

JOURNAL ARTICLE

The role of vowel length and glottalization in German learners' perception of the English coda stop voicing contrast

Eva Reinisch^{1,2} and Joshua Penney³

¹ Institute of Phonetics and Speech Processing, Ludwig Maximilians University Munich, DE

² Acoustics Research Institute, Austrian Academy of Sciences, Vienna, AT

³ Centre for Language Sciences, Department of Linguistics, Macquarie University, Sydney, AU

Corresponding author: Eva Reinisch (Eva.Reinisch@oeaw.ac.at)

In German, the voicing contrast in word-final stops is neutralized towards the voiceless sound. We tested how German learners of English use in perception two phonetic cues to this contrast in English: the duration of the vowel preceding the stop and the partial glottalization of this vowel. While a longer vowel cues the voiced sound of the contrast, glottalization enhances the voiceless sound, which should be 'easy' for learners as word-finally it is the default in German. We asked whether cueing the 'easy' sound would nevertheless affect learners' word identification. Learners categorized two English minimal pairs along vowel duration continua with either a fully modal vowel or the last 25% of the vowel glottalized. Learners gave more voiced-stop responses as vowel duration increased. They also used glottalization by giving fewer voiced-stop responses for the glottalized continua. A second experiment demonstrated that the glottalization was not merely perceived as a change in the vowel+closure duration ratio. When the glottalized portion of the vowels was set to silence learners gave even fewer voiced-stop responses than in the glottalized condition. Results suggest that learners can use a phonetic cue to a second language sound contrast even if it enhances the familiar 'easy' sound.

Keywords: second language learning; speech perception; acoustic cues; glottalization; word-final stop voicing

1. Introduction

When learning a second language (L2) later in life, one of the most difficult tasks for learners is to acquire the sound system of the L2. This is because learning an L2 often requires the establishment of new L2 phonological sound contrasts that may involve one or more sounds that are not part of the native language (L1) sound inventory. For instance, German learners of English have to establish a new vowel category /æ/ that in English (but not German) contrasts with /ɛ/ in order to distinguish between words such as *bed* and *bad*. Learning such sound contrasts in perception and production has been shown to be very difficult for learners (e.g., Bohn & Flege, 1992; Eger & Reinisch, 2017, 2019; Flege, Bohn, & Jang, 1997; Iverson & Evans, 2007; Llompарт & Reinisch, 2017, 2019, in press). In other cases, learners do not have to establish a new phonological contrast in the L2 but they have to learn an existing L1 contrast in a new word position in the L2—a position where the relevant sounds do not occur in the L1 or the relevant distinction is neutralized. A case in point here is the voicing distinction in English word-final obstruents (e.g., /t-/ /d/ in *beat* versus *bead*). This distinction has been shown to be of substantial difficulty for many learners of English from a wide variety of native languages such as Mandarin, Cantonese,

Japanese, Korean, Dutch, or German (e.g., Baker, 2010; Bent, Badlow, & Smith, 2008; Crowther & Mann, 1992; Broersma, 2005; Flege & Hillenbrand, 1987; Flege & Wang, 1989; Flege & Port, 1981; Hayes-Harb, Smith, Bent, & Bradlow, 2008).

Native speakers of languages that neutralize the word-final voicing contrast (e.g., native speakers of German, Dutch, Polish, or Russian) have been shown to transfer this neutralization in favour of the voiceless sounds when learning to produce English words and consequently have difficulty distinguishing the sounds of this contrast in perception (Broersma, 2005; Eger & Reinisch, 2017, 2019; Smith, Hayes-Harb, Bruss, & Harker, 2009). Since in word-final position the voiced obstruent is typically the 'new' and hence 'difficult' sound of the contrast for learners of English, the question that most studies address is what combination of acoustic-phonetic cues to the voiced sound can learners use to differentiate it from the familiar, 'easy' voiceless sound (e.g., Eger & Reinisch, 2017). The present study took the opposite perspective and addressed whether acoustic-phonetic enhancement of the difference between the sounds of the contrast would affect perceptual differentiation, even if this enhancement marked the familiar voiceless sound. Specifically, we asked whether German learners of English could use glottalization of the preceding vowel as a cue to voicelessness of a word-final stop in order to differentiate it from a word-final voiced stop.

The English word-final voicing contrast in stops is well known for its many acoustic cues (Lisker, 1986; Raphael, 1981) that are not only found in the critical segments (e.g., /t/ versus /d/) but also in the preceding segments, like, most prominently, the duration of a preceding vowel. A vowel preceding /d/ is typically longer than a vowel preceding /t/, with vowel duration in the voiced coda context being up to double the duration of the vowel in the voiceless coda context (Chen, 1970; Fowler, 1992; Klatt, 1976; Lisker, 1974; Raphael, 1972; Raphael, Dorman, Freeman, & Tobin, 1975). Additional cues to voicing are a shorter closure duration and longer sustenance of voicing during the closure (though closure voicing is often weak or absent; Docherty, 1992; Keyser & Stevens, 2006; Klatt, 1976; Westbury & Keating, 1986), and shorter and less intense release bursts. In addition, a lower fundamental frequency (F0) and/or a drop in the first formant (F1) frequency in the preceding vowel indicate a voiced stop. In addition to these cues to stop voicing, in some varieties of English the presence of glottalization may specifically signal a *voiceless* stop. Glottalization, as we use the term in this paper, refers to the addition of glottal constriction to an oral stop, resulting in irregular, laryngealized voicing of the preceding vowel, and, in some cases, a full glottal stop (Garellek, 2015; Huffman, 2005; Penney, Cox, Miles, & Palethorpe, 2018b; Seyfarth & Garellek, 2015).

Glottalization preceding voiceless stops has been observed in multiple varieties of British English (Docherty & Foulkes, 1999; Gordeeva & Scobbie, 2013; Mathisen, 1999; Roach, 1973; Tollfree, 1999), American English (Huffman, 2005; Pierrehumbert, 1995; Seyfarth & Garellek, 2015), and also in Australian English (Penney et al., 2018b). Critically, for Australian English a recent change has been observed suggesting that glottalization is becoming an increasingly prominent cue to distinguish the final voicing contrast while at the same time the importance of the duration of the preceding vowel has started to decrease (Penney et al., 2018b). In an apparent time study, Penney et al. (2018b) found that while both younger and older native speakers of Australian English produced glottalization primarily preceding voiceless (rather than voiced) coda stops, younger speakers produced more glottalization than older speakers. Conversely, younger speakers used vowel duration to cue coda voicing to a lesser extent than older speakers, pointing to a possible change in cue weighting in the language system. In addition, both age groups produced the least amount of glottalization in high vowel contexts, and the vowel duration cue was used most clearly in combination with high vowels. This suggests that glottalization may serve as an important cue to stop voicelessness, particularly if preceding vowel duration is not an informative cue.

Native English listeners also appear to be sensitive to glottalization in perception. Garellek (2011) found that American English listeners identified /t/ faster and more accurately when glottalization was present than when it was absent. In an eye-tracking task, Chong and Garellek (2018) found that glottalization hindered perception of voiced but not voiceless codas, demonstrating that American English listeners associate glottalization with voiceless stops. Penney, Cox, and Szakay (2018a, submitted) found that Australian English listeners are also perceptually aware of glottalization, and interpret glottalization as a cue to stop voicelessness. In a task in which stimuli co-varying in vowel duration and the presence of glottalization were presented, listeners perceived more voiceless codas when glottalization was present than when it was absent, even when this occurred in combination with extended preceding vowel duration, which would otherwise signal a voiced coda. This effect was found both for older and younger Australian English listeners, despite older speakers using glottalization less than younger speakers in production (Penney et al., 2018b). These results suggest that L1 Australian English speakers/listeners use glottalization to cue voiceless stops, in both production and perception.

Learners of English are faced with the sum of the different acoustic cues when learning to master the word-final voicing contrast. In the literature on second language production and perception, the focus has naturally been on the traditionally most important cue of vowel duration (e.g., Baker, 2010; Bent et al., 2008; Crowther & Mann, 1992; Broersma, 2005; Flege & Hillenbrand, 1986; Flege & Port, 1981; Hayes-Harb et al., 2008), with some studies considering closure duration (e.g., Baker, 2010), F1 offset (Crowther & Mann, 1992) or the presence of a burst (in combination with the use of vowel duration; Broersma, 2005; Flege, 1989; Flege & Wang, 1989). Findings on learners' use of these cues suggest that learners with native languages that make use of at least one of these cues in their L1—in other positions of the word or even for other sound contrasts—are typically better in perception and more target-like in production of the English final voicing contrast than learners who are entirely unfamiliar with a given cue. For instance, Dutch learners can draw on their knowledge about longer vowel duration preceding voiced than voiceless stops in word medial position to interpret the English final voicing contrast (e.g., Broersma, 2005). But also experience with phonemic vowel duration in the L1 helps using the vowel duration cue for the detection of English final stop voicing (Crowther & Mann, 1992; Flege & Hillenbrand, 1987; Hayes-Harb et al., 2008; but see Flege & Port, 1981, who do not find such an effect). Similarly, learners with L1s that allow for stops in word-final position even without a voicing contrast (e.g., Cantonese versus Mandarin) appear to have an advantage in using cues such as vowel duration when the stop burst is missing as supposedly they are used to paying attention to the presence of word-final obstruents (Flege, 1989; Flege & Wang, 1989). However, despite these advantages due to familiarity with a given cue, even advanced L2 learners hardly ever match English native speakers in their use of this cue (e.g., Broersma, 2005). In the present study we tested German learners of English on their perception of the English word-final voicing contrast. We thereby focused not only on the role of familiarity with a given cue to the contrast but assessed whether learners would use a cue which enhances the voiceless category that is the default in their L1 rather than the 'difficult' voiced English sound that does not occur in word-final position in German. Testing the use of an additional cue to the 'easy' sound contributes to our understanding of L2 learning since previous studies have shown that good versus poor German learners of English differ in how well they are able to use cues to the 'difficult' rather than 'easy' sounds (Eger & Reinisch, 2017, 2019; though note that these studies looked at natural productions and hence a combination of cues).

German is known for final obstruent devoicing such that *Rat* ('advice') and *Rad* ('wheel') are homophones [ʁa:t] while a phonological difference surfaces in other morphological

forms such as their plurals *Rat-Räte* [ʁɛ:tə] vs. *Rad-Räder* [ʁɛ:də]. However, a number of studies on German (and other languages with final obstruent devoicing) have suggested that in production this neutralization is sometimes incomplete (see Nicenboim, Roettger, & Vasishth, 2018, for a critical discussion). Remaining differences have been shown in the duration of the preceding vowel, the closure duration, the duration of closure voicing, and the burst and aspiration durations (O'Dell & Port, 1983; Port, Mitleb, & O'Dell, 1981; Port & O'Dell, 1985; Piroth & Janker, 2004; Roettger, Winter, Grawunder, Kirby, & Grice, 2014). In perception, it has been shown that German listeners can use these cues when asked to categorize word-final obstruents, even though they perform just above chance level (Kleber, John, & Harrington, 2010; Port & Crawford, 1989; Port & O'Dell, 1985; Roettger et al., 2014). Therefore, given this low accuracy especially in identifying the voiced obstruents, it has been established that acoustic differences due to incomplete neutralization are unlikely to be functionally relevant in German (e.g., Port & O'Dell, 1985; Kleber et al., 2010; Roettger et al., 2014; see also Mitterer & Reinisch, 2017).

