on a hearer’s comprehension, we would argue that teachers do have a certain responsibility to ensure that the input to the students is as close to the target norms as possible with word-final obstruents in general and especially with /s~z/. We would also encourage native speakers to become more aware of these issues and try to fill in the gaps when hearing the speech of people who exhibit WF omission of consonants.

In terms of methodology, the study could not control for whether the output was created ‘online’, was read from a script, or was semi-prompted by notes, and so on. It is certainly true that during their presentations, the participants could read from their presentation slides if they wished, and in fact, these aided the transcriptions in cases of lexical indeterminacy. There were no observable errors with morphological {-s} on the slides, which implies that the tendency for elision overrides visual input from reading-aloud tasks. It may also be the case that the input to the students has consisted largely of written language. However, because the participants chose their own topics and content and had been working with these words for months, it is not as if they were lacking the opportunity to rehearse. These points form part of the continuing research agenda.

5. Concluding Remarks

The present study is situated in the analysis of data rather than in any attempt to refute or support certain theoretical perspectives. Notwithstanding this, the patterns revealed by the data, in general, do support the view that frequency is a key variable, as would be expected in a connectionist approach to learning. We discussed how various factors can potentially affect accuracy, such as L1 influence, frequency, markedness at the phonological and morphological levels, as well as ease of articulation. Individual interlanguage instability and perhaps task-type effects are in play as well. Entrenchment of forms with errors occurs frequently and may be hard for an individual to overcome. While we can observe these patterns, it is impossible to tease apart the relative effects with any *predictive* certainty, as the history of the field suggests. However, the current paper

has succeeded in quantifying some general tendencies that do have some good predictive power, with some meaningful results within a robust methodology.

This paper focuses on the quantitative results for an analysis of authentic L2 data but sits within a project that has the aim of forming a strong base of evidence to aid in the development of teaching materials. Further empirical study into the effects of targeted mitigation in the classroom is therefore required. From additional observations not presented above, it is clear that it is not WF /s~z/ that is ‘the problem’ *per se*, but the issue concerns syllable coda /s~z/. Indeed, the other fricatives and the affricates are also frequently omitted. The data here were not large enough for analysis of other interesting avenues of analysis in the speech of VLEs, such as the effect of a following word starting with /s/, another alveolar or coronal, for example. Studies could look at voicing contrasts, fortis plosive aspiration, and pre-fortis vowel clipping, for example. We did not investigate the issue of syntactic context, e.g., whether accuracy is affected when the word is phrase or utterance final because we did not observe an obvious effect in the initial examination of the data. Style shifting effects also need to be explored for /s~z/, with comparisons between casual conversations, workplace interactions, and word lists. What balance of sources forms the input may also be relevant. There is also much scope for future in-class studies, particularly using mitigation materials and control groups with large numbers of learners and high-quality recordings.

There is clearly a need for such research in order to inform teaching in Vietnam and, given the obvious similarity in these issues, across the whole SE Asian area. The people learning English in these places are desperately keen to improve their pronunciation, and we English language professionals have a significant role to play in helping them achieve their goals. We hope that this study and future research will help highlight the problems and propose solutions to contribute to this improvement agenda and help learners in their goal of improving the lives of themselves and their families, which is surely what they are learning English for.

Author Contributions: Conceptualization, S.J.D.; Methodology, S.J.D.; Software, S.J.D.; Validation, S.J.D.; Formal analysis, S.J.D.; Investigation, S.J.D.; Resources, L.N.C.L.; Data curation, L.N.C.L.; Writing—original draft, S.J.D.; Writing—review & editing, S.J.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Ethics Committee of Plymouth Marjon University (protocol code EP104, final version approved on 10 July 2020).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The transcribed data, along with the annotations, are available on request from the lead author.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Table A1. The detailed raw results for all sixteen participants.

| Participant Number | Lexical | Morphological | Preceding Vowel | Preceding Consonant | Total Errors |
|--------------------|---------|---------------|-----------------|---------------------|--------------|
| 1 | 45.9% | 34.5% | 43.7% | 36.2% | 41.9% |
| 2 | 21.5% | 67.7% | 25.7% | 68.8% | 36.5% |
| 3 | 5.0% | 31.9% | 12.1% | 28.6% | 18.3% |
| 4 | 6.5% | 41.7% | 6.2% | 48.7% | 25.1% |
| 5 | 3.4% | 28.2% | 8.3% | 27.1% | 15.3% |
| 6 | 4.9% | 17.8% | 7.1% | 24.5% | 11.5% |

Table A1. Cont.

| Participant Number | Lexical | Morphological | Preceding Vowel | Preceding Consonant | Total Errors |
|--------------------|---------|---------------|-----------------|---------------------|--------------|
| 7 | 26.6% | 73.3% | 30.1% | 64.9% | 41.6% |
| 8 | 10.5% | 90.0% | 17.0% | 81.1% | 34.1% |
| 9 | 31.7% | 46.3% | 31.9% | 58.1% | 38.0% |
| 10 | 19.5% | 20.6% | 18.3% | 26.1% | 19.9% |
| 11 | 12.1% | 51.9% | 25.3% | 41.9% | 29.7% |
| 12 | 11.0% | 39.1% | 12.7% | 48.4% | 22.4% |
| 13 | 4.2% | 33.5% | 9.6% | 35.0% | 22.9% |
| 14 | 26.3% | 42.7% | 26.2% | 51.5% | 32.6% |
| 15 | 32.3% | 77.3% | 33.3% | 76.7% | 48.8% |
| 16 | 6.7% | 24.5% | 7.0% | 29.3% | 15.3% |
| Mean | 16.8% | 45.1% | 19.7% | 46.7% | 28.4% |
| Median | 11.6% | 40.4% | 17.6% | 45.2% | 27.4% |
| St. Deviation | 12.4 | 20.8 | 11.1 | 18.1 | 10.8 |

Notes

- Only P1 had more lexical absence than morphological. This was probably an entrenched error because it was almost all due to two high-frequency words with high error rates; 10 out of 27 instances of *Vietnamese* were incorrect, and all 16 instances of *voice* had no /s~z/. These also account for the preceding phoneme anomaly in their data.
- My thanks to Dr Jonathon Waddington at Plymouth Marjon University for his help in the statistical analysis of the data.
- Again, due to the word *Vietnamese* (lexical + preceding vowel), P1 alone breaks the pattern, where all others performed worst on this combination than elsewhere.
- Examples of target /ɪz/ were discarded because /l/ is typically velarised or vocalised before a consonant, and its status in this environment as an alveolar is indeterminate.

References

- Albright, Adam, and Bruce Hayes. 2004. Modeling Productivity with the Gradual Learning Algorithm: The Problem of Accidental Gaps in Phonological Theory. In *Proceedings of the thirtieth Annual Meeting of the Berkeley Linguistics Society*. Edited by Marc Ettliger, Nicholas Fleischer and Mischa Park-Doob. Berkeley: Berkeley Linguistics Society, pp. 37–48.
- Altenberg, Evelyn. 2005. The judgment, perception, and production of consonant clusters in a second language. *International Review of Applied Linguistics in Language Teaching* 43: 53–80. [CrossRef]
- Alves, Mark. 2006. Linguistic research on the origins of the Vietnamese language: An overview. *Journal of Vietnamese Studies* 1: 104–30. [CrossRef]
- Amos, Jenny, Jonathon R. Kasstan, and Wyn Johnson. 2020. Reconsidering the variable context: A phonological argument for (t) and (d) deletion. *English Today* 36: 6–13. [CrossRef]
- Archibald, John. 2022. Segmental and Prosodic Evidence for Property-by-Property Transfer in L3 English in Northern Africa. *Languages* 7: 28. [CrossRef]
- Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.
- Aronoff, Mark. 1994. *Morphology by Itself: Stems and Inflectional Classes*. Cambridge, MA: MIT Press.
- Baese-Berk, Melissa M., and Arthur G. Samuel. 2016. Listeners beware: Speech production may be bad for learning speech sounds. *Journal of Memory and Language* 89: 23–36. [CrossRef]
- Baranowski, Maciej, and Danielle Turton. 2020. TD-deletion in British English: New evidence for the long-lost morphological effect. *Language Variation and Change* 32: 1–23. [CrossRef]
- Bates, Elizabeth, and Brian MacWhinney. 1987. Competition, variation, and language learning. In *Mechanisms of Language Acquisition*. Edited by Brian MacWhinney. Mahwah: Lawrence Erlbaum Associates, Inc., pp. 157–93.
- Benua, Laura. 2005. *Transderivational Identity: Phonological Relations between Words*. New York: Routledge.
- Blench, Roger. 2015. The origins of nominal classification markers in MSEA languages. In *Mainland Southeast Asian Languages*. Edited by Nicholas J. Enfield and Bernard Comrie. Berlin: De Gruyter, pp. 558–87.
- Bloomfield, Leonard. 1933. *Language*. New York: Holt.
- BNC Consortium. 2007. *The British National Corpus, XML Edition*. Oxford Text Archive. Available online: <http://hdl.handle.net/20.500.12024/2554> (accessed on 17 July 2021).
- Bod, Rens, Jennifer Hay, and Stefanie Jannedy, eds. 2003. *Probabilistic Linguistics*. Cambridge, MA: MIT Press.
- Bohn, Ocke-Schwen, and Rikke L. Bundgaard-Nielsen. 2009. Second language speech learning with diverse inputs. In *Input Matters in SLA*. Edited by Thorsten Piske and Martha Young-Scholten. Bristol: Multilingual Matters, pp. 207–18.