The cues that are found to differentiate the voiced versus voiceless word-final stops in German are overwhelmingly the same as those that are found in the English contrast (though with a larger difference between the two categories in English as opposed to the small difference due to incomplete neutralization in German). Smith and colleagues (Smith & Peterson, 2012; Smith et al., 2009) asked German learners of English and English learners of German to produce English minimal word pairs and orthographically similar German words (e.g., English *toad-tote* and German *Tod-tot* 'death'-'dead'). Results suggest that despite considerable variability between participants, in general, the voiceless sounds were produced similarly in English and German, and German learners of English produced the same cues for incomplete voicing neutralization in German and the contrast between word-final voiced versus voiceless stops in English. These cues were mainly the duration of the preceding vowel and closure duration with longer vowel and shorter closure duration cueing the voiced sound. German learners tended to produce a larger distinction between the voiced and voiceless categories in English than in German, but differentiated between the English categories to a lesser extent than native speakers of English.

As for the perception of cues to the English word-final stop voicing contrast, previous studies (Smith et al., 2009; see also Eger & Reinisch, 2017, 2019, for the same language pair) compared native English listeners' and German learners' perception of native versus learners' productions of English minimal pairs. They found that German learners were able to identify the intended word and hence use available cues to the word-final stop voicing distinction, although—maybe unsurprisingly—native English listeners outperformed the learners in the identification of native and non-native productions. However, these studies focused on the combination of a set of cues to the voicing distinction: vowel duration, stop closure duration, voiced portion of the closure, and the burst and aspiration duration of the stop. The relative contribution of each acoustic-phonetic cue to the perception of the contrast hence remains unknown.

One thing that Eger and Reinisch (2017, 2019) did test for German learners of English was whether learners differed in perception (goodness ratings and word identification) depending on whether the intended word contained the familiar 'easy' voiceless stop or the 'difficult' voiced stop (in addition to sounds from other L2 contrasts). Results showed that learners were better at identifying words containing voiceless rather than voiced stops—likely because any neutralization of the contrast would go by default towards the voiceless category (Smith et al., 2009). Cues to the voiced category were detected better the larger their difference to the 'default' voiceless values were (e.g., the longer the duration of the preceding vowel, the shorter the closure duration, etc.), and the more proficient learners

were themselves in producing large differences between the sounds of the contrasts. The authors suggest that this is because only proficient learners have mental representations of the voiced sounds that are similar to native speakers' representations and hence have more precise expectations about the cues that indicate a final voiced stop. The representation of the voiceless stop can simply be transferred from the L1 and is hence similar for all learners.

The present study addressed how German learners of English use two specific cues to the English word-final voicing contrast in stops: the duration of the preceding vowel and the glottalization of this vowel. Since previous studies suggest that learners can transfer their use of phonetic cues for a given contrast across positions in the word (e.g., Broersma, 2005) and across different phoneme contrasts (e.g., Crowther & Mann, 1992), we expected German learners to be quite able to use a longer vowel duration as a cue to the voiced stop. This is because, firstly, as discussed above, a longer vowel duration is involved in incomplete neutralization of the German contrast in word-final position (i.e., the cue exists in the L1 to some extent). Secondly, vowel duration is an important cue to word-medial stop voicing in German, especially if the cue of aspiration duration is diminished due to a lateral or nasal release (e.g., in *leiden* [laɪdŋ] versus *leiten* [laɪtŋ] 'to suffer'-'to guide'; Kohler, 1977). Thirdly, German uses duration as a cue to distinguish between tense versus lax vowels. Therefore, Germans should be familiar with vowel duration as a cue from their L1 from various contexts. Given that vowel duration is typically reported as one of the most important cues to final stop voicing in English (Klatt, 1976; Raphael, 1972; Raphael et al., 1975), we expect our learners to use a longer vowel duration to categorize English word-final stops as voiced.

Whether or not learners would use the second cue under investigation, glottalization, is harder to predict. Glottalization so far has not been the focus of studies on the acquisition of this contrast by L2 learners hence data for comparison are scarce. In German glottalization is used as a variant of glottal stop that canonically precedes foot-initial vowels (i.e., vowels that native speakers of German typically consider the onset of 'vowel-initial' words; Wiese, 1996, but see Mitterer & Reinisch, 2015, who show that glottalization is essential in the recognition of these words in connected speech). In addition, Germans are not entirely unfamiliar with the concept of glottalization as a possible realization of stops in German. Words such as *können* [kœnŋ] ('being able to') and *könnten* [kœntŋ] ('(they) may be able to') may in production solely be distinguished by the presence versus absence of glottalization in the nasal instead of a [t] (Kohler, 2000, 2001; see also John & Harrington, 2007) and listeners are sensitive to this cue in perception (Kohler, 2001). However, in neither of these cases of glottalization is a voicing distinction at stake. Therefore, German learners of English may rely on glottalization less in perceiving the English word-final stop voicing contrast than on vowel duration where a more direct link to the L1 is found. Critically, glottalization has been shown to cue the *voiceless* sound for native English speakers (Chong & Garellek, 2018; Penney et al., submitted). That is, for German learners of English it would enhance the sound of the contrast that is already similar to their L1 (i.e., which is the default in word-final position in German) and should hence be 'easy' to learn. Note also that in the studies by Eger and Reinisch (2017, 2019) good versus poor German learners of English mainly differed in their use of cues to the difficult sounds. Given this, it might be the case that learners ignore the glottalization cue or may even be confused by the presence of glottalization and respond to glottalization in unknown ways. In fact, given that the glottalized portion of the vowel is voiced (albeit irregularly) it could be possible that, in contrast to native English listeners, German learners take glottalization as closure voicing and hence as cueing a voiced coda. To our knowledge this

issue has not been previously explored. Experiment 1 therefore focuses on learners' use of glottalization to distinguish the English voicing contrast, and Experiment 2 presents a control to further assess how learners interpret or weight the cue of glottalization.

2. Experiment 1

Experiment 1 set out to test German learners' perception of the English word-final voicing contrast in stops using the minimal pairs *bart-bard* and *beat-bead* (i.e., pairs with a low versus high vowel preceding the /t-/d/ contrast). Specifically, it compared the use of duration of the preceding vowel and the presence of glottalization on the vowel. Duration was manipulated along a 9-step vowel continuum, and glottalization was present versus absent on the last 25% of the vowel. All other cues to the stop voicing distinction were neutralized. We expected that German learners of English would perceive a longer vowel duration as cueing the voiced stop, and we speculated that they may interpret the presence of glottalization on the last portion of the vowel as a cue to the voiceless sound. It remains to be shown whether and how the two cues interact in L2 perception. Predictions about possible interactions are difficult, since in line with Penney et al. (submitted) we use a relative proportion of glottalization on the vowel (mirroring L1 production; Penney et al., 2018b). Therefore, a longer vowel duration in the glottalized condition will have a comparatively longer glottalized portion of the vowel than shorter vowels.

2.1. Methods

2.1.1. Participants

Participants were recruited to be native speakers of German between 18 and 40 years of age, without known speech, language, or hearing issues. Nineteen participants¹ took part on-site at the Institute of Phonetics at the University of Munich. They took part voluntarily after completing another unrelated experiment. They were all students at the University of Munich and on average 22.3 years old ($SD = 3.1$, range 18–29). Another set of participants was recruited according to the same criteria to participate in the experiment via the internet. This was done to increase the number of participants in the experiment as well as in order to compare results collected on-site versus via the web to assess whether it is feasible to conduct web experiments on phonetic cues.

Participants for the web experiment were recruited via the same electronic university newsletter via which on-site participants were recruited. In addition, the link was spread via colleagues and friends who posted the link on other (university) mailing lists throughout Germany. Across Experiments 1 and 2, a total of 123 participants started the web experiment of which 61 completed all trials (see below; via the web both experiments were run simultaneously and participants were assigned randomly). Twenty-nine web participants completed Experiment 1, of which data from 28 participants were retained for analyses (18 female, 10 male, none other). Data from one participant who reported a native language other than German (i.e., Bulgarian) were excluded for this reason. The average age of the web participants in Experiment 1 was 27.2 years ($SD = 5.2$, range 18–38).

In addition to completing the experiment, all participants were required to answer a number of questions on their frequency of use and exposure to English. While the on-site participants filled in an extensive language-background questionnaire, the number and type of questions was restricted for the web participants to keep the overall duration of the experiment as short as possible. Therefore, only comparable questions between the two sets of participants will be reported. All participants were asked whether they aimed at speaking

¹ Note that information about the gender of the on-site participants was accidentally lost when anonymizing the data.

a specific variety of English. Eight of the on-site and four of the web participants indicated that they aimed at sounding British or Irish, six of the on-site participants aimed at sounding American, two of the web participants aimed at speaking with an Australian or New Zealand accent, and the rest indicated not to model their English on any specific variety.

All other questions involved a rating that had to be answered on a scale from 1–7. Answers were re-coded where necessary to match reported values between the on-site and web-based questionnaires. The questions were “How often do you hear English?” (1 = not at all, 7 = very often), “How strong do you consider your own German accent when speaking English?” (1 = not noticeable, 7 = very strong), “How annoyed are you when you hear someone else speaking English with a German accent?” (1 = not at all, 7 = very much), “How often do you write English?” (1 = not at all, 7 = very often), “How often do you read in English?” (1 = not at all, 7 = very often). Given the focus of the present study on glottalization and the use of stimuli from an Australian speaker (see below), we also asked specifically about participants' exposure to and familiarity with Australian English. A summary of the mean ratings (and standard deviations in brackets) is given in **Table 1**.

A qualitative comparison of the two groups of participants suggests that on-site participants were somewhat younger than the web participants, had a stronger desire to model their English to a specific variety, but had overall less exposure to English than the web participants. Specifically, they reported being less familiar with Australian English. Note, however, that this difference could in part be enhanced due to slight differences in how the questions were asked. While the on-site participants were asked how often they are actually exposed to Australian English, web participants were asked more generally about their familiarity with the variety. Analyses reported below will take these potential group differences into account.

2.1.2. Material

Stimuli were based on a subset of stimuli from Penney et al. (submitted). The minimal word pairs *bart-bard* and *beat-bead* were selected to compare the perception of final stop voicing following low and high vowels (i.e., /v:/ versus /i:/). As glottalization tends to occur cross-linguistically more frequently on low than high vowels (Brunner & Żygis, 2011; Hejná & Scanlon, 2015; Malisz, Żygis, & Pompino-Marschall, 2013; Penney et al., 2018b; Pompino-Marschall & Żygis, 2010), it could be hypothesized that glottalization may either be easier to use in combination with low vowels since it is more common or with high vowels because it may stand out more since it is less usual. Only long vowels were chosen for this analysis, in order to limit any potential effects that may be related to

Table 1: Summary of the language-background questionnaires for the on-site and web participants in Experiment 1 and 2. All judgments were given on a 7-point scale where 1 indicated “not at all” and 7 “very much” or “very often.” The table shows the mean ratings and the standard deviation in brackets.