- Bui, Thanh-Son. 2016. Pronunciations of consonants /d/ and /θ/ by adult vietnamese EFL learners. *Indonesian Journal of Applied Linguistics* 6: 125–34. [\[CrossRef\]](#)
- Bybee, Joan L., and Paul J. Hopper. 2001. Introduction to Frequency and the emergence of linguistic structure. In *Frequency and the Emergence of Linguistic Structure. Typological Studies in Language* 45. Edited by Joan L. Bybee and Paul J. Hopper. Amsterdam: John Benjamins, pp. 1–24.
- Chomsky, Noam, and Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper & Row.
- Colantoni, Laura, Jeffrey Steele, and Paola Escudero. 2015. *Second Language Speech: Theory and Practice*. Cambridge: Cambridge University Press. [\[CrossRef\]](#)
- Collins, Beverley, and Inger M. Mees. 2013. *Practical Phonetics and Phonology: A Resource Book for Students*. London: Routledge.
- Corder, Stephen Pit. 1967. The significance of learner's errors. *International Review of Applied Linguistics in Language Teaching* 5: 161–70. [\[CrossRef\]](#)
- Cruttenden, Alan. 2014. *Gimson's Pronunciation of English*. London: Routledge.
- Cummins, Jim. 1979. Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research* 49: 222–51. [\[CrossRef\]](#)
- Derwing, Tracey M., and Murray J. Munro. 2015. *Pronunciation Fundamentals: Evidence-Based Perspectives for L2 Teaching and Research*. Amsterdam: John Benjamins Publishing Company.
- De Villiers, Jill G., and Valerie E. Johnson. 2007. The information in third-person/s: Acquisition across dialects of American English. *Journal of Child Language* 34: 133–58. [\[CrossRef\]](#)
- Disney, Steve. 2016. Another Visit to BE Supposed to from a Diachronic Constructionist Perspective. *English Studies* 97: 892–916. [\[CrossRef\]](#)
- Doughty, Catherine, and Jessica Williams, eds. 1998. *Focus on Form in Classroom Second Language Acquisition*. New York: Cambridge University Press.
- Eckman, Fred R. 1977. Markedness and the contrastive analysis hypothesis. *Language Learning* 27: 315–30. [\[CrossRef\]](#)
- Eckman, Fred R. 1981. On the naturalness of interlanguage phonological rules. *Language Learning* 31: 195–216. [\[CrossRef\]](#)
- Eckman, Fred R. 1987. The reduction of word-final consonant clusters in interlanguage. In *Sound Patterns in Second Language Acquisition*. Edited by Allan James and Jonathan Leather. Dordrecht: Foris, pp. 143–62.
- Edwards, Janice. G. H. 2011. Deletion of /t, d/ and the Acquisition of Linguistic Variation by Second Language Learners of English. *Language Learning* 61: 1256–301. [\[CrossRef\]](#)
- Ellis, Nick C. 2002. Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24: 143–88. [\[CrossRef\]](#)
- Ellis, Nick C. 2008. Usage-based and Form-focused Language Acquisition: The Associative Learning of Constructions, Learned-attention, and the Limited L2 Endstate. In *The Handbook of Cognitive Linguistics and Second Language Acquisition*. Edited by Peter Robinson and Nick C. Ellis. London: Routledge, pp. 372–405.
- Ellis, Nick C. 2013. Frequency-based accounts of second language acquisition. In *The Routledge Handbook of Second Language Acquisition*. London: Routledge, pp. 193–210.
- Ellis, Nick C., and Laura Collins. 2009. Input and Second Language Acquisition: The Roles of Frequency, Form, and Function. Introduction to the Special Issue. *The Modern Language Journal* 93: 329–35. [\[CrossRef\]](#)
- Ellis, Nick C., and Richard Schmidt. 1997. Morphology and longer distance dependencies: Laboratory research illuminating the A in SLA. *Studies in Second Language Acquisition* 19: 145–71. [\[CrossRef\]](#)
- Enfield, Nicholas J., and Bernard Comrie. 2015. Mainland Southeast Asian languages: State of the art and new directions. In *Languages of Mainland Southeast Asia: The State of the Art*. Edited by Nicholas J. Enfield and Bernard Comrie. Berlin: De Gruyter Mouton, pp. 1–20. [\[CrossRef\]](#)
- Flege, James E. 2009. Give input a chance! In *Input Matters in SLA*. Edited by Thorsten Piske and Martha Young-Scholten. Bristol: Multilingual Matters, pp. 175–90.
- Flege, James E., and Wieke Eefting. 1987. Production and perception of English stops by native Spanish speakers. *Journal of Phonetics* 15: 67–83. [\[CrossRef\]](#)
- Flege, James E., Murray J. Munro, and Linda Skelton. 1992. Production of the word-final English /t-/d/ contrast by native speakers of English, Mandarin, and Spanish. *The Journal of the Acoustical Society of America* 92: 128–43. [\[CrossRef\]](#)
- Fullana, Noemí, and Joan C. Mora. 2009. Production and perception of voicing contrasts in English word-final obstruents: Assessing the effects of experience and starting age. *Recent Research in Second Language Phonetics/Phonology: Perception and Production*, 97–117.
- Gass, Susan M., and Larry Selinker. 2010. Second Language Acquisition. *Annual Review of Applied Linguistics* 25: 46–73. [\[CrossRef\]](#)
- Gass, Susan M., Jennifer Behney, and Luke Plonsky. 2020. *Second Language Acquisition: An Introductory Course*, 5th ed. London: Routledge.
- Goldwater, Sharon, and Mark Johnson. 2006. Learning English Morphology. *Advances in Neural Information Processing Systems* 18: 681–88.
- Greenberg, Joseph. H. 1965. Some generalizations concerning initial and final consonant sequences. *Linguistics* 3: 5–34. [\[CrossRef\]](#)
- Gut, Ulricke. 2007. First language influence and final consonant clusters in the new Englishes of Singapore and Nigeria. *World Englishes* 26: 346–59. [\[CrossRef\]](#)
- Haspelmath, Martin, and Andrea D. Sims. 2010. *Understanding Morphology*, 2nd ed. London: Routledge.

- Hayes, Bruce, Kathryn Zuraw, Péter Siptár, and Zsófia Londe. 2003. Stochastic OT: Learning and Typology. *Phonology* 20: 273–377.
- Holliday, Adrian. 2006. Native-speakerism. *ELT Journal* 60: 385–87. [CrossRef]
- Honey, Peter. J. 1987. Vietnamese Speakers. In *Learner English: A Teacher's Guide to Interference and Other Problems*, 1st ed. Edited by Michael Swann and Bernard Smith. Cambridge: Cambridge University Press, pp. 238–51.
- Hwa-Froelich, Deborah, Barbara Hodson, and Helen Edwards. 2002. Characteristics of Vietnamese Phonology. *American Journal of Speech-Language Pathology* 11: 264–73. [CrossRef]
- James, Carl. 1998. *Errors in Language Learning and Use: Exploring Error Analysis*. London: Routledge.
- Kiparsky, Paul. 1982. From Cyclic Phonology to Lexical Phonology. In *The Structure of Phonological Representations*. Edited by Harry van der Hulst and Norval Smith. Dordrecht: Foris Publications, pp. 131–75.
- Kirby, James P. 2011. Vietnamese (Hanoi Vietnamese). *Journal of the International Phonetic Association* 41: 381–92. [CrossRef]
- Kueser, Juliana B., Laurence B. Leonard, and Patricia Deevy. 2018. Third person singular-s in typical development and specific language impairment: Input and neighbourhood density. *Clinical Linguistics and Phonetics* 32: 232–48.
- Labov, William. 1972. *Sociolinguistic Patterns*. Oxford: Blackwell.
- Lado, Robert. 1957. *Linguistics across Cultures*. Ann Arbor: Ann Arbor University of Michigan Press.
- Larsen-Freeman, Diane. 1975. The acquisition of grammatical morphemes by adult ESL students. *TESOL Quarterly* 9: 409–30. [CrossRef]
- Larsen-Freeman, Diane. 1991. Second language acquisition research: Staking out the territory. *TESOL Quarterly* 25: 315–50. [CrossRef]
- Larsen-Freeman, Diane. 2002. Making sense of frequency. *Studies in Second Language Acquisition* 24: 275–85. [CrossRef]
- Leung, Alex Ho-Cheong. 2014. Input Multiplicity and the Robustness of Phonological Categories in Child L2 Phonology Acquisition. In *Proceedings of the International Symposium on the Acquisition of Second Language Speech Concordia Working Papers in Applied Linguistics* 5. Montreal: Concordia University.
- Leung, Alex Ho-Cheong, Martha Young-Scholten, Wael Almurashi, Saleh Ghadanfari, Chloe Nash, and Olivia Outhwaite. 2023. (Mis) perception of consonant clusters and short vowels in English as a foreign language. *International Review of Applied Linguistics in Language Teaching* 61: 731–64. [CrossRef]
- Lieven, Elena, and Michael Tomasello. 2008. Children's first language acquisition from a usage-based perspective. In *Handbook of Cognitive Linguistics and Second Language Acquisition*. Edited by Peter Robinson and Nick C. Ellis. New York: Routledge, pp. 168–96.
- Loi, Dao Phuong. 2018. An Analysis of Vietnamese EFL Students' Pronunciation of English Affricates and Nasals. *International Journal of English Linguistics* 8: 298–306. [CrossRef]