	Experiment 1		Experiment 2	
	on-site	web	on-site	web
frequency hearing English	4 (1.71)	5.6 (1.22)	4.17 (1.88)	5.15 (1.71)
strength of own German accent in English	4.61 (1.19)	4.04 (1.67)	4.05 (1.16)	3.38 (1.31)
annoyance by others' German accent	3.83 (1.76)	4.89 (1.64)	4.44 (1.76)	5.44 (1.34)
frequency writing English	2.33 (1.57)	4.39 (1.93)	2.5 (1.58)	4.47 (1.81)
frequency reading English	4.44 (1.29)	5.89 (1.29)	4.06 (1.3)	5.59 (1.62)
familiarity with Australian English	1.05 (0.93)	2.29 (1.25)	1 (0.97)	2.66 (1.52)

inherent vowel length, and to maximize the perceivable differences in duration, as long vowels demonstrate a greater range of duration across coda voicing contexts compared to short vowels (Cox & Palethorpe, 2011; Penney et al., 2018b).

The stimuli were based on natural speech tokens recorded by a 25-year old female native speaker of Australian English who lived in Sydney. Recordings were made in a sound attenuated booth and the tokens with the best recording quality of multiple repetitions of each of the four words were selected as source tokens for further manipulation. Tokens were selected to have a fully modal articulation of the vowel, and a falling intonation contour. In addition, the speaker produced each vowel with sustained creaky voice. These productions were used to create the glottalized tokens for the experiment. Note that 'real' glottalization could not be used for our purpose as then potentially additional cues to the stop's phonological voicing could have been retained.

The four source tokens were then further edited in Praat (Boersma & Weenink, 2018). That is, in the end two continua were created for each minimal pair, one in which the onset /b/ and the vowel were originally produced in the context of a voiceless coda and one in the context of a voiced coda. This was to control for any potential effects of the source that could not be taken care of by the manipulations described below. Results will be reported collapsed across sources. For each of the four source tokens, any cues to final stop voicing other than the ones to be investigated in the present study were removed. The closure duration was set to silence in all instances and cut to a duration that was the mean coda closure duration across voiceless and voiced codas following the respective preceding vowel, as reported in Penney et al. (2018b). The closure duration following /ɛ:/ (i.e., for *bart-bard*) was 75 ms, and following /i:/ (i.e., for *beat-bead*) 82 ms. The bursts were replaced by one low-amplitude burst taken from an originally voiced token. Piloting of the stimuli suggested that the selected burst was perceived as ambiguous between voiced and voiceless by native speakers of Australian English. Additionally, the cue to voicing due to a drop in F1 in the vowel preceding the closure was removed by cutting back the respective portion of the vowel.

An equally spaced 9-step vowel duration continuum was created from the four modal source tokens using the Pitch Synchronous Overlap and Add (PSOLA) function in Praat (Boersma & Weenink, 2018). The endpoint durations were based on previously collected production data (reported in Penney et al., 2018b) such that the shortest duration of the continuum (i.e., step 1) was two standard deviations shorter than the reported mean duration of the relevant vowel as produced in a voiceless coda context, and the longest step (i.e., step 9) was two standard deviations greater than the mean duration of the vowel in a voiced coda context. The continua hence spanned a rather large range of durations, which was for /ɛ:/ from 165 to 370 ms (step size = 25.6 ms), and for /i:/ from 104 to 340 ms (step size = 29.5 ms). For each continuum step for each token, F0 was set to 265 Hz at the beginning of the vowel and to 203 Hz at vowel offset. These values were based on the speaker's mean F0 values across all tokens that were recorded for the experiment in Penney et al. (submitted). The intensity contours of all tokens were equalized and scaled to 70dB.

In order to test the effects of glottalization, a 'glottalized' set of each of the continua was created. This was done by manually replacing the last 25% of the vowel with glottalization which was taken from the creaky productions of the speaker that were recorded for this purpose. The chosen proportion of glottalization for the vowels was based on the Australian English production study reported in Penney et al. (2018b). The method of splicing natural portions of creaky vowels into the continua was piloted in Penney et al. (submitted) and assessed to result in the most naturally sounding glottalized

tokens compared to other methods such as creating a sudden drop in the F0 contour using PSOLA.² Portions of creaky vowels were chosen for the glottalized part of the vowels to avoid the accidental addition of cues to a voiceless coda that might have been present in natural glottalization, which generally occurs preceding voiceless stops. **Figure 1** shows spectrograms of the modal and glottalized conditions used in Experiment 1 (as well as the additional glottalization = silence condition tested in Experiment 2) for the middle step of the continua for the source *bart*.

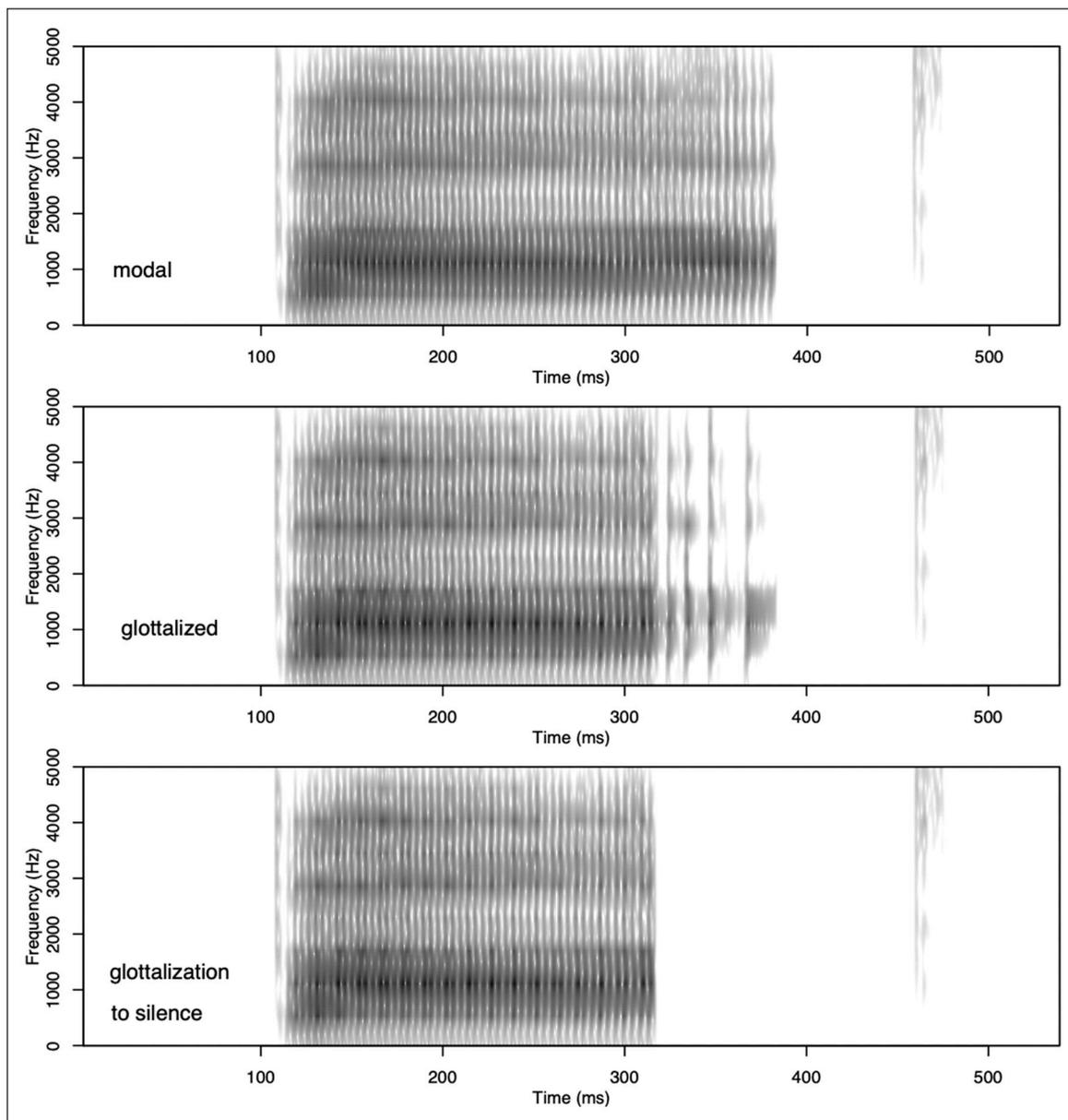


Figure 1: Spectrograms illustrating the different conditions. They show the middle step of the continuum of the source *bart* in the condition with a fully modal vowel (baseline condition; top panel), with the last 25% of the vowel glottalized (experimental condition in Experiment 1; middle panel), and with the glottalized portion of the vowel set to silence (experimental condition in Experiment 2; lower panel).

² Note that by splicing in portions of the creaky vowel we also created a radical drop in F0. However, the natural irregularities in glottal pulsing in the creaky tokens made the glottal portion sound more natural than only an artificially created drop in F0.

2.1.3. Procedure

On-site participants were seated in a sound-attenuated booth in front of a notebook computer and were fitted with headphones. They used the computer keyboard (number-keys 1 and 0 for the left and right response option respectively) to log their response. Web participants used equipment as described below. All participants received written instructions in English. To ensure that all participants were familiar with the four words that were used in the experiment, they received written explanations of the words and were shown pictures of the objects (e.g., Bart, a proper name, with the picture of the cartoon character Bart Simpson). They were told that on each trial they would hear an English word and their task was to decide which of two words they heard. On each trial, first the two response options were presented orthographically on the screen. The allocation of words to the left and right side of the screen was constant for a given participant but counterbalanced over participants. After 500 ms the word was presented auditorily at a comfortable listening level. Five hundred milliseconds after the participants logged their response the next trial started automatically. There was no time limit to respond. The two minimal word pairs (*bart-bard* and *beat-bead*) were presented blocked with the order counterbalanced across participants. Before each block, participants were presented three times with naturally produced tokens of the two words for this block (i.e., natural endpoint stimuli). For each minimal pair all nine continuum steps (built from the two sources) of the baseline modal-vowel condition and the experimental condition with the glottalized vowel were presented four times each in a fully randomized order for a total of 144 trials per minimal pair (9 steps × 2 sources × 2 conditions × 4 repetitions). The experiment for on-site participants was implemented in PsychoPy2 (Peirce, 2007).

The web-based experiment was tightly modelled on the on-site experiment using the software Percy (Draxler, 2014). When asked about the environment and equipment that web participants used, 13 reported to undertake the experiment at home, and five in their office. Twenty-four reported using headphones (as suggested in the instructions), while four used speakers. Almost all of them (27 of 28) used a notebook or desktop computer (i.e., using keys to log their responses as did the on-site participants). Only one participant reported using a tablet and hence the touchscreen to log responses. Independent of the location (on-site versus web), the experiment took approximately 15 minutes to complete.

2.2. Results

Results were analyzed using generalized linear mixed-effects models as implemented in the lme4 package (version 1.1-21, Bates et al., 2015) in R (version 3.5.2, R Core Team, 2018). The dependent variable was final stop voicing (voiced versus voiceless stop, with voiced coded as 1 and voiceless as 0) for which a logistic linking function was used. The first model tested listeners' use of vowel duration and glottalization in categorizing the words depending on whether they completed the experiment on-site in the lab versus via the internet. That is, fixed factors were Vowel Duration (continuum steps coded to be centred on zero and re-scaled to range from -1 to 1), Glottalization (modal vowel versus glottalized vowel coded as 0.5 and -0.5 respectively), and Testing Location (on-site versus web coded as -0.5 and 0.5). All interactions between the three fixed factors were included. Any interaction involving Testing Location would suggest that the two sets of participants behaved differently, possibly due to some differences in their exposure to English as indicated by the answers to the language background questionnaire (see **Table 1**). A random intercept was estimated for participants with random slopes for Vowel Duration, Glottalization, and their interaction over participants (i.e., all within-participant factors; Barr, Levy, Scheepers, & Tily, 2013).

Figure 2 illustrates the results and **Table 2** reports the statistics. Participants based their responses on the overall duration of the vowel with more voiced-stop responses the

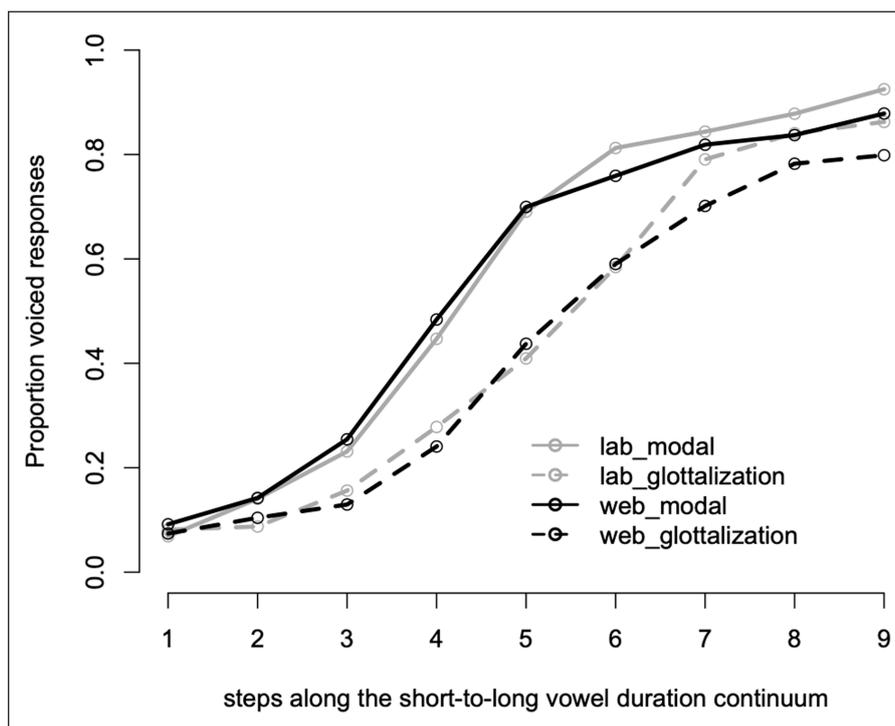


Figure 2: Proportion voiced responses over continuum steps in Experiment 1 comparing results from the on-site participants (grey lines) and the web participants (black lines). Solid lines represent the modal vowel condition, the dashed lines the glottalized condition.

Table 2: Summary of the statistics for the mixed-effects model testing the effects of vowel duration, glottalization, and testing location (on-site in the lab versus via the web) in Experiment 1.

Factor	<i>b</i>	<i>z</i>	<i>P</i>
intercept	0.059	0.58	.563
location	-0.055	-0.27	.786
glottalization	0.918	10.53	<.001
vowel duration	3.136	13.95	<.001
location : glottalization	0.101	0.59	.557
location : vowel duration	-0.287	-0.64	.522
glottalization : vowel duration	0.215	2.17	.030
location : glottalization : vowel duration	-0.096	-0.48	.630

longer the vowel (mean difference between responses to the shortest versus longest vowel 78.2%). They also used the glottalization such that the glottalized versions of the continua received fewer voiced-stop responses (mean difference between modal and glottalized continua 11.5%). The interaction between Vowel Duration and Glottalization suggests that the effect of Glottalization was larger the longer the vowel. Critically, there was no effect of Testing Location or any interaction involving this factor. This suggests that participants that were tested on-site in the lab versus via the internet did not differ in their responses regarding their use of vowel duration and glottalization.

Given the lack of effect of Testing Location, data were collapsed across this factor to test whether there was a difference in the use of glottalization depending on the vowel, that is between the minimal pairs with the vowel /ɐ:/ versus /i:/. Vowel Duration and Glottalization were entered as fixed factors and coded as before. The factor Vowel was coded such that results for the vowel /ɐ:/ were mapped onto the intercept. Main effects

of Vowel Duration and Glottalization then reflect results for /ɛ:/ and interactions with the factor Vowel indicate differences of these effects for the vowel /i:/. A random intercept was estimated for participants with random slopes for Vowel Duration, Glottalization, and their interaction over participants.

The statistics are reported in Table 3 and Figure 3 illustrates the results. As in the results reported above, for the minimal pair containing the vowel /ɛ:/, participants based their responses on the overall duration of the vowel with more voiced-stop responses the longer the vowel (cf. effect of Vowel Duration) and they used glottalization such that fewer voiced responses were given if glottalization was present (effect of Glottalization).

Table 3: Summary of the statistics for the mixed-effects model testing the effects of vowel duration, glottalization, and vowel (/ɛ:/ in the minimal pair *bart-bard* and /i:/ in *beat-bead*) collapsed across all participants in Experiment 1.

Factor	<i>b</i>	<i>z</i>	<i>p</i>
intercept	0.083	0.80	.421
vowel_i	-0.057	-1.21	.228
vowel duration	2.736	12.02	<.001
glottalization	0.784	8.07	<.001
vowel_i : vowel duration	0.865	9.63	<.001
vowel_i : glottalization	0.321	3.40	<.001
vowel duration : glottalization	-0.03	-0.24	.808
vowel_i : vowel duration : glottalization	0.562	3.14	.002

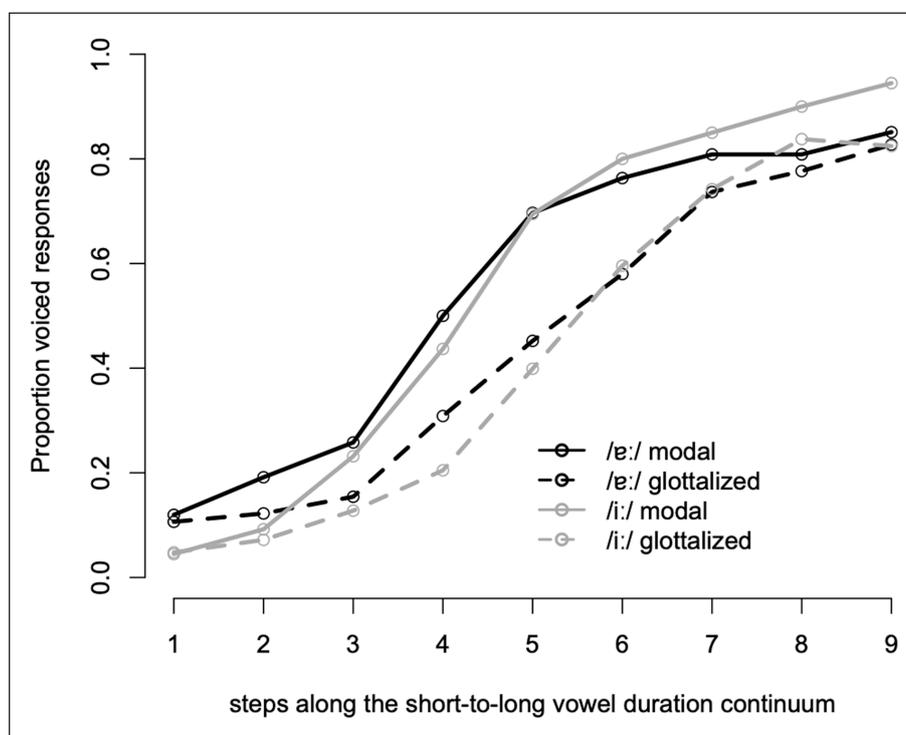


Figure 3: Proportion voiced responses over continuum steps in Experiment 1 comparing results between the two minimal pairs. Results for *bart-bard* with the low vowel /ɛ:/ are presented in black lines and for *beat-bead* with the high vowel /i:/ in grey lines. Solid lines represent the modal vowel condition, the dashed lines the glottalized condition.

Critically, all two-way interactions involving the factor Vowel (i.e., the comparison to the vowel /i:/) and the three-way interaction between all factors were significant suggesting differences in the effects of Vowel Duration and Glottalization between the two minimal pairs. Specifically, the interaction between Vowel and Vowel Duration suggests that the effect of vowel duration was stronger for /i:/ than /e:/. In other words, the categorization functions were steeper for *beat-bead* than *bart-bard*. The interaction between Vowel and Glottalization suggests that the effect of glottalization was larger for /i:/ than /e:/, and the three-way interaction suggests that the finding of a larger effect of glottalization at longer vowel durations was larger for /i:/ than /e:/. In fact, this interaction was not significant for /e:/ as indicated by the non-significant interaction between Vowel Duration and Glottalization (see also **Figure 3**).

2.3. Discussion

Experiment 1 set out to test whether German learners of English would use the vowel duration as well as the glottalization of the vowel preceding a word-final stop to distinguish the English word-final voicing contrast. Importantly, to native speakers of English glottalization marks the voiceless sound, that is, the sound of the contrast that is supposedly 'easy' for the learners as voiceless sounds are the default in word-final position in German. Since previous studies on L2 sound learning only focused on the use of cues to the unfamiliar sound we hence asked whether enhancing the familiar sound of the L2 contrast would also affect differentiation. Alternatively, German learners might have perceived glottalization as closure voicing, hence cueing a voiced coda. However, this was clearly not the case. Results showed that learners used the glottalization cue by giving fewer voiced responses in the glottalization condition than the condition with the fully modal vowels. Glottalization was used in addition to the overall duration of the vowel for both vowel contexts (where more voiced responses were given as vowel duration increased).

For the vowel /i:/ there was an even larger effect of glottalization at overall longer vowel durations. Since in the glottalization condition it was always the last 25% of the vowel that was glottalized, at longer vowel durations the absolute duration of glottalization was longer than at shorter vowel durations. Hence the glottalization was likely perceptually more salient at longer vowel durations. Note however, that at the same time a longer vowel duration points in the opposite direction as glottalization, namely towards a voiced coda.

This raises the question of how the glottalized portion of the vowel is perceived in the first place. Specifically, it could be the case that glottalization is not perceived as a cue on its own but as a change in the duration of the vowel or stop closure (or a combination of both). If the closure duration was perceived as 'lengthened' by the glottalized portion of the vowel this would predict the same results as we find. Similarly, listeners may have relied on only the modal portion of the vowel, which was 25% shorter than the respective vowel of the fully-modal vowel continuum. Again, this would predict our result of fewer voiced responses in the glottalized condition. The same holds for the overall resulting change in the ratio between the vowel + closure duration. German learners of English are familiar with such a trading relation between vowel and closure duration from their L1 where it is used to cue the stop voicing distinction in word-medial position. The question then is whether L2 learners would treat glottalization as a change in vowel or stop closure duration or whether the presence of glottalization has a (quantitatively) different effect.