- Major, Roy C. 2001. *Foreign Accent: The Ontogeny and Phylogeny of Second Language Phonology*. London: Routledge.
- Major, Roy C. 2004. Gender and stylistic variation in second language phonology. *Language Variation and Change* 16: 169–88. [CrossRef]
- Major, Roy C. 2008. Transfer in second language phonology: A review. In *Phonology and Second Language Acquisition. Studies in Bilingualism* 36. Edited by Jette G. Hansen Edwards and Mary L. Zampini. Amsterdam and Philadelphia: John Benjamins Publishing Company, pp. 63–94. [CrossRef]
- Major, Roy C., and Edward Kim. 1996. The similarity differential rate hypothesis. *Language Learning* 46: 465–96. [CrossRef]
- Maniwa, Kazumi, Allard Jongman, and Tessa Wade. 2009. Acoustic characteristics of clearly spoken English fricatives. *The Journal of the Acoustical Society of America* 125: 3962–73. [CrossRef]
- McCarthy, John J. 2002. *A Thematic Guide to Optimality Theory*. Cambridge: Cambridge University Press.
- McLeod, Sharynne, Jean Van Doorn, and Valerie A. Reed. 2001. Normal Acquisition of Consonant Clusters. *American Journal of Speech-Language Pathology* 10: 265–74. [CrossRef]
- Moran, Mary Jane. 1993. Final consonant deletion in African American children speaking Black English: A closer look. *Language, Speech, and Hearing Services in Schools* 24: 161–66. [CrossRef]
- Nguyễn-Dăng-Liêm. 1970. *Vietnamese Pronunciation*. Hawaii: University of Hawaii Press.
- Nu, Duong Thi. 2009. Mistake or Vietnamese English. *VNU Journal of Science, Foreign Languages* 25: 41–50. Available online: <https://js.vnu.edu.vn/FS/article/view/2233> (accessed on 21 March 2022).
- Odlin, Terence. 1989. *Language Transfer*. Cambridge: Cambridge University Press.
- Osborne, Alan G. 1996. Final cluster reduction in English L2 speech: A case study of a Vietnamese speaker. *Applied Linguistics* 17: 164–181. [CrossRef]
- Piske, Thorsten, and Marie Young-Scholten, eds. 2008. *Input Matters in SLA*. Bristol: Multilingual Matters.
- Piske, Thorsten, Ian R. A. MacKay, and James E. Flege. 2001. Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics* 29: 191–215. [CrossRef]
- Prince, Alan, and Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden: Blackwell.
- Prince, Alan, and Paul Smolensky. 1997. Optimality: From neural networks to universal grammar. *Science* 275: 1604–10. [CrossRef] [PubMed]
- Raymond, William D., Esther L. Brown, and Alice F. Healy. 2016. Cumulative context effects and variant lexical representations: Word use and English final t/d deletion. *Language Variation and Change* 28: 175–202. [CrossRef]
- Sato, Charlene. 1990. The Syntax of conversation in interlanguage development. In *Language Development* 11. Edited by Sascha W. Felix, Jürgen Meisel and Henning Wode. Tübingen: Guntar Narr Verlag, pp. 75–92.