Penney et al. (submitted) explored this possibility in their study of L1 Australian English listeners and found that listeners responded differently to stimuli with fully modally voiced (i.e., non-glottalized) vowels than they did to glottalized stimuli in which the modally voiced portion of the vowel was the same length as the vowels in the non-glottalized

stimuli. That is, they compared for instance responses to step 7 from the glottalized series of the /ɛ:/ vowel continuum (with a modal vowel duration of 239 ms) to step 4 of the fully modal vowel duration continuum (with a vowel duration of 241 ms). They found that listeners gave fewer voiced responses for the vowel with glottalization at step 7 than for the modal vowel at step 4, showing that the effect of glottalization was not due to the shorter modal portion of the vowel. Rather, given the same duration of the modal portion of the vowel, for L1 listeners of Australian English, the addition of a glottalized portion of the vowel enhanced the perception of voicelessness. Together with findings of an additional condition in which only closure duration was manipulated and in which they found that L1 listeners were not very sensitive to changes in closure duration this was interpreted as indicating that L1 listeners perceive glottalization as a distinct cue to stop voicelessness.

Nevertheless, it is possible that L2 learners perceive glottalization differently to native listeners. German learners are known to be sensitive to vowel duration, stop closure duration, and their trading relation from their L1 stop voicing contrast in word medial position. As mentioned above, they may perceive the glottalized portions of the vowel as a lengthened closure which would also predict our results. To test this possible explanation with our German learners, an additional experiment with a new group of learners was conducted. The experiment was identical to Experiment 1 except that the perception of the vowel duration continuum with a fully modal vowel was compared to the perception of a continuum in which the previously glottalized portion of the vowel had been set to silence. Given that a shorter modal vowel and a longer closure duration is expected to point towards a voiceless coda, we expected differences in the perception of coda voicing between the fully modal vowel condition (identical to the modal vowel condition in Experiment 1) and the new experimental condition with glottalization set to silence (i.e., resulting in a shorter modal vowel and longer closure). Critically, we hypothesized that if the glottalized portion of the vowel was perceived as lengthening the closure duration (or as changing the vowel + closure duration ratio) then responses to the new condition should be similar to the glottalized condition in Experiment 1. If a difference was found, then the effect of glottalization in Experiment 1 could be interpreted such that glottalization is treated as a separate cue to coda voicelessness that is different from the mere trading relation between vowel and closure duration.

3. Experiment 2

3.1. Method

3.1.1. Participants

Participants were recruited according to the same criteria and from the same pool of participants as in Experiment 1. None had participated in Experiment 1. Participants that were tested on-site at the University of Munich were assigned to Experiment 2 after a set of 18 participants³ had completed Experiment 1. Participants in the web experiment were randomly assigned to either Experiment 1 or 2. Eighteen participants completed Experiment 2 on-site, and data from 32 participants who completed the web experiment (27 female, 5 male, none other) were included in the analyses. Data from three further web participants were excluded, one due to a large age difference to the rest of the participants (this participant reported being 54 years of age while the upper limit for recruitment was 40 and all other participants were below this age), and two who reported native languages other than German (one Russian, one French).

³ Note that the 19th participant in Experiment 1 was tested last of all and arbitrarily assigned to Experiment 1 rather than 2. As for the participants in Experiment 1, the information about gender was lost during anonymization of the data.

The average age of the on-site participants in Experiment 2 was 23.2 ($SD = 2.6$, range 19–27) and of the web participants 28.3 ($SD = 4.8$, range 19–39). Six of the on-site and three of the web participants reported to model their English on a British variety, eight of the on-site and two of the web participants aimed at sounding American, three of the web participants aimed at sounding Australian, and the rest reported not to model their English on any specific variety. Answers to the other background questions that both groups filled in are reported in **Table 1** (along with the responses of participants from Experiment 1). Again, a number of differences can be seen between the on-site and web participants, most prominently in their age, use of English in reading and writing, as well as their overall reported familiarity with Australian English.

3.1.2. Material and Procedure

The materials and procedure of Experiment 2 were almost identical to Experiment 1 with the main difference being the experimental condition regarding glottalization. That is, for the baseline condition including the fully modal vowel the same stimuli were used as in Experiment 1. The vowel duration continua in which the last 25% of the vowel was glottalized, however, were manipulated such that the last 25% of the vowel was set to silence (we henceforth refer to this condition as 'glottalization = silence'). This was done using a Praat script (Boersma & Weenink, 2018). **Figure 1** (bottom panel) illustrates the manipulation compared to the baseline modal condition and the glottalized condition in Experiment 1.

Note that in the baseline condition with the fully modal vowel, only the vowel duration varies over the nine continuum steps with the closure duration being constant at 75 ms following /e:/ and 82 ms following /i:/ (see Method of Experiment 1). In contrast, the new stimuli in Experiment 2 with the glottalized portion of the vowel set to silence varied in their duration of the modal portion of the vowel as well as the duration of the closure. In other words, in the baseline condition with the fully modal vowel the shortest vowel duration for /e:/ was 165 ms with the closure of 75 ms and the longest vowel duration 370 ms with the same closure. Adding 25% of these vowels to the closure results in a ratio of 124 ms modal vowel to 116 ms closure at the short end of the continuum, and 278 ms modal vowel to 168 ms closure at the long end. For /i:/ the endpoints of continua with the fully modal vowel were 104 ms vowel to 82 ms closure at the short end and 340 ms to 82 ms closure at the long end. Adding 25% of the vowel to the closure results in ratios of 78 ms vowel to 108 ms closure at the short end, and 255 ms vowel to 167 ms closure at the long end of the continuum.

The task and procedure were identical to Experiment 1. Participants were first familiarized with the words and their meanings, and received a few trials with naturally spoken stimuli followed by the experimental stimuli. They had to indicate whether the word they heard ended in a voiced or voiceless stop by deciding which word of the minimal pair they heard. Stimuli were again blocked by vowel and the order was counterbalanced across participants. On-site participants were tested in the same sound attenuated booth and with the same equipment as the participants in Experiment 1. Web participants in Experiment 2 reported doing the experiment at home or in their office (17 versus 5 participants). Twenty-six used headphones as instructed, six used speakers. Twenty-seven used a notebook or desktop computer entering their responses via the keyboard, five used a tablet or smartphone and entered responses via the touchscreen. Since Experiment 2 had the same number of conditions and trials as Experiment 1 it also took 15 minutes to complete.

3.2. Results

Results were analyzed as in Experiment 1 using generalized linear mixed-effects models with a logistic linking function. The first model tested listeners' use of vowel duration and the effect of the experimental condition in which the glottalization had been set to

silence, depending on whether participants completed the experiment on-site in the lab versus via the web. That is, fixed factors were Vowel Duration (continuum steps coded to be centred on zero and re-scaled to range from -1 to 1), Glottalization = Silence (modal vowel versus vowel with glottalization set to silence; coded as 0.5 and -0.5 respectively), Testing Location (on-site versus web, coded as -0.5 and 0.5), and all interactions. A random intercept was estimated for participants with random slopes for Vowel Duration, Glottalization = Silence, and their interaction over participants.

Figure 4 illustrates the results and Table 4 reports the statistics. As in Experiment 1 participants based their responses on the overall duration of the vowel with more voiced-stop responses the longer the vowel (mean difference between longest and shortest vowel 79.6%). They also used the longer closure duration due to the silenced glottalized vowel 79.6%). They also used the longer closure duration due to the silenced glottalized

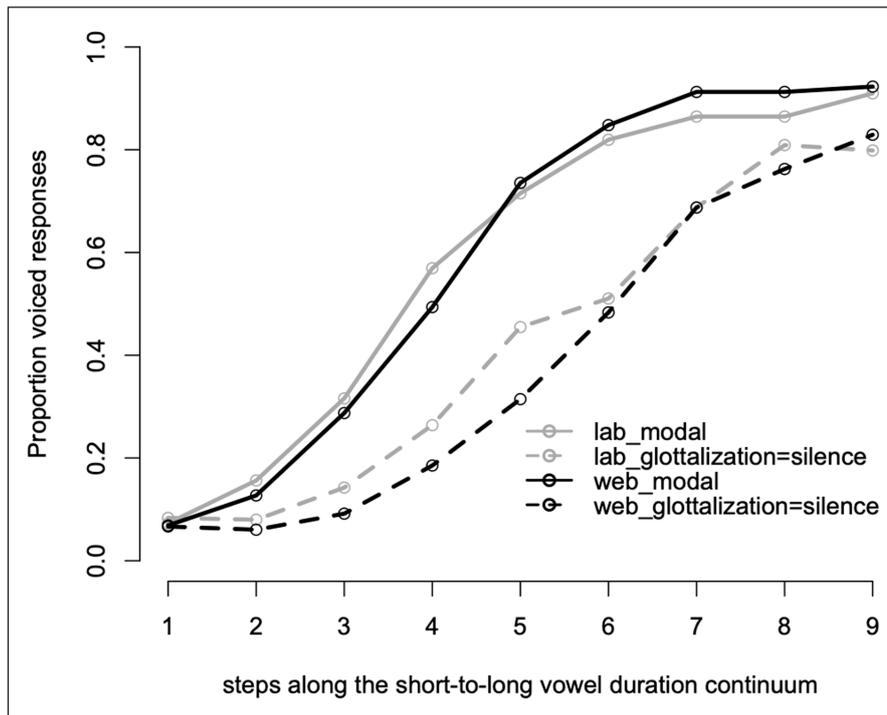


Figure 4: Proportion voiced responses over continuum steps in Experiment 2 comparing results from the on-site participants (grey lines) and the web participants (black lines). Solid lines represent the modal vowel condition, the dashed lines the condition in which the last 25% of the vowel (corresponding to the glottalized portion in Experiment 1) was set to silence.

Table 4: Summary of the statistics for the mixed-effects model testing the effects of vowel duration, glottalization=silence, and testing location (on-site in the lab versus via the web) in Experiment 2.

Factor	b	z	p
intercept	0.058	0.57	.566
location	-0.136	-0.67	.503
glottalization	1.424	14.37	<.001
vowel duration	3.067	17.87	<.001
location : glottalization	0.355	1.81	.071
location : vowel duration	0.309	0.91	.365
glottalization : vowel duration	0.428	4.60	<.001
location : glottalization : vowel duration	0.142	0.76	.445

portion of the vowel such that fewer voiced-stop responses were given if silence was added (mean difference between modal and glottalization = silence continua 18.8%). The interaction between Vowel Duration and Glottalization = Silence suggests that the effect of Glottalization = Silence was larger the longer the vowel was. Critically, there was no effect of Testing Location and none of the interactions involving this factor reached significance at the $p < .05$ level. This suggests that participants that were tested on-site in the lab versus via the internet did not statistically differ in their responses including their use of vowel duration and the longer closure due to silencing the glottalized portion of the vowel (see **Figure 4**, which shows that both types of cues were used).

Having established that participants did not differ based on their location, the main question could be addressed as to whether participants' responses in Experiment 2 differed from those in Experiment 1. Given the differences in the effect of Glottalization in Experiment 1 between the two vowels, the comparison between Experiment 1 and 2 was done separately for the two minimal pairs.⁴ The experiment was coded such that results from Experiment 2 were mapped onto the intercept. Effects of Vowel Duration and Glottalization = Silence then corresponded to these effects in Experiment 2. Interactions with Experiment would further indicate that the respective effects differed from those in Experiment 1. The random effects structure included a random intercept for participants with random slopes for Vowel Duration, Glottalization = Silence, and their interaction over participants.

Results are illustrated in **Figure 5** and statistics are reported in **Table 5**. The effects of Vowel Duration and Glottalization = Silence for both vowels confirm what has been shown in the analysis above, namely that participants in Experiment 2 used vowel duration as

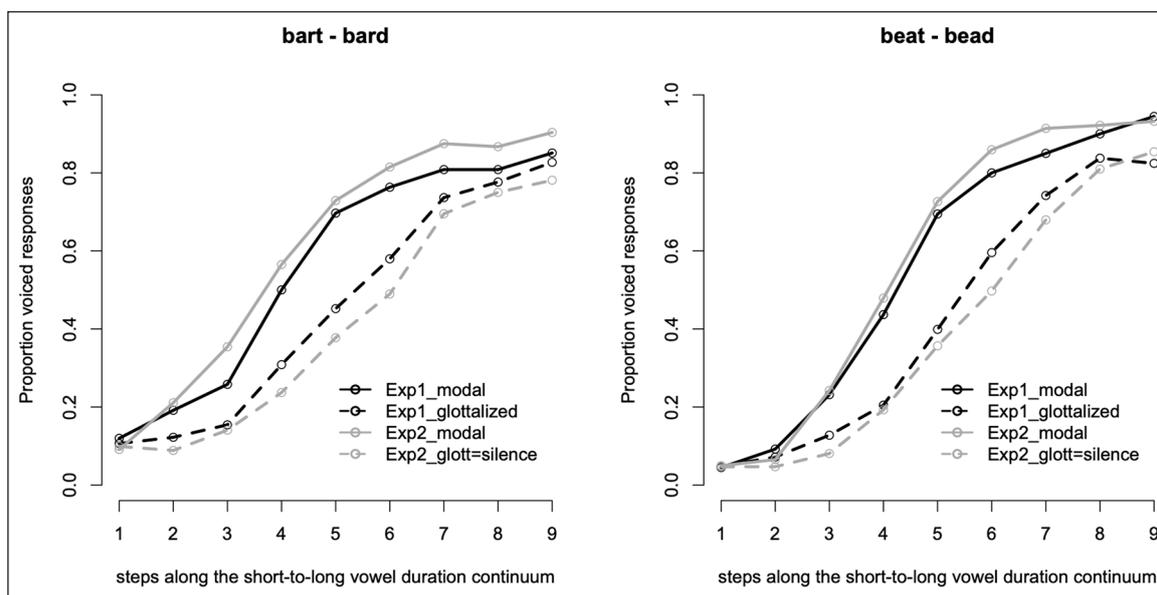


Figure 5: Results split by minimal pair/vowel quality for the proportion voiced-stop responses over continuum steps in Experiment 1 (black lines) and Experiment 2 (grey lines) for the baseline condition with the fully modal vowel (solid lines) and the experimental conditions with the glottalized vowel in Experiment 1 (black dashed line), and with glottalization set to silence in Experiment 2 (grey dashed lines).

⁴ We originally fitted a more complex model including all interactions with Vowel, however, this model did not converge. In addition, such a complex model would be hard to interpret and would distract from the main focus of the cross-experiment comparison of effects of Vowel Duration and Glottalization condition (i.e., Glottalization in Experiment 1 versus Glottalization = Silence in Experiment 2). Therefore, we report the simpler models conducted separately for each minimal pair.

Table 5: Summary of the statistics for the mixed-effects models for each vowel, testing the effects of vowel duration, and glottalization/glottalization=silence across experiments (the latter factor is abbreviated as 'glottalization').

Factor	bart – bard			beat – bead		
	<i>b</i>	<i>z</i>	<i>p</i>	<i>b</i>	<i>z</i>	<i>p</i>
intercept	0.094	0.82	.415	-0.010	-0.08	.936
vowel duration	2.875	12.35	<.001	3.705	18.26	<.001
glottalization	1.479	12.58	<.001	1.557	14.70	<.001
Experiment	0.016	0.10	.924	0.013	0.08	.941
vowel duration : glottalization	0.366	2.77	.006	0.659	4.10	<.001
vowel duration : experiment	0.037	0.11	.910	-0.10	-0.35	.725
glottalization : experiment	-0.648	-3.93	<.001	-0.457	-3.11	.002
vowel duration : glottalization : experiment	-0.355	-1.89	.059	-0.124	-0.55	.582

well as the longer closure duration due to the silenced glottalized portion of the vowel in categorizing the word-final stops as voiced versus voiceless. The longer the vowel (both /ɛ:/ and /i:/) the more voiced-stop responses were given and fewer voiced responses were given in the glottalization = silence condition than baseline condition with the fully modal vowel. For both vowels we find that the effect of Glottalization = Silence is larger at longer vowel durations (cf. interaction with Vowel Duration). Critically, for both vowels an interaction between Glottalization = Silence and Experiment was found indicating that the effect of Glottalization = Silence in Experiment 2 was larger than the effect of Glottalization in Experiment 1.⁵

3.3. Discussion

In Experiment 2 we set out to test whether results in Experiment 1 could be explained by listeners perceiving the glottalized portion of the vowel in the experimental condition in a similar manner as a lengthened stop closure duration (or change in vowel + closure duration ratio). This was done by comparing effects of glottalization in Experiment 1 to effects of a condition in which the glottalized portion of the vowels was set to silence. While both types of manipulations led German learners of English to perceive the respective tokens as voiceless more often than tokens of the baseline condition with the fully modal vowel, the effect was significantly larger in Experiment 2 than in Experiment 1. This suggests that the glottalized portion of the vowel in Experiment 1 is at least to some extent perceived as a cue to voicelessness that has a different perceptual effect than simply a longer closure.

4. General Discussion

The present study set out to test whether and to what extent German learners of English use vowel duration and the glottalization of the vowel as cues to the English word-final stop voicing contrast. The cue of vowel duration was chosen because it was considered familiar to our German learners from their L1. Results with regard to this cue were in line with previous studies on learners from a variety of language backgrounds (Crowther & Mann, 1992; Flege & Hillenbrand, 1987; Hayes-Harb et al., 2008), and showed that learners gave

⁵ Given that the continua with modal vowels were identical between the two experiments, additional analyses followed up on apparent differences for this condition as indicated by the solid lines in Figure 5. The models reported above were hence re-fitted with Glottalization coded such that the modal vowel condition was mapped onto the intercept and the effect of Experiment referred to only the modal vowel condition. This effect was not significant for either the *bart-bard* continuum ($b = 0.32, z = 1.53, p = .125$) or *beat-bead* ($b = 0.23, z = 1.18, p = .240$).

more voiced stop responses the longer the duration of the preceding vowel was. The cue of glottalization was chosen because it is less well known as a cue to the English stop voicing contrast, and although German learners are not entirely unfamiliar with this cue, it does not have counterparts in the learners' L1 that are as clear as those for vowel duration. Most importantly, in contrast to other well-studied cues to the word-final stop voicing distinction, glottalization cues the sound that for our learners is typically considered the familiar, 'easy' sound of the contrast since German as well as other languages neutralize this contrast word-finally towards the voiceless sound (Fourakis & Iverson, 1984). Given that previous studies have focused on cues to the 'difficult' sounds, and good versus poor learners mainly differed in their recognition of words with the difficult sounds in various contrasts (Eger & Reinisch, 2019), it was not clear from the outset that L2 learners would use glottalization as a cue to voicelessness. Results of the present study showed that German learners of English were able to use glottalization in perception such that they reported hearing the word ending in the voiceless sound more often if the preceding vowel was partially glottalized. This effect was on top of the effect of vowel duration.

One additional reason for testing the use of glottalization as a cue to the voicelessness of the final stop was that while it has been shown to be used by native speakers of American English (Chong & Garellek, 2018; Garellek, 2011) and Australian English (Penney et al., 2018a, submitted), in Australian English it appears to be becoming an increasingly important cue to this contrast (Penney et al., 2018b). In Australian English, glottalization promotes perception of stop voicelessness, even when paired with extended preceding vowel duration, which is otherwise a strong cue to final stop voicing (Penney et al., 2018a, submitted). Although the German learners showed increased perception of final stop voicelessness when glottalization was present, the effect was smaller than what has been shown for native listeners (Penney et al., 2018a, submitted). Results from the conditions in the study with Australian listeners that matched the stimuli used in the present Experiment 1 (note that the study by Penney, Cox, and Szakay had a different focus and tested several additional conditions) showed that Australian listeners had an effect of glottalization of 30% as calculated by the difference between voiced responses in the modal versus glottalized condition averaged across the steps of the duration continuum. For the Germans the effect of glottalization across vowel durations was merely 11.5% (see also **Figures 2 and 3**). In sum, for both native speakers of Australian English as well as German learners of English glottalization is a cue to the voiceless stop in perception, however, based on this descriptive comparison our learners utilized the glottalization cue less than native speakers.

Critically, Experiment 2 demonstrated that the effect of glottalization for our learners in Experiment 1 was indeed due to the glottalized portion of the vowel rather than to listeners perceiving a change in the relation between vowel and closure duration—a relation that cues the stop voicing contrast in word-medial position in German. We hypothesized that if learners simply perceived the glottalized stimuli as having a shorter modal vowel and longer closure then setting the glottalized portion of the vowel to silence should yield the same results since this weighting of cues would also point towards the voiceless stop. However, it was found that relative to the fully modally voiced vowel condition, in the glottalization=silence condition (Experiment 2) significantly more voiceless responses were given than in the glottalized condition in Experiment 1. This suggests that learners assess glottalization as a cue to voicelessness that is quantitatively different from the effect of a prolonged closure. Importantly, this effect of more voiceless responses in the glottalization=silence condition points to another difference with the native Australian listeners in Penney et al. (submitted). In their comparison of responses to continuum steps from the fully modal versus glottalized vowel continua with matched durations of the modal portions of the vowels, listeners gave more voiceless responses in

the glottalized condition. In contrast, the German learners gave more voiceless responses in the glottalization = silence condition in Experiment 2 than in the glottalized condition in Experiment 1.⁶ This suggests that while neither L1 Australian English listeners nor German learners of English appear to simply perceive the presence of glottalization as a change in the vowel + closure duration ratio, the perceptual effects differ. This likely reflects the overall different degrees of reliance on the glottalization cue as discussed above (as well as an apparent diminished reliance of Australian listeners on vowel duration).

These overall effects for the German learners were found for both minimal word pairs that we tested—*bart-bard* with the low vowel /ɛ:/ and *beat-bead* with a high vowel /i:/ preceding the stop contrast. Previous studies have shown that low vowels are more likely to be glottalized than high vowels (Brunner & Żygis, 2011; Hejná & Scanlon, 2015; Malisz et al., 2013; Penney et al., 2018b; Pompino-Marschall & Żygis, 2010) and Australian English listeners show a larger effect of glottalization as a cue to voicelessness in a low than high vowel context (Penney et al., submitted). Interestingly, for our learners the opposite was found (despite the identical materials as in the Australian study). Learners showed a larger effect of vowel duration (as evidenced in steeper categorization curves) and a larger effect of glottalization as a cue to voicelessness for the minimal pair with the vowel /i:/ (*beat-beat*) than /ɛ:/ (*bart-bard*). In addition, for the vowel /i:/ but not /ɛ:/ the effect of glottalization was larger at longer vowel durations (and hence longer portions of glottalization). A possible explanation for this discrepancy between the German learners and native listeners of Australian English could be that due to the higher frequency of glottalization on low than high vowels cross-linguistically, the glottalization stood out more in the high vowel context. Differences in the lexical frequency and learners' differential familiarity with the four words should have been diminished by familiarizing participants with the four words right before the experiment. They received written explanations of the concepts, saw pictures, and during the first few trials they listened to natural productions of the words that included all cues that the speaker would naturally produce.