- Schreier, Daniel. 2005a. Initial Cluster Reduction in English. In *Consonant Change in English Worldwide: Synchrony Meets Diachrony*. Edited by Daniel Schreier. Basingstoke: Palgrave Macmillan, pp. 56–125.
- Schreier, Daniel. 2005b. Final Cluster Reduction in English. In *Consonant Change in English Worldwide: Synchrony Meets Diachrony*. Edited by Daniel Schreier. Basingstoke: Palgrave Macmillan, pp. 126–97.
- Selinker, Larry. 1972. Interlanguage. *IRAL-International Review of Applied Linguistics in Language Teaching* 10: 209–31. [CrossRef]
- Slobin, Dan I. 1992. *Psycholinguistics*, 2nd ed. Glenview: Scott, Foresman and Company.
- Swan, Michael, and Bernard Smith. 1987. *Learner English: A Teacher's Guide to Interference and Other Problems*, 1st ed. Cambridge: Cambridge University Press.
- Swan, Michael, and Bernard Smith. 2001. *Learner English: A Teacher's Guide to Interference and Other Problems*, 2nd ed. Cambridge: Cambridge University Press.
- Sweller, John, Paul Ayres, and Slava Kalyuga. 2011. *Cognitive Load Theory*. New York: Springer.
- Tang, Giang. 2007. Cross-Linguistic Analysis of Vietnamese and English with Implications for Vietnamese Language Acquisition and Maintenance in the United States. *Journal of Southeast Asian American Education and Advancement* 2: 1–33. [CrossRef]
- Tarone, Elaine, and Zhao Hong Han. 2014. *Interlanguage*. Amsterdam: John Benjamins Publishing Company.
- Taylor, Rodger. 2008. Prototypes in Cognitive Linguistics. In *The Handbook of Cognitive Linguistics and Second Language Acquisition*. Edited by Peter Robinson and Nick C. Ellis. London: Routledge, pp. 39–65.
- Temperley, Mark S. 1983. The articulatory target for final-s clusters. *TESOL Quarterly* 17: 421–36. [CrossRef]
- Torkildsen, Janne von Koss, and Jessica S. Horst. 2019. *International Handbook of Language Acquisition*. London: Routledge.
- Traugott, Elizabeth C. 1989. On the Rise of Epistemic Meanings in English: An Example of Subjectification in Semantic Change. *Language* 65: 31–55. [CrossRef]
- Trofimovich, Pavel, Laura Collins, Walcir Cardoso, Joanna White, and Marlise Horst. 2012. A Frequency-Based Approach to L2 Phonological Learning: Teacher Input and Student Output in an Intensive ESL Context. *TESOL Quarterly* 46: 176–87. Available online: <http://www.jstor.org/stable/41576035> (accessed on 31 March 2024). [CrossRef]
- Trudgill, Peter. 1974. *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.
- Trudgill, Peter. 1992. Norwich Revisited: Recent Linguistic Changes in an English Urban Dialect. In *Thirty Years of Linguistic Evolution: Studies in Honour of René Dirven on the Occasion of His 60th Birthday*. Edited by Martin Pütz. Amsterdam: John Benjamins Publishing Company, pp. 361–78.
- Trudgill, Peter. 1997. Third-Person Singular Zero: African-American English, East Anglian Dialects and Spanish Persecution in the Low Countries. *Folia Linguistica Historica* 31: 139–48. [CrossRef]
- VanPatten, Bill. 1996. *Input Processing and Grammar Instruction in Second Language Acquisition*. New York: Ablex.
- Wells, John C. 1990. Syllabification and Allophony. In *Studies in the Pronunciation of English: A Commemorative Volume in Honour of A.C. Gimson*. Edited by Susan Ramsaran. London: Routledge, pp. 76–86.
- Westergaard, Marit. 2021. Microvariation in multilingual situations: The importance of property-by-property acquisition. *Second Language Research* 37: 379–407. [CrossRef]
- Westergaard, Marit, Natalia Mitrofanova, Roksolana Mykhaylyk, and Yulia Rodina. 2017. Crosslinguistic influence in the acquisition of a third language: The Linguistic Proximity Model. *International Journal of Bilingualism* 21: 666–82. [CrossRef]
- Yang, Charles, and Rens Bod. 2008. Probabilistic linguistics. *Journal of Linguistics* 44: 205–8. [CrossRef]
- Zsiga, Elizabeth C. 2013. *The Sounds of Language: An Introduction to Phonetics and Phonology*. Malden: Wiley-Blackwell.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.