Importantly, glottalization as a cue to differentiate between the sounds of the English word-final stop voicing contrast specifically marks the voiceless sound (Chong & Garellek, 2018; Docherty & Foulkes, 1999; Gordeeva & Scobbie, 2013; Huffman, 2005; Penney et al., 2018a, submitted; Penney et al., 2018b; Pierrehumbert, 1995; Roach, 1973). The present results then show that learners can also use a cue that marks or 'enhances' the sound that for them is the familiar, 'easy' one of the L2 contrast. This suggests that it is the general enhancement of the contrast, in this case via an additional cue to the 'easy' sound, that helps in L2 perception regardless of the learners' ability to establish the new 'difficult' sound as a distinct category from the 'stable' familiar sound that is likely a simple transfer from the L1 (Flege, 1995; Best & Tyler, 2008). A general enhancement of a given contrast may be especially relevant in tasks with a phonetic focus such as the present phonetic identification task where on every trial, learners had to make a forced-choice decision about the voicing of the final sound (i.e., by picking one of two orthographically presented words).

It will remain speculation as to how learners may perform in other speech perception tasks, specifically tasks in which the focus is on lexical processing and word meaning rather than phonetic perception. Previous studies have shown that learners may be very

⁶ We acknowledge that in the comparison of continuum steps for the Australian listeners, the closure duration stayed constant, while setting the glottalized portion of the vowels to silence added to the closure possibly providing an additional cue to voicelessness. Note, however, that in another experiment of Penney et al. (submitted), Australian listeners did not appear to be very sensitive to a manipulation of closure duration. Comparisons of continuum steps from Experiment 1 similar to the comparisons for the Australian listeners confirm the results of Experiment 2. Unlike the Australian listeners, German learners gave fewer voiceless responses for the short modal vowels than for vowels that had the same modal vowel duration but glottalization added.

good at phonetic tasks such as identification or discrimination of difficult sound contrasts even if their ability to differentiate between the respective sounds in lexical tasks (e.g., lexical decision or spoken-word recognition as shown by eye-tracking paradigms) is much lower (e.g., Amengual, 2016; Darcy, Daidone, & Kojima, 2013; Díaz, Mitterer, Broersma, & Sebastián-Gallés, 2012; Llompart & Reinisch, 2019, in press; Sebastián-Gallés & Baus, 2005; Sebastián-Gallés, Echeverría, & Bosch, 2005). That is, the ability to phonetically identify sounds of a difficult L2 contrast does not automatically mean that it is also faithfully encoded in the L2 lexicon. Assessing the use of a phonetic cue to the 'easy' sound of a difficult L2 contrast in lexical access will hence remain for future research.

Another issue that may be relevant with regard to learners' use of L2 phonetic cues are individual differences. In the present case we tested two groups of participants, one on-site at the university and the other via the web. Answers to the background questionnaire about their use of English suggested certain group-level differences, for instance, in their desire to model their English on a specific accent or their overall exposure to English. However, no overall differences could be found between the two groups for either the use of the familiar cue of vowel duration or the 'unfamiliar' cue of glottalization. Note, however, that although such group-level differences are typically easier to detect than correlations at the individual learner level (e.g., Llompart & Reinisch, in press), ratings from language background questionnaires rarely predict effects in phonetic tasks—likely due to the overall good performance and relatively small variability in phonetic tasks (compared to lexical tasks; Eger & Reinisch, 2017; Llompart & Reinisch, 2019).

Looking at the comparison between on-site and web participants from another perspective, the lack of difference found in the present study suggests that conducting phonetic experiments via the web may be a feasible method to use. Conducting an experiment via the web may be useful if a large number of participants is required and the experiment is relatively short. Note that our manipulation of glottalizing the last 25% of the vowel could be considered rather subtle, yet participants' reactions in the web experiment mirrored those from the lab, despite their different equipment and our lack of control of the experimental situation. The only drawback that we noticed is that many participants complained about the experiment being not engaging enough, which via the web led to a large proportion of interruptions. Only half of the web participants who started the experiment (i.e., 61 out of 123) completed the full number of trials (which took about 15 minutes). Nevertheless, web experiments may be a suitable way to increase the total number of participants, which may be specifically useful for testing second language learners who as a group are typically quite variable in their language background as well as their performance in at least some tasks.

In sum, the present study showed that German learners of English use glottalization in the vowel preceding a word-final stop as a cue to voicelessness of this sound. Critically, the glottalization in the vowel is perceived as quantitatively different from a simple change in the ratio of vowel and closure duration since a prolonged closure leads to even more voiceless responses than glottalizing the same portion of the vowel. We suggest that at least in L2 perception tasks with a phonetic focus, learners rely on cues that enhance the contrast regardless of whether the cue marks the 'easy,' familiar or the 'difficult,' unfamiliar sound. This is despite the fact that—similarly to other phonetic cues to L2 contrasts—learners' reliance on glottalization may differ from native speakers' usage of this cue.

Acknowledgements

This project was funded by a grant from the German Research Foundation (DFG; grant nr. RE3047/1-1) to the first author and an Australian Government Endeavour Research Fellowship (6331_2018) to the second author. We would like to thank Rosa Hufschmidt for

running the on-site participants and Christoph Draxler for help with the web experiment. We would like to thank Miquel Llompart for comments on an earlier version of the manuscript.

We also thank the Linguistics in Open Access foundation (<https://www.lingoa.eu/>) for financial support of open access publication of *Laboratory Phonology*.

Competing Interests

The authors have no competing interests to declare.

References

- Amengual, M. (2016). The perception of language-specific phonetic categories does not guarantee accurate phonological representations in the lexicon of early bilinguals. *Applied Psycholinguistics*, 37(5), 1221–1251. DOI: <https://doi.org/10.1017/S0142716415000557>
- Baker, W. (2010). Effects of age and experience on the production of English word-final stops by Korean speakers. *Bilingualism: Language and Cognition*, 13, 263–278. DOI: <https://doi.org/10.1017/S136672890999006X>
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. DOI: <https://doi.org/10.1016/j.jml.2012.11.001>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1). DOI: <https://doi.org/10.18637/jss.v067.i01>
- Bent, T., Bradlow, A. R., & Smith, B. L. (2008). Production and perception of temporal patterns in native and non-native speech. *Phonetica*, 65, 131–147. DOI: <https://doi.org/10.1159/000144077>
- Best, C. T., & Tyler, M. D. (2008). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn, & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13–34). Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/llt.17.07bes>
- Boersma, P., & Weenink, D. (2018). *Praat: Doing phonetics by computer* (Version 6.0.36). <http://www.praat.org/>
- Bohn, O.-S., & Flege, J. E. (1992). The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition*, 14(2), 131–158. DOI: <https://doi.org/10.1017/S0272263100010792>
- Broersma, M. (2005). Perception of familiar contrasts in unfamiliar positions. *Journal of the Acoustical Society of America*, 117, 3890–3901. DOI: <https://doi.org/10.1121/1.1906060>
- Brunner, J., & Żygis, M. (2011). Why do glottal stops and low vowels like each other? In W. Lee, & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences, Hong Kong, 17–21 August* (pp. 376–379).
- Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica*, 22, 129–159. DOI: <https://doi.org/10.1159/000259312>
- Chong, A., & Garellek, M. (2018). Online perception of glottalized coda stops in American English. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 9(1), 4. DOI: <https://doi.org/10.5334/labphon.70>
- Cox, F., & Palethorpe, S. (2011). Timing differences in the VC rhyme of standard Australian English and Lebanese Australian English. In *Proceedings of the 17th International Congress of Phonetic Sciences, Hong Kong, 17–21 August* (pp. 528–531).

- Crowther, C. S., & Mann, V. (1992). Native language factors affecting use of vocalic cues to final consonant voicing in English. *Journal of the Acoustical Society of America*, 92, 711–722. DOI: <https://doi.org/10.1121/1.403996>
- Darcy, I., Daidone, D., & Kojima, C. (2013). Asymmetric lexical access and fuzzy lexical representations in second language learners. *The Mental Lexicon*, 8(3), 372–420. DOI: <https://doi.org/10.1075/ml.8.3.06dar>
- Díaz, B., Mitterer, H., Broersma, M., & Sebastián-Gallés, N. (2012). Individual differences in late bilinguals' L2 phonological processes: From acoustic-phonetic analysis to lexical access. *Learning and Individual Differences*, 22(6), 680–689. DOI: <https://doi.org/10.1016/j.lindif.2012.05.005>
- Docherty, G. (1992). *The timing of voicing in British English obstruents*. Berlin, Germany: Foris. DOI: <https://doi.org/10.1515/9783110872637>
- Docherty, G., & Foulkes, P. (1999). Sociophonetic variation in 'glottals' in Newcastle English. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, & A. C. Bailey (Eds.), *Proceedings of the 14th International Congress of Phonetic Sciences, San Francisco, USA, August 1–7* (pp. 1037–1040).
- Draxler, C. (2014). Online experiments with the Percy software framework – experiences and some early results. In N. Calzolari, et al. (Eds.). *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)* (pp. 235–240). Reykjavik, Iceland: European Language Resources Association (ELRA).
- Eger, N. A., & Reinisch, E. (2017). The role of acoustic cues and listener proficiency in the perception of accent in nonnative sounds. *Studies in Second Language Acquisition*, online first view. DOI: <https://doi.org/10.1017/S0272263117000377>
- Eger, N. A., & Reinisch, E. (2019). The impact of one's own voice and production skills on word recognition in a second language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45, 552–571. DOI: <https://doi.org/10.1037/xlm0000599>
- Flege, J. E. (1989). Chinese subjects' perception of the word-final English /t-/d/ contrast: Performance before and after training. *Journal of the Acoustical Society of America*, 86, 1684–1697. DOI: <https://doi.org/10.1121/1.398599>
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). Timonium, MD: York Press.
- Flege, J. E., Bohn, O. S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, 25(4), 437–470. DOI: <https://doi.org/10.1006/jpho.1997.0052>
- Flege, J. E., & Hillenbrand, J. (1987). A differential effect of release bursts on stop voicing judgments of native French and English listeners. *Journal of Phonetics*, 15, 203–208.
- Flege, J. E., & Port, R. (1981). Cross-language phonetic interference: Arabic to English. *Language and Speech*, 24(2), 125–146. DOI: <https://doi.org/10.1177/002383098102400202>
- Flege, J. E., & Wang, C. (1989). Native-language phonotactic constraints affect how well Chinese subjects perceive the word-final English /t-/d/ contrast. *Journal of Phonetics*, 17, 299–315.
- Fourakis, M., & Iverson, G. K. (1984). On the 'incomplete neutralization' of German final obstruents. *Phonetica*, 41, 140–149. DOI: <https://doi.org/10.1159/000261720>
- Fowler, C. A. (1992). Vowel duration and closure duration in voiced and unvoiced stops: There are no contrast effects here. *Journal of Phonetics*, 20, 143–165.
- Garellek, M. (2011). The benefits of vowel laryngealization on the perception of coda stops in English. *UCLA Working Papers in Phonetics*, 109, 31–39.

- Garellek, M. (2015). Perception of glottalization and phrase-final creak. *The Journal of the Acoustical Society of America*, 137, 822–831. DOI: <https://doi.org/10.1121/1.4906155>
- Gordeeva, O. B., & Scobbie, J. M. (2013). A phonetically versatile contrast: Pulmonic and glottalic voicelessness in Scottish English obstruents and voice quality. *Journal of the International Phonetic Association*, 43, 249–271. DOI: <https://doi.org/10.1017/S0025100313000200>
- Hayes-Harb, R., Smith, B. L., Bent, T., & Bradlow, A. R. (2008). The interlanguage speech intelligibility benefit for native speakers of Mandarin: Production and perception of English word-final voicing contrasts. *Journal of Phonetics*, 36(4), 664–679. DOI: <https://doi.org/10.1016/j.wocn.2008.04.002>
- Hejrná, M., & Scanlon, J. (2015). New laryngeal allophony in Manchester English. In J. Stuart-Smith, J. Scobbie, & A. Turk (Eds.), *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, Scotland, 10–14 August*.
- Huffman, M. K. (2005). Segmental and prosodic effects on coda glottalization. *Journal of Phonetics*, 33, 335–362. DOI: <https://doi.org/10.1016/j.wocn.2005.02.004>
- Iverson, P., & Evans, B. G. (2007). English vowel training with different first-language vowel systems. *The Journal of the Acoustical Society of America*, 121(5), 3072–3072. DOI: <https://doi.org/10.1121/1.4781875>
- John, T., & Harrington, J. (2007). Temporal alignment of creaky voice in neutralised realisations of an underlying, post-nasal voicing contrast in German. In H. van Hamme, & R. van Son (Eds.), *Proceedings of Interspeech 2007* (pp. 986–989). Antwerp, Belgium.
- Keyser, S. J., & Stevens, K. N. (2006). Enhancement and overlap in the speech chain. *Language*, 82, 33–62. DOI: <https://doi.org/10.1353/lan.2006.0051>
- Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *The Journal of the Acoustical Society of America*, 59, 1208–1221. DOI: <https://doi.org/10.1121/1.380986>
- Kleber, F., John, T., & Harrington, J. (2010). The implications for speech perception of incomplete neutralization of final devoicing in German. *Journal of Phonetics*, 38, 185–196. DOI: <https://doi.org/10.1016/j.wocn.2009.10.001>
- Kohler, K. (1977). The production of plosives. *Arbeitsberichte des Instituts für Phonetik der Universität Kiel*, 8, 30–110.
- Kohler, K. J. (2000). Linguistic and paralinguistic functions of non-modal voice in connected speech. In P. Hoole (Ed.), *Proceedings of 5th seminar of speech production: Models and data* (pp. 89–92). Bavaria, Germany: Kloster Seeon, 1–4 May.
- Kohler, K. J. (2001). Plosive-related glottalization phenomena in read and spontaneous speech. A stød in German. To Honour Eli Fischer-Jørgensen. *Travaux du Cercle Linguistique de Copenhague*, 31, 174–211.
- Lisker, L. (1974). On explaining vowel duration variation. *Glossa*, 8, 233–246.
- Lisker, L. (1986). ‘Voicing’ in English: A catalogue of acoustic features signaling /b/ versus /p/ in trochees. *Language and Speech*, 29, 3–11. DOI: <https://doi.org/10.1177/002383098602900102>
- Llompert, M., & Reinisch, E. (2017). Articulatory information helps encode lexical contrasts in a second language. *Journal of Experimental Psychology: Human Perception and Performance*, 43(5), 1040–1056. DOI: <https://doi.org/10.1037/xhp0000383>
- Llompert, M., & Reinisch, E. (2019). Imitation in a second language relies on phonological categories but does not reflect the productive usage of difficult sound contrasts. *Language and Speech*, 62, 594–622. DOI: <https://doi.org/10.1177/0023830918803978>

- Llompарт, M., & Reinisch, E. (in press). Robustness of phonolexical representations relates to phonetic flexibility for difficult second language sound contrasts. *Bilingualism: Language and Cognition*. online first view. DOI: <https://doi.org/10.1017/S1366728918000925>
- Malisz, Z., Żygis, M., & Pompino-Marschall, B. (2013). Rhythmic structure effects on glottalisation: A study of different speech styles in Polish and German. *Laboratory Phonology*, 4, 119–158. DOI: <https://doi.org/10.1515/lp-2013-0006>
- Mathisen, A. G. (1999). Sandwell, West Midlands: Ambiguous perspectives on gender patterns and models of change. In P. Foulkes, & G. J. Docherty (Eds.), *Urban Voices: Accent studies in the British Isles* (pp. 107–123). London, UK: Arnold.
- Mitterer, H., & Reinisch, E. (2015). Letters don't matter: No effect of orthography on the perception of conversational speech. *Journal of Memory and Language*, 85, 116–134. DOI: <https://doi.org/10.1016/j.jml.2015.08.005>
- Mitterer, H., & Reinisch, E. (2017). Surface forms trump underlying representations in functional generalizations in speech perception: The case of German devoiced stops. *Language, Cognition and Neuroscience*, 32, 1133–1147. DOI: <https://doi.org/10.1080/23273798.2017.1286361>
- Nicenboim, B., Roettger, T. B., & Vasishth, S. (2018). Using meta-analysis for evidence synthesis: The case of incomplete neutralization in German. *Journal of Phonetics*, 70, 39–55. DOI: <https://doi.org/10.1016/j.wocn.2018.06.001>
- O'Dell, M., & Port, R. (1983). Discrimination of word-final voicing in German. *The Journal of the Acoustical Society of America*, 73, S31–S31. DOI: <https://doi.org/10.1121/1.2020331>
- Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1), 8–13. DOI: <https://doi.org/10.1016/j.jneumeth.2006.11.017>
- Penney, J., Cox, F., Miles, K., & Palethorpe, S. (2018b). Glottalisation as a cue to coda stop voicing in Australian English. *Journal of Phonetics*, 66, 161–184. DOI: <https://doi.org/10.1016/j.wocn.2017.10.001>
- Penney, J., Cox, F., & Szakay, A. (2018a). Weighting of coda voicing cues: Glottalisation and vowel duration. In *Proceedings of the 19th Annual Conference of the International Speech Communication Association, INTERSPEECH 2018, September 2–6, Hyderabad, India* (pp. 1422–1426). DOI: <https://doi.org/10.21437/Interspeech.2018-1677>
- Penney, J., Cox, F., & Szakay, A. (submitted). Effects of glottalisation and preceding vowel duration on the perception of coda stop voicing.
- Pierrehumbert, J. (1995). Prosodic effects on glottal allophones. In O. Fujimura, & M. Hirano (Eds.), *Vocal fold physiology: Voice quality control* (pp. 39–60). San Diego, CA: Singular.
- Piroth, H. G., & Janker, P. M. (2004). Speaker-dependent differences in voicing and devoicing of German obstruents. *Journal of Phonetics*, 32(1), 81–109. DOI: [https://doi.org/10.1016/S0095-4470\(03\)00008-1](https://doi.org/10.1016/S0095-4470(03)00008-1)
- Pompino-Marschall, B., & Żygis, M. (2010). Glottal marking of vowel-initial words in German. *ZAS Papers in Linguistics*, 52, 1–17.
- Port, R., & Crawford, P. (1989). Incomplete neutralization and pragmatics in German. *Journal of Phonetics*, 17, 257–282.
- Port, R., Mitleb, F., & O'Dell, M. (1981). Neutralization of obstruent voicing in German is incomplete. *The Journal of the Acoustical Society of America*, 70(S1), S13–S13. DOI: <https://doi.org/10.1121/1.2018716>
- Port, R., & O'Dell, M. (1985). Neutralization of syllable-final voicing in German. *Journal of Phonetics*, 13, 455–471.
- Raphael, L. J. (1972). Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *The Journal of the Acoustical Society of America*, 51, 1296–1303. DOI: <https://doi.org/10.1121/1.1912974>

- Raphael, L. J. (1981). Durations and contexts as cues to word-final cognate opposition in English. *Phonetica*, 38, 126–147. DOI: <https://doi.org/10.1159/000260019>
- Raphael, L. J., Dorman, M. F., Freeman, F., & Tobin, C. (1975). Vowel and nasal duration as cues to voicing in word-final stop consonants: Spectrographic and perceptual studies. *Journal of Speech, Language, and Hearing Research*, 18(3), 389–400. DOI: <https://doi.org/10.1044/jshr.1803.389>
- R Core Team. (2018). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from www.r-project.org/
- Roach, P. J. (1973). Glottalization of English /p/, /t/, /k/ and /tʃ/ – a reexamination. *Journal of the International Phonetic Association*, 3, 10–21. DOI: <https://doi.org/10.1017/S0025100300000633>
- Roettger, T. B., Winter, B., Grawunder, S., Kirby, J., & Grice, M. (2014). Assessing incomplete neutralization of final devoicing in German. *Journal of Phonetics*, 43, 11–25. DOI: <https://doi.org/10.1016/j.wocn.2014.01.002>
- Sebastián-Gallés, N., & Baus, C. (2005). On the relationship between perception and production in L2 categories. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones* (pp. 279–292). Hillsdale, NJ: LEA.
- Sebastián-Gallés, N., Echeverría, S., & Bosch, L. (2005). The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals. *Journal of Memory and Language*, 52(2), 240–255. DOI: <https://doi.org/10.1016/j.jml.2004.11.001>
- Seyfarth, S., & Garellek, M. (2015). Coda glottalization in American English. In J. Stuart-Smith, J. Scobbie, & A. Turk (Eds.), *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, Scotland, 10–14 August*.
- Smith, B. L., Hayes-Harb, R., Bruss, M., & Harker, A. (2009). Production and perception of voicing and devoicing in similar German and English word pairs by native speakers of German. *Journal of Phonetics*, 37(3), 257–275. DOI: <https://doi.org/10.1016/j.wocn.2009.03.001>
- Smith, B. L., & Peterson, E. A. (2012). Native English speakers learning German as a second language: Devoicing of final voiced stop targets. *Journal of Phonetics*, 40, 129–140. DOI: <https://doi.org/10.1016/j.wocn.2011.09.004>
- Tollfree, L. (1999). South East London English: Discrete versus continuous modelling of consonantal reduction. In P. Foulkes, & G. J. Docherty (Eds.), *Urban Voices: Accent studies in the British Isles* (pp. 163–184). London, UK: Arnold.
- Westbury, J. R., & Keating, P. A. (1986). On the naturalness of stop consonant voicing. *Journal of Linguistics*, 22, 145–166. DOI: <https://doi.org/10.1017/S0022226700010598>
- Wiese, R. (1996). *The Phonology of German*. Oxford, UK: Clarendon Press.

How to cite this article: Reinisch, E., & Penney, J. 2019 The role of vowel length and glottalization in German learners' perception of the English coda stop voicing contrast. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 10(1):18, pp. 1–26. DOI: <https://doi.org/10.5334/labphon.176>

Submitted: 22 October 2018

Accepted: 08 October 2019

Published: 24 October 2019

Copyright: © 2019 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.



Laboratory Phonology: Journal of the Association for Laboratory Phonology is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS