Explicit pronunciation instruction in the second language classroom

An acoustic analysis of German final devoicing

James M. Stratton
University of British Columbia

The present study uses an acoustic analysis to examine the effects of implicit and explicit pronunciation instruction on the acquisition of German final devoicing in the L2 classroom. Twenty-nine English-speaking L2 learners of German at a North American university were assigned to an implicit or explicit condition. Learner speech samples were recorded, following a pre/post/delayed-post-test design. Four acoustic correlates of final and medial obstruent voicing were analyzed to establish the degree to which underlyingly voiced word-final stops were phonetically devoiced. Results indicate that learners in the explicit condition significantly outperformed learners in the implicit condition, with all four acoustic measures signaling significantly greater word-final devoicing by the post-test in the explicit condition. Orthography, declarative knowledge, and level of awareness are hypothesized as factors that influenced the acquisition process. The study calls for additional acoustic work on the effects of different instructional practices on German L2 pronunciation.

Keywords: final devoicing, explicit pronunciation instruction, second language acquisition, German

1. Literature review

1.1 Pronunciation instruction

Although many studies suggest that L2 learners can benefit from explicit pronunciation instruction (Lee, Jang & Plonsky, 2015; Thomson & Derwing, 2015; Olson & Offerman, 2021), there is a clear language bias in the literature, with English and Spanish as a second language accounting for more than two thirds of the languages studied (Thomson & Derwing, 2015; Levis, 2019). The effects of explicit

https://doi.org/10.1075/jslp.22038.str | Published online: 20 July 2023
Journal of Second Language Pronunciation 9:1 (2023), pp. 71–102. ISSN 2215-1931 | E-ISSN 2215-194X

instruction on German L2 speech have received comparatively less attention. While there is little reason to believe that L2 learners of German would respond differently to explicit pronunciation instruction than learners of other languages, previous studies on German have found mixed results (Dlaska & Krekeler, 2013; Martin, 2018; McCandless & Winitz, 1986; Peltekov, 2020; Roccamo, 2015;). Some studies have reported that explicit pronunciation instruction or explicit pronunciation feedback made some learners more comprehensible (McCandless & Winitz, 1986; Dlaska & Krekeler, 2013; Roccamo, 2015; Martin, 2018), while others found no significant difference in comprehensibility and intelligibility ratings (Peltekov, 2020). Previous research on German may have also been affected by various methodological confounds, such as the lack of a pre-test and post-test, the use of different instructors for different experimental groups, not controlling for time on task, and comparing different learner populations (Stratton, 2022, p. 43).

Moreover, previous classroom-based studies on German (Dlaska & Krekeler, 2013; Martin, 2018; McCandless & Winitz, 1986; Peltekov, 2020; Roccamo, 2015) have relied heavily on impressionistic ratings to examine improvements in L2 pronunciation as a result of explicit instruction. While impressionistic ratings can measure changes in learner comprehensibility, their focus on larger strings of speech as opposed to individual segments, could mean that learners improve their L2 pronunciation, but such changes are not necessarily perceptible to native speaker raters. While one could argue that less noticeable improvements in pronunciation are less important to acquire, if improvements are not observed because of the methodological decision to evaluate pronunciation impressionistically, the research question regarding the effects of explicit pronunciation instruction on improvements in L2 pronunciation is nevertheless affected.

Therefore, the present study uses an acoustic analysis to examine changes in final devoicing in the speech of L2 learners of German after receiving either implicit or explicit pronunciation instruction. Since acoustic correlates of voicing are well established in the literature and have been used in several studies of final devoicing (Charles-Luce, 1985; Dmitrieva, Jongman & Sereno, 2010; Piroth & Janker, 2004; Port & Crawford, 1989; Port & O'Dell, 1985; Smith et al., 2009), the focus on final devoicing offers a unique opportunity to study the effects of different pedagogical practices on changes in L2 pronunciation. While it is currently unclear which aspects of German pronunciation carry the highest functional load, that is, the elements that are more likely to impact comprehensibility (Munro & Derwing, 2006), Peltekov (2020) suggests the failure to devoice word-final stops in German is associated with accentedness (p. 14). Since many learners care about accentedness (Derwing & Munro, 2015, pp. 131–152; Levis, 2016), as it can have social ramifications (Gluszek & Dovidio, 2010; Munro, 2003), reducing accentedness where possible appears to be an important goal for L2 learners

(Derwing & Munro, 2015, pp. 131–152). Research also suggests that many learners want to learn all aspects of pronunciation regardless of functional load (Huensch & Thompson, 2017; Sturm, Miyamoto & Suzuki, 2019) and reducing accentedness may increase learner confidence and reduce anxiety (Stratton, accepted).

1.2 German final devoicing

Final devoicing describes the process by which underlyingly voiced obstruents, most commonly, stops /b, d, g/, become systematically devoiced [p, t, k] in syllable-final or word-final position. English permits voiced stops in word-final position, so if English-speaking L2 learners of German use English phonology as a template, it is hypothesized that they will fail to devoice German stops word-finally. Because German makes no overt orthographic distinction between word-final (e.g., *Tag* 'day' [taːk]) and non-word-final stops (e.g., *Tage* 'days [taːgə]), it is hypothesized that orthography will also interfere with the acquisition of German final devoicing (Hayes-Harb, Brown & Smith, 2018; Hayes-Harb & Barrios, 2021).[1] In a laboratory study on the influence of orthographic input on the acquisition of German final devoicing, Hayes-Harb, Brown and Smith (2018) exposed native speakers of English, with no prior knowledge of German, to pseudowords ending in underlyingly voiced obstruents. Participants presented with both auditory and orthographic input produced more voicing than participants who received only auditory input, suggesting that orthography interfered with the perceptual status of the underlyingly voiced obstruents.

1.3 Acoustic properties

To measure the production of stop consonants, spectral and temporal acoustic measurements can be used, including duration, voicing, and aspiration. Voiced stops are associated with the presence of a voice bar on the spectrogram (Hogan & Rozsypal, 1980), the presence of periodic glottal pulsing in the waveform (Colantoni, Steele & Escudero, 2015, p. 187), continuation of voicing into closure (Charles-Luce & Dinnsen, 1987; Lousada, Jesus & Hall, 2010; Port & O'Dell, 1985), and a durationally shorter closure and release (Port & O'Dell, 1985; Port & Crawford, 1989). In contrast, voiceless stops are associated with shorter voicing into closure and a durationally longer closure and release (Ibid). These acoustic correlates have been used widely to measure the degree of final devoicing in

1. It should be noted that historically an orthographic distinction can be found between word-final and non-word-final stops (cf. Middle High German *tak* 'day' vs. *tages* 'of the day'). Middle High German was spoken between 1050–1350 CE.

languages such as German (Charles-Luce, 1985; Piroth & Janker, 2004; Port & Crawford, 1989; Port & O'Dell, 1985; Smith et al., 2009), Russian (Dmitrieva, Jongman & Sereno, 2010), Polish (Slowiaczek & Dinnsen, 1985), and Dutch (Simon, 2010). The phenomenon of "boundary-lengthening" should also be noted, as segments at boundary domains tend to be longer than domain-medial ones (e.g., Turk & Shattuck-Hufnagel, 2007). Therefore, word-final voiced stops in German are durationally longer not only because they are devoiced, but also because they are domain-final.

Vowels preceding voiced stops are typically durationally longer than vowels preceding voiceless stops (House & Fairbanks, 1953; Piroth & Janker, 2004; Port & O'Dell, 1985; Simon, 2010), which can be used as a perceptual cue for determining laryngeal status, particularly for word-final segments (Mack, 1982; Simon, 2010). This durational difference has been well documented in English (House & Fairbanks, 1953; Simon, 2010), German (Charles-Luce, 1985; Port & O'Dell, 1985; Piroth & Janker, 2004), and in other languages (Fischer-Jørgensen, 1964).

1.4 Research question

The goal of the present study is three-fold. First, it seeks to improve the visibility of German in research on L2 pronunciation instruction. Second, it seeks to examine the effects of implicit and explicit pronunciation instruction on the acquisition of German final devoicing. Third, by employing an acoustic analysis of learner speech before and after intervention, it aims to promote the use of acoustic work on classroom-based research on L2 German pronunciation instruction. The research question addressed is as follows.

RQ: As measured by four acoustic correlates of voicing (closure duration, release duration, preceding vowel duration, and duration of voicing into closure), is there a statistically significant difference between learners who received explicit pronunciation instruction (explicit condition) and learners who received implicit pronunciation instruction (implicit condition) in the production of underlyingly voiced German word-final and non-word-final stops?

H: Given the positive effects of explicit pronunciation instruction in studies of other languages (Lee, Jang & Plonsky, 2015), and the fact that the Skill Acquisition Theory (DeKeyser, 2020) points to the important role of declarative knowledge, the acoustic correlates of voicing will suggest that significantly greater word-final devoicing took place in speech produced by learners in the explicit condition after receiving explicit pronunciation instruction. The lack of an orthographic distinction between word-final and non-final German stops will interfere with the acquisition of this

phonological rule, but explicit instruction on final devoicing will help learners circumvent the interference.

2. Materials and methods

2.1 Participants

Two sections of third-semester German at a North American university took part in this study, divided into two learning conditions: explicit ($n = 16$) and implicit ($n = 13$).[2] To account for the instructor as a potential confound, the same instructor taught both sections. English was the L1 of 70 percent of learners in the explicit condition ($n = 11/16$) and 84 percent of learners in the implicit condition ($n = 11/13$). Mandarin Chinese ($n = 1/16$), Vietnamese ($n = 1/16$), Malay ($n = 1/16$), and Spanish ($n = 1/16$) made up the L1 for the remaining 30 percent of the explicit condition ($n = 4/16$), and Mandarin Chinese was the L1 of the remaining 16 percent of learners ($n = 2/13$) in the implicit condition. Based on their TOEFL and English proficiency admission scores, all learners who were non-L1 speakers of English were proficient speakers of the language. None of the learners had final devoicing in their L1.[3] As for the mean exposure to German, learners in the implicit condition had learned German for slightly longer ($M = 2$ years) than learners in the explicit condition ($M = 1.5$ years).

2.2 Procedure

Over a 16-week semester, learners completed a pre-test (during weeks 1–2), post-test (weeks 7–8), and delayed-post-test (weeks 15–16) to document their German pronunciation. During the first week of class, students selected a time to meet with the instructor outside of regular class hours to complete the production pre-test. The same procedure was followed for the post-test (during weeks 7–8) and delayed-post-test (during weeks 15–16). During their selected timeslot, students were taken to a sound-attenuated booth. Twenty-four slides on a computer were presented to learners, one by one (Appendix A). Learners were instructed to read

2. The choice to use sections of third-semester German was due to sample convenience.

3. Several languages with systematic final devoicing are reported in Section 1.3. While some evidence suggests that a degree of final devoicing (i.e., "variable final devoicing") is present for some lexical items in specific varieties of Spanish, especially those in contact with Catalan (Hualde & Eager, 2016), this was not the case for the L1 speaker of Argentinian Spanish in this study. Any influence of the learners' L1 was controlled for by including the L1 as a fixed factor in each statistical model and including each learner as a random intercept.

the stimuli on each slide aloud at a tempo comfortable to them.[4] Each production test was completed individually, outside of the classroom, to maximize the quality of the acoustic recordings. Their speech was recorded in Praat (Boersma & Weenink, 2019) using a Snowball ICE Microphone and was digitized at a sampling rate of 44,100 Hz, with a 16-bit resolution. The 24 slides were randomized for each learner using a random generator. Other than the order, the slides were identical across all three tests.[5]

2.3 Stimuli

The 24 slides consisted of target stimuli that were deliberately chosen to test final devoicing (Table 1). Twelve slides contained target stimuli in isolation (i.e., words by themselves) and twelve slides contained target stimuli in context. While data for other target sounds were also collected (e.g., fricatives and vowels), the analysis of these segments is beyond the scope of this study.[6] The 24 slides consisted of 188 words, containing 36 underlyingly voiced stops, of which 23 surface as voiceless because they occurred in word-final position, and 13 surface as voiced as they were not word-final. A list of the target stimuli containing the German stops in word-final and non-word-final position is provided in Table 1. There were nine words containing the underlyingly voiced /b/, of which six surface as voiceless (*lob*, *lieb*, *gib*, *gelb*, *ob*, *Brob*) and three surface as voiced (*loben*, *lieben*, *Brobe*). Six of these words alternated phonologically: *loben/lob*, *lieb/lieben*, *Brob/Brobe*. There were 19 words containing the underlyingly voiced /d/, of which 12 surface as voiceless (*Rad*, *Land*, *Kind*, *Hund*, *fand*, *Grund*, *Hemd*, *wird*, *sind*, *Deutschland*, *erminkeld*, *Pind*) and seven surface as voiced (*Räder*, *Länder*, *Kinder*, *Hunde*, *finden*, *Freunden*, *Pinde*).[7] Twelve alternated phonolog-

4. Because measurements for rate of speech were not normalized, one could argue that any durational difference between the two learning conditions is due to a difference in speech rate, that is, learners in one group talked faster than learners in the other. To account for this interference, each learner was included as a random factor in the statistical analysis.

5. To confirm the reliability of the four acoustic correlates of voicing, a 26-year-old female L1 speaker of German from Freiburg completed the same production task (Stratton, 2022). Statistical analyses found significant differences for all four parameters, suggesting, in line with previous research, that these are robust cues for measuring the presence or absence of German word-final devoicing (see Stratton, 2022, pp. 89–102).

6. For more information on the acquisition of other aspects of German L2 speech, see Stratton (2022, pp. 69–119).

7. The author acknowledges that, phonetically speaking, stops following voiced segments such as nasals (e.g., *Hund*) and liquids (e.g., *erminkeld*) may not be devoiced completely due to voicing assimilation. However, words containing nasals were included because they were frequent in the learning material.

ically: *Rad*/*Räder*, *Land*/*Länder*, *Kind*/*Kinder*, *Hund*/*Hunde*, *fand*/*finden*, *Pind*/*Pinde*. Eight words contained underlyingly voiced /g/, three surface as voiced (*Kriege*, *Teige*, *Piege*), and five surface as voiceless (*Krieg*, *Teig*, *mag*, *Tag*, *Pieg*), of which three pairs alternated: *Krieg*/*Kriege*, *Teig*/*Teige* and *Pieg*/*Piege*. Seven target words were pseudowords, included to test knowledge of the underlying phonological rule (*Brob*, *Brobe*, *Pieg*, *Piege*, *Pind*, *Pinde*, *erminkeld*), each containing one of the three underlyingly voiced word-final and non-word-final stops.[8] The target words (Table 1) were chosen because they contain medial and word-final stops, because they were appropriate for learners' proficiency level, and because they occurred frequently in the material for the course.[9]

Table 1. Target words containing underlyingly voiced stops

Stop	Word	Word-final	Gloss	IPA	Non-word-final	Gloss	IPA
/b/	Real	lob	'praise'	[loːp]	loben	'to praise'	[loːbn̩]
		lieb	'dear'	[liːp]	lieben	'to love'	[liːbn̩]
		gib	'give'	[giːp]			
		gelb	'yellow'	[gɛlp]			
		ob	'whether'	[ʔɔp]			
	Pseudo	Brob		[bʁoːp]	Brobe		[bʁoːbə]
/d/	Real	Rad	'wheel'	[ʁaːt]	Räder	'wheels'	[ʁɛːdɐ]
		Land	'land'	[lant]	Länder	'countries'	[lɛndɐ]
		Kind	'child'	[kɪnt]	Kinder	'children'	[kɪndɐ]
		Hund	'dog'	[hʊnt]	Hunde	'dogs'	[hʊndə]
		fand	'found'	[fant]	finden	'to find'	[fɪndn̩]
		Grund	'reason'	[gʁʊnt]	Freunden	'friends-DAT'	[frɔɣ̞ndn̩]
		Hemd	'shirt'	[hɛmt]			
		wird	'will'	[vɪʁ̞t]			
		sind	'are'	[zɪnt]			
		Deutschland	'Germany'	[dɔɣ̞t͡ʃlant]			
	Pseudo	Pind		[pɪnt]	Pinde		[pɪndə]
		erminkeld		[ɛa̯mɪŋkɛlt]			

8. While learners were not told explicitly that the 24 slides contained pseudowords, all pseudowords occurred in isolation (i.e., seven of the twelve slides containing only one word were pseudowords).

9. The uneven proportion of voiced-voiceless stops was due to course material. Since the training was conducted in the classroom, only stimuli relevant to course content were included. However, in retrospect, the author acknowledges the usefulness of matching, where possible, each word-final stop (e.g., *Tag* 'days') to a non-word-final counterpart (e.g., *Tage* 'days') even though some words (e.g., *ob* 'whether'), do not have such alternations (**obe*).

Table 1. *(continued)*

Stop	Word	Word-final	Gloss	IPA	Non-word-final	Gloss	IPA
/g/		Krieg	'war'	[kʁiːk]	Kriege	'wars'	[kʁiːgə]
		mag	'may'	[maːk]			
		Tag	'day'	[taːk]			
		Teig	'dough'	[taɪ̯k]	Teige	'dough- DAT'	[taɪ̯gə]
	Pseudo	Pieg		[piːk]	Piege		[piːgə]

2.4 Data analysis

For the acoustic analysis, recordings were annotated in Praat (Boersma & Weenink, 2019) using the TextGrid function (see Figure 1). To examine German final devoicing, following previous acoustic analyses (Dmitrieva, Jongman & Sereno, 2010), four temporal measures were taken: the duration of the closure in word-final and non-word-final stops, the duration of release in word-final and non-word-final stops, the duration of the vowel preceding word-final and non-word-final stops, and the duration of voicing into closure in word-final and non-word-final stops. For the analysis of vowel duration, to avoid interference of vowel height/quality only the seven contrastive word pairs that did not contain nasals were analyzed (*lob-loben*, *lieb-lieben*, *Brob-Brobe*, *Rad-Räder*, *Krieg-Kriege*, *Teig-Teige*, *Pieg-Piege*). Vowels were measured from the onset of the first formant on the spectrogram until the end of the second formant and the abrupt drop in waveform amplitude. Stop closure was taken from the end of the preceding vowel, nasal, or lateral, to the start of the release. Stop release was taken from the end of the closure until the end of visible noise on the spectrogram. Voicing into closure was measured from the end of the preceding vowel, lateral or nasal, until the end of periodic vibrations in the waveform. These measurements were taken for both word-final and non-word-final stops so that temporal comparisons could be made for both environments.

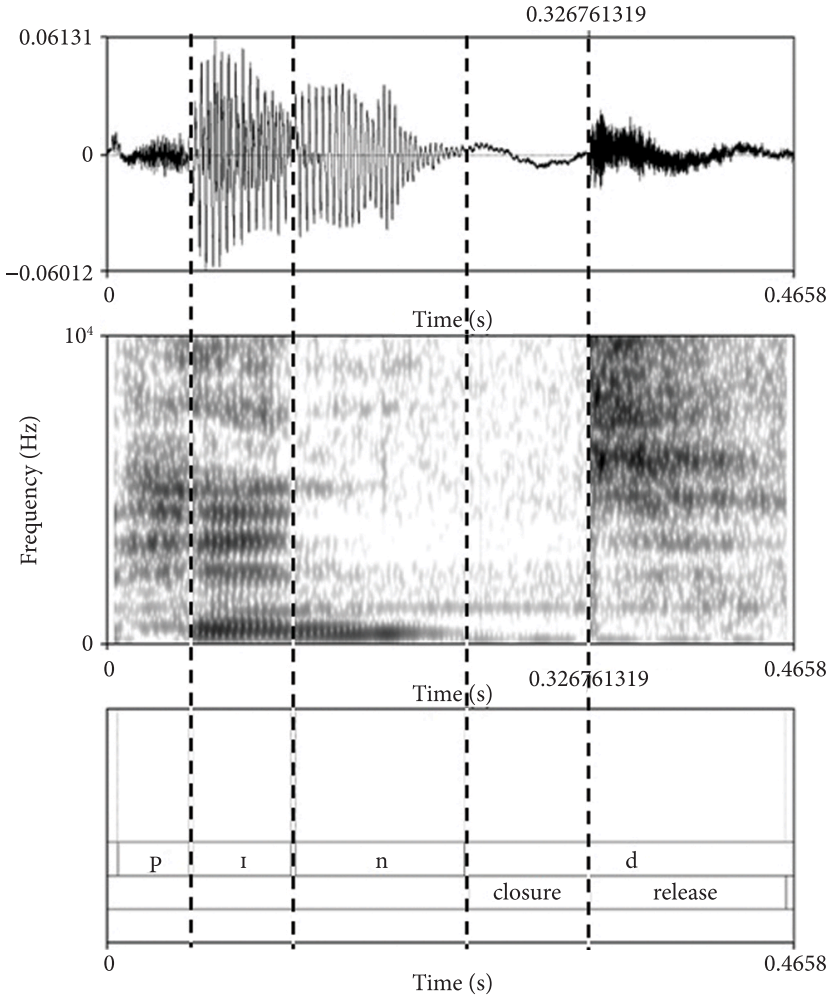

Figure 1. Sample of acoustic annotation in Praat[10]

10. /pɪnd/ is the phonemic, not allophonic representation.

2.5 Intervention

During weeks 3–6, both learning conditions completed six twenty-minute training sessions on pronunciation. The explicit condition completed six twenty-minute explicit training sessions on applied German phonetics and phonology, taught in English. These training sessions included instruction on manner and place of articulation, final devoicing, and other relevant aspects of pronunciation.[11] For pedagogical reasons, final devoicing was simplified to include only stop consonants. Learners in the explicit condition were given the following explanation: "when b, d, g appear at the end of a word in German, they are pronounced as p, t, k", with several illustrative examples.

In contrast, the implicit condition received six twenty-minute training sessions which drew attention to final devoicing implicitly, through the medium of German. For instance, in one of the sessions, the implicit condition played *Bingo* and had a table containing words ending in word-final (e.g., *Tag*) and non-word-final stops (e.g., *Tage*). The instructor read the word aloud (e.g., [taːk]) and, if present, learners crossed it off their sheet (e.g., *Tag*). If a learner had a full line/full house, after shouting *Bingo*, they read their words aloud to the instructor, one-by-one. Including words with word-final and non-word-final stops provided an opportunity to notice the difference in pronunciation between domain-final and non-domain-final stops. These learners also received implicit corrective feedback, such as recasts: a reformulation of the learner's utterance minus the error.[12] While both learning conditions received six twenty-minute training sessions, learners in the explicit condition spent less time on final devoicing than learners in the implicit conditions. Instruction on final devoicing in the explicit condition was limited to one twenty-minute training session and a follow-up review. A summary of the lessons can be found in Appendix B (explicit condition) and C (implicit condition).[13] With the exception of the pseudowords, learners in both learning conditions encountered all target words (Table 1) during their training sessions.

11. For information on the effects of the instruction on aspects of pronunciation beyond final devoicing, see Stratton (2022).

12. For example, if a learner said *ich habe einen Hund* 'I have a dog' [ʔɪç haːbə ʔaɪ̯nən hʊnd], the instructor would reply, *ah, du hast einen Hund* 'ah, you have a dog' [aː du hast ʔaɪ̯nən hʊnt].

13. A more detailed overview of the training sessions, as well as example activities, can be found on the journal site under supplementary materials. See also Stratton (2022).

3. Results

3.1 Final devoicing

The four acoustic correlates of voicing were analyzed (closure duration, release duration, preceding vowel duration, voicing into closure) in SPSS26 (IMB Corp., Armonk, NY). In the LMMs (Linear Mixed Models) which follow, an $\alpha = .05$ was used as the criterion for significance, and Cohen's d was used to measure effect size using the benchmarks of Plonsky and Oswald (2014): small ($d = .40$), medium ($d = .70$), large ($d = 1.0$). All models were run with the same random effects structure: a random intercept for LEARNER and a random intercept for WORD. The significance of the fixed factors and interactions were assessed using ANOVA tests, and all pairwise comparisons were carried out using Sidak correction. Reported confidence intervals (*CI*) are at 95% confidence.

3.1.1 *Closure duration*

According to previous research (Port & O'Dell, 1985; Port & Crawford, 1989; Smith et al., 2009), voiceless stops have a durationally longer closure than voiced stops. Because underlyingly voiced word-final stops are devoiced in German, word-final stops are expected to have durationally longer closures than their non-word-final counterparts. To examine CLOSURE DURATION, an LMM was run with four fixed factors: GROUP, POSITION, TIME, LEARNER-L1. All possible interactions were included in the model. The factor GROUP had two levels (explicit condition, implicit condition), POSITION had two levels (word-final, non-word-final), TIME had three levels (pre-test, post-test, delayed-post-test), and LEARNER-L1 had two levels (English L1, non-English L1).

The results demonstrated a significant effect of GROUP F (1, 2099) = 201.491, $p = .001$, POSITION F (1, 2099) = 6.024, $p = .001$, and TIME F (2, 2099) = 18.455, $p = .001$, but not LEARNER-L1 F (1, 2099) = .073, $p = .789$. The effect of GROUP was due to significantly longer closures in the explicit condition ($M = 50$ ms, $SD = 40$ ms) compared to the implicit condition ($M = 33$ ms, $SD = 27$ ms), with an effect size of $d = .50$ ($CI = .24, 1.24$). The effect of POSITION was due to significantly longer closures for word-final stops ($M = 51$ ms, $SD = 40$ ms) than for non-word-final stops ($M = 32$ ms, $SD = 27$ ms), with an effect size of $d = .56$ ($CI = .03, 1.08$). According to post-hoc pairwise comparisons, closures were significantly longer in the post-test ($M = 44$ ms, $SD = 40$ ms) and the delayed-post-test ($M = 44$ ms, $SD = 38$ ms) compared to the pre-test ($M = 40$ ms, $SD = 30$ ms).

As for interactions, there was a significant effect of GROUP × POSITION F (1, 2099) = 43.227, $p = .001$, GROUP × TIME F (2, 2099) = 25.824, $p = .001$, POSITION × TIME F (2, 2099) = 22.303, $p = .001$, and GROUP × POSITION × TIME

F (2, 2099) = 27.937, p = .001. The three-way interaction suggests the joint effects of POSITION × TIME on CLOSURE DURATION were not the same across the two learning conditions. As Figure 2 shows, the difference in closure duration for word-final and non-final stops in the explicit condition was not statistically significant in the pre-test, but this changed by the post and delayed-post-test in the direction of longer word-final closures. In contrast, as Figure 3 indicates, the implicit condition demonstrated a moderate effect of POSITION, but the size of this effect (ca. 10 ms) did not change from pre-test to post to delayed-post-test. Comparisons of marginal and conditional R^2 values indicate that 4.9 percent of variance in the model was due to the random effect LEARNER (SD = 8.4) and 11 percent was due to the random effect WORD (SD = 10.8). The variance for WORD (Estimate = 265, SD = 16) was higher than the variance for LEARNER (Estimate = 54, SD = 7): items *Brob* (intercept = 36) and *Brobe* (intercept = 24) diverged the most, suggesting some learners were responding differently to pseudowords. A follow-up model run using only the pseudowords demonstrated a three-way interaction of GROUP × POSITION × TIME F (2, 494) = 13.120, p = .001, with learners in the explicit condition producing significantly longer closures for word-final stops than non-word-final stops compared to the implicit condition.

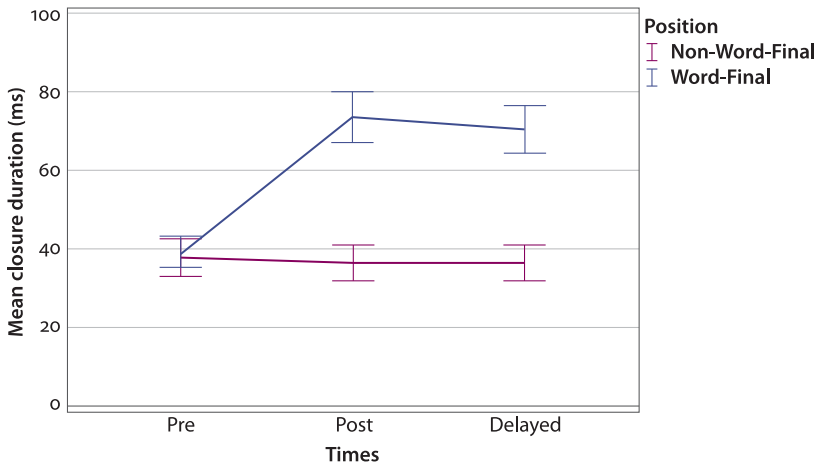

Figure 2. Closure duration in the explicit condition[14]

14. Whiskers in the figures represent the 95% confidence intervals.

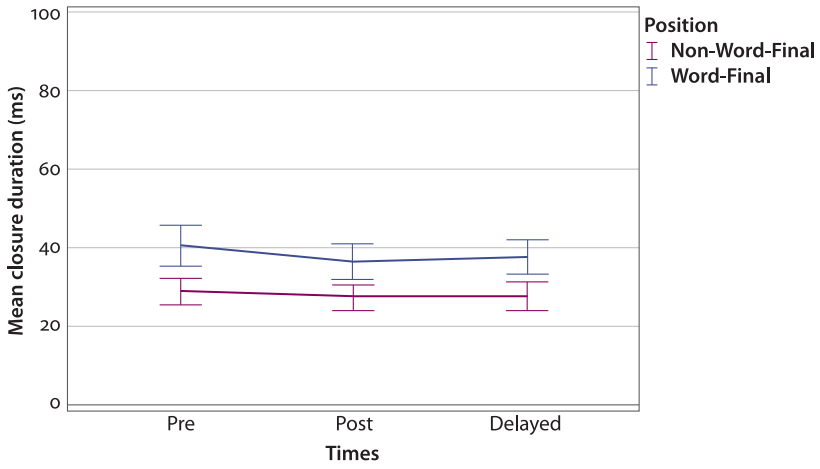

Figure 3. Closure duration in the implicit condition

To confirm the effect of GROUP, two follow-up models were run. First, an LMM was run on the data from the explicit condition, with POSITION and TIME as fixed factors, and a two-way interaction of POSITION × TIME. The model found a significant effect of POSITION F (1,098) = 12.732, p = .001, TIME F (1, 098) = 33.716, p = .001, and POSITION × TIME F (2, 098) = 36.977, p = .001. Second, the same model was run on the implicit condition, but no significant effects nor interactions were found. These results therefore confirm that the implicit condition's CLOSURE DURATION as a function of POSITION remained consistent over time. In contrast, in the explicit condition, there was an interaction between POSITION × TIME; specifically, there was an effect of POSITION in the post and delayed-post-test compared to the pre-test. Since shorter closure duration is a correlate of voicing and longer closure duration is a correlate of voicelessness, evidence from this parameter suggests that the intervention for learners in the explicit condition had a significant effect on the acquisition of German final devoicing, as they increased the CLOSURE DURATION of word-final stops by 34 ms from pre-test (M = 39 ms, SD = 27) to post-test (M = 73 ms, SD = 46 ms), with an effect size d = .90 (CI = .17, 1.63); longer duration in this context being more characteristic of a phonologically devoiced stop.

A parallel coordinate plot of the within-group mean closure duration is provided in Figure 4. Each line represents the average mean closure duration of underlyingly voiced word-final stops for each learner from pre-test to delayed-post-test. This plot shows that while each learner in the explicit condition increased their average closure duration, this was not true for all learners in the implicit condition. Three learners in the implicit condition increased their closure

duration from pre-test to post-test, but most learners decreased their closure duration, suggesting that some learners did not attend to nor acquire this durational cue to voicing.

3.1.2 *Release duration*

According to previous research (Port & O'Dell, 1985; Port & Crawford, 1989; Smith et al., 2009), voiceless stops are associated with durationally longer releases than voiced stops. Because underlyingly voiced word-final stops are devoiced in German, it is expected that the release duration of word-final stops will be longer than the release duration of non-word-final voiced stops. To examine RELEASE DURATION, an LMM was run with four fixed factors: GROUP, POSITION, TIME, and LEARNER-L1, and all possible interactions. The model found a significant effect of GROUP F (1, 2116) = 195.879, p = .001, POSITION F (1, 2116) = 34.297, p = .001, and TIME F (2, 2,116) = 39.042, p = .001, but not LEARNER-L1 F (1, 2,116) = 3.103, p = .078. The effect of GROUP was due to a significantly longer RELEASE DURATION in the explicit condition (M = 47 ms, SD = 57 ms) than in the implicit condition (M = 23 ms, SD = 31 ms), with an effect size of d = .5 (CI = .43, .62). The effect of POSITION was due to a significantly longer RELEASE DURATION for word-final stops (M = 19 ms, SD = 31 ms) than for non-word-final stops (M = 52 ms, SD = 56 ms), with an effect size of d = .73 (CI = .64, .82). According to post-hoc pairwise comparisons, releases were significantly longer in the post-test (M = 41 ms, SD = 54 ms) and the delayed-post-test (M = 41 ms, SD = 54 ms) compared to the pre-test (M = 24 ms, SD = 30 ms), while the post and delayed-post-test did not differ from each other.

As for interactions, there was a significant effect of GROUP × POSITION F (1, 2116) = 97.190, p = .001, GROUP × TIME F (2, 2116) = 40.895, p = .001, POSITION × TIME F (2, 2116) = 25.871, p = .001, and GROUP × POSITION × TIME F (2, 2166) = 43.455, p = .001. The three-way interaction suggests the joint effects of POSITION × TIME on RELEASE DURATION were different across the two learning conditions. As Figures 5 and 6 show, in the explicit condition, the difference in RELEASE DURATION for word-final and non-final stops in the pre-test was not statistically significant. However, by the post and delayed-post-test, the difference was statistically significant in the direction of longer word-final releases. In contrast, in the implicit condition, the RELEASE DURATION did not change significantly across the three tests, differing in word-final position by an average of only 2 ms. Based on this measure, the results suggest that learners in the implicit condition did not improve as a result of the pedagogical intervention. Comparisons of marginal and conditional R^2 values indicate that 4 percent of variance in the model was due to the random effect LEARNER (SD = 10.9) and 6 percent due to WORD (SD = 12.8). The variance for WORD (Estimate = 155,

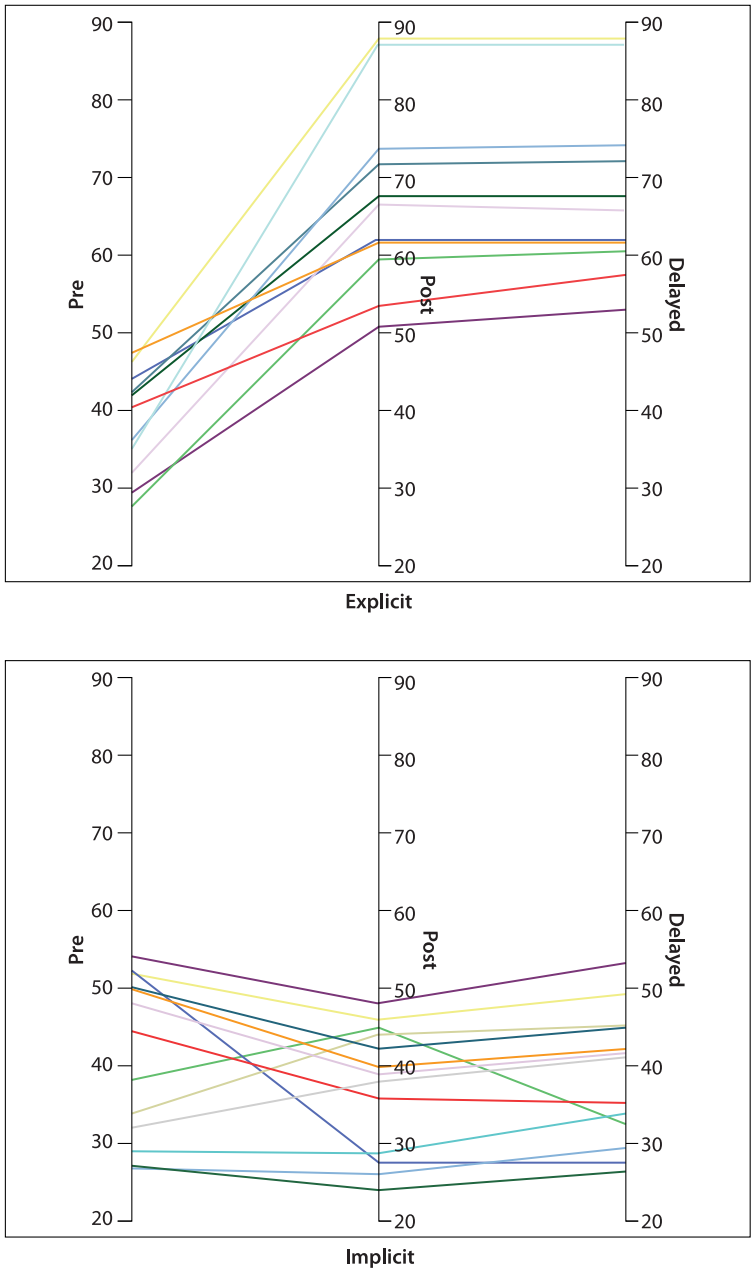

Figure 4. Parallel coordinate plot of individual mean differences in closure duration of underlyingly voiced word-final stops

SD = 12) was higher than the variance for LEARNER (Estimate = 71, *SD* = 11): items *Pind* (intercept = 25) and *Pieg* (intercept = 26) diverged the most. A follow-up model was run on the pseudowords using the same fixed and random effects. The model found a three-way interaction of GROUP × POSITION × TIME F (2, 154) = 3.337, p = .02; the explicit condition produced significantly longer releases over time for stops in word-final pseudowords than for stops in non-word-final counterparts compared to the implicit condition.

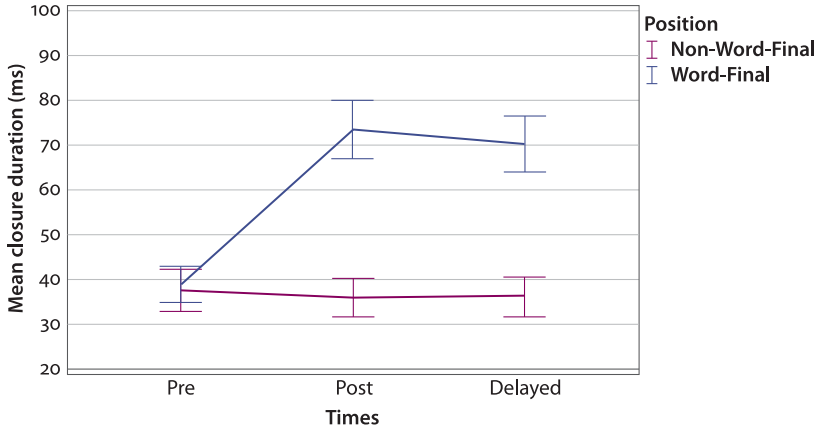

Figure 5. Release duration in the explicit condition

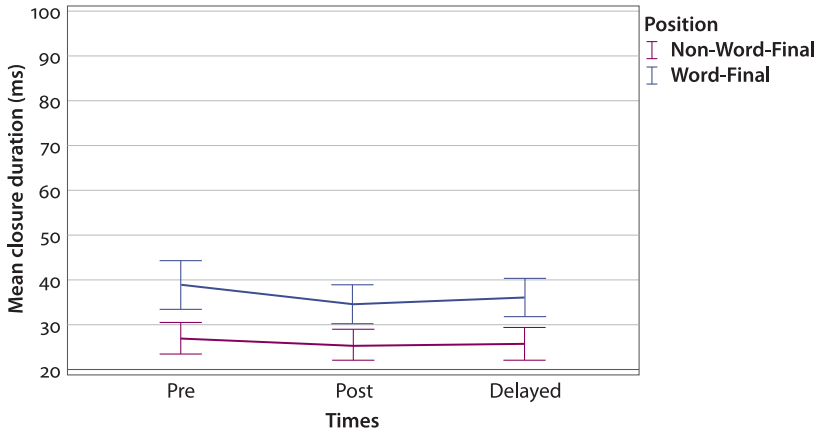

Figure 6. Release duration in the implicit condition

To confirm the effect of GROUP, two follow-up models were run, one LMM on the RELEASE DURATION in the explicit condition, and one LMM on the RELEASE DURATION in the implicit condition. In both models, POSITION

and TIME were run as fixed factors, and a two-way interaction of POSITION and TIME was included. As expected, the explicit model found a significant effect of POSITION F (1,115) = 292.232, p = .001, TIME F (1,115) = 54.920, p = .001, and POSITION × TIME F (1,115) = 46.453, p = .001. In comparison, while the implicit model found a significant effect of POSITION F (1,001) = 70.235, p = .001, no significant effect of TIME F (1,001) = .189, p = .828 was found, nor was there an interaction of POSITION × TIME F (1,001) = 1.667, p = .189. The significant difference in position in the implicit model is expected regardless of voicing due to lengthening effects. While this difference remains stable, in the implicit model it did not change significantly over time. The coordinate plot of the within-group differences in Figure 7 illustrates how the RELEASE DURATION increased in the explicit condition, but not in the implicit condition. These follow-up models therefore confirm that although the implicit model made a distinction in RELEASE DURATION between word-final and non-word-final stops, the RELEASE DURATION did not change significantly over time.

3.1.3 *Preceding vowel duration*

According to previous acoustic analyses (Port & O'Dell, 1985; Simon, 2010), vowels preceding voiceless stops are associated with durationally shorter vowels than preceding voiced stops. Because underlyingly voiced word-final stops are devoiced in German, their preceding vowels should be durationally shorter than their non-word-final counterparts. Like with the previous measures, to examine the duration of vowels preceding word-final and non-word-final stops in the two learning conditions, an LMM was run with the same four fixed factors and interactions. Results demonstrated a significant effect of GROUP F (1, 1225) = 57.367, p = .001, but the fixed factors POSITION F (1, 1225) = 2.950, p = .086, and TIME F (2, 1225) = .536, p = .872 were non-significant. The effect of GROUP was due to durationally shorter preceding vowels in the explicit condition (M = 158 ms, SD = 56 ms) compared to the implicit condition (M = 178 ms, SD = 57 ms), with an effect size of d = .35 (CI = –.38, 1.09). Comparisons of marginal and conditional R^2 values indicate that 7 percent of variance in the model was due to the random effect LEARNER (SD = 11) and 12 percent was due to WORD (SD = 13). The variance for WORD (Estimate = 769, SD = 27) was higher than for LEARNER (Estimate = 343, SD = 19). Items *Pieg* (intercept = –35, p < .005) and *lob* (intercept = –36, p < .004) diverged the most; negative intercept values in this context indicating durationally shorter preceding vowels. One learner in the implicit condition significantly diverged from the norm (Estimate = 38, p < .001), producing durationally shorter preceding vowels for four of the non-word-final stops than word-final counterparts, suggesting some confusion with attending to this durational cue.

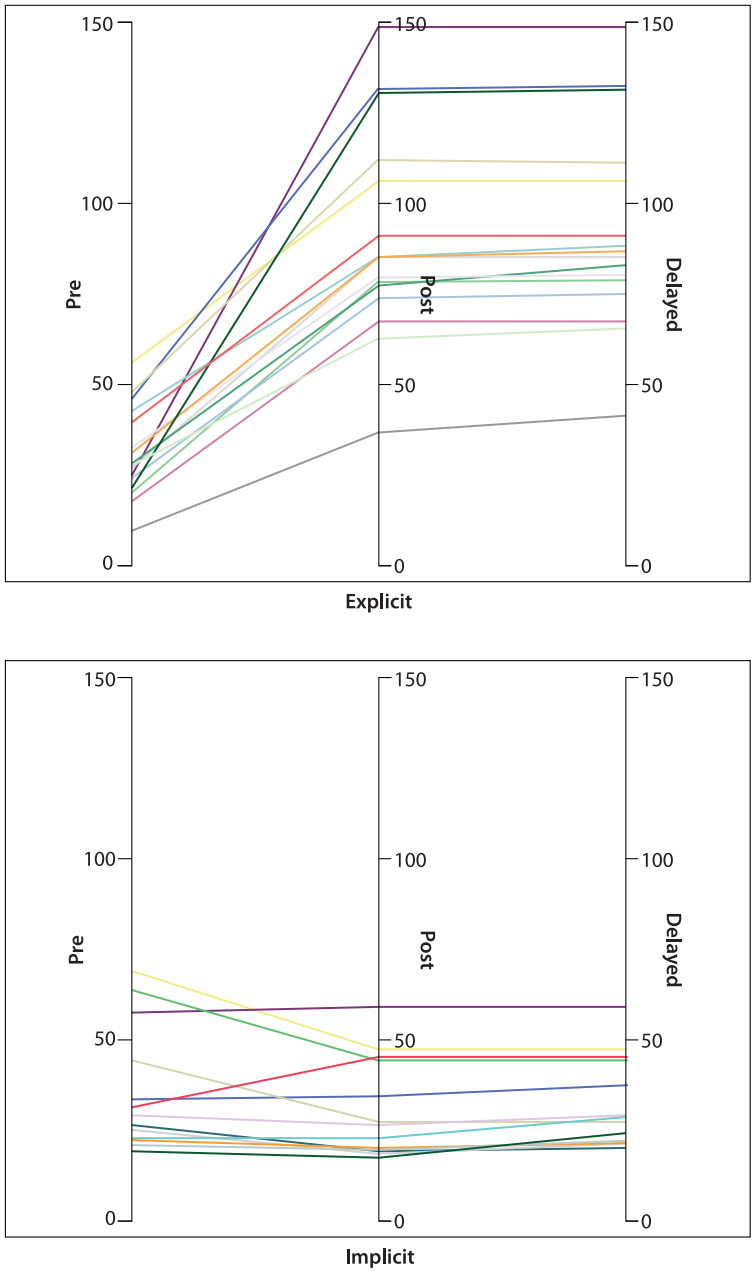

Figure 7. Parallel coordinate plot of individual mean differences in release duration of underlyingly voiced word-final stops

Although POSITION was not significant by itself, the interaction of POSITION × TIME was significant F (2, 1225) = 4.706, p = .009. As Figure 8 shows, in the explicit condition pre-test, vowels preceding word-final stops were durationally longer (M = 176 ms, SD = 61 ms) than vowels before non-word-final stops (M = 135 ms, SD = 46 ms), with an effect size of d = .76 (CI = .04, 1.48). However, by the post-test the duration of vowels preceding word-final stops decreased by an average of 22 ms (M = 154 ms, SD = 40 ms). In contrast, as Figure 9 illustrates, in the implicit condition, the duration of vowels preceding word-final stops remained stable from pre-test (M = 179 ms, SD = 57 ms) to delayed-post-test (M = 177 ms, SD = 57 ms). Since longer preceding vowel duration is more characteristic of phonologically voiced stops, the gradient movement over time that is more characteristic of voicing suggests that the instruction learners in the explicit condition received had a greater effect on final devoicing than the implicit instruction. While longer vowels are expected for word-final stops because of boundary lengthening, vowels preceding word-final stops in the explicit condition were significantly longer than in the implicit condition following the pedagogical intervention. Individual differences by learner across the two learning conditions is reported in Figure 10.

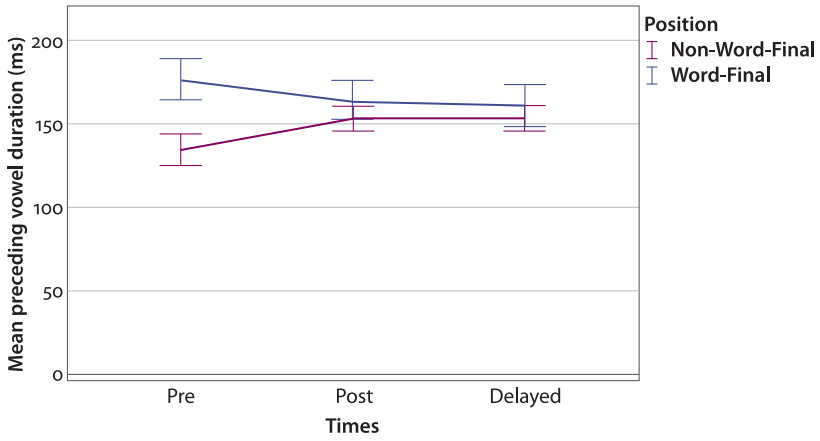

Figure 8. Preceding vowel duration of stops in the explicit condition

3.1.4 *Voicing into closure*

Longer voicing into closure is associated with voiced stops (Port & O'Dell, 1985; Charles-Luce & Dinnsen, 1987; Piroth & Janker, 2004). Vowels preceding underlyingly voiced word-final stops in German should therefore be shorter than underlyingly voiced non-word-final stops. To examine the VOICING INTO CLOSURE for word-final and non-word-final stops, an LMM was run, using

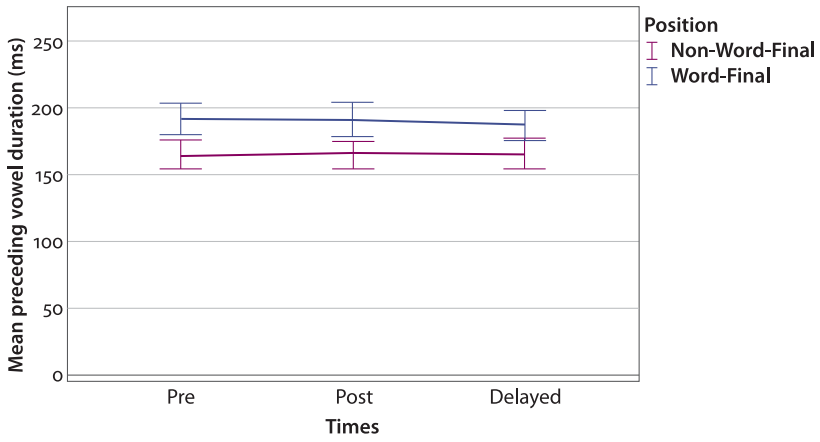

Figure 9. Preceding vowel duration of stops in the implicit condition

the same aforementioned factors and interactions. The model found a significant effect of POSITION F (1, 2122) = 195.645, p = .001, GROUP F (1, 2122) = 35.561, p = .001, and TIME F (2, 2122) = 71.734, p = .001. The effect of POSITION was due to significantly longer voicing into closure for non-word-final stops (M = 42 ms, SD = 8 ms) when compared with word-final stops (M = 37 ms, SD = 19 ms). The effect of GROUP was due to a significant difference in voicing into closure duration between the implicit (M = 39 ms, SD = 9 ms) and explicit condition (M = 28 ms, SD = 20 ms), with a medium effect size of d = .71 (CI = –.04, 1.46). Post-hoc pairwise comparisons indicate that the effect of TIME was due to significant differences in voicing into closure duration from pre-test (M = 42 ms, SD = 7 ms), to post-test (M = 28 ms, SD = 19 ms), to delayed-post-test (M = 23 ms, SD = 19 ms). As for interactions, TIME × GROUP was significant F (2, 2122) = 46.373, p = .001, as well as GROUP × POSITION F (1, 2122) = 92.975, p = .001 and TIME × POSITION F (2, 2122) = 41.609, p = .001. There was also a three-way interaction of TIME × GROUP × POSITION F (2, 2122) = 33.579, p = .001. Comparisons of marginal and conditional R^2 values indicate that 5 percent of variance in the model was due to the random effect LEARNER (SD = 5.5) and 6 percent was due to random effect WORD (SD = 6.4).

To explore the effect of GROUP further, two follow-up LMMs were run on the explicit and implicit condition. The model for the explicit condition found a significant effect of TIME F (2, 1,494) = 287.696, p = .001, POSITION F (1, 1494) = 921.452, p = .001, and TIME × POSITION F (2, 1,494) = 196.000, p = .001. As Figure 11 shows, in the pre-test, word-final stops (M = 41 ms, SD = 9 ms) and non-word-final stops (M = 43 ms, SD = 5 ms) in the explicit condition differed only minimally, suggesting little difference in attention to this acoustic cue of voicing.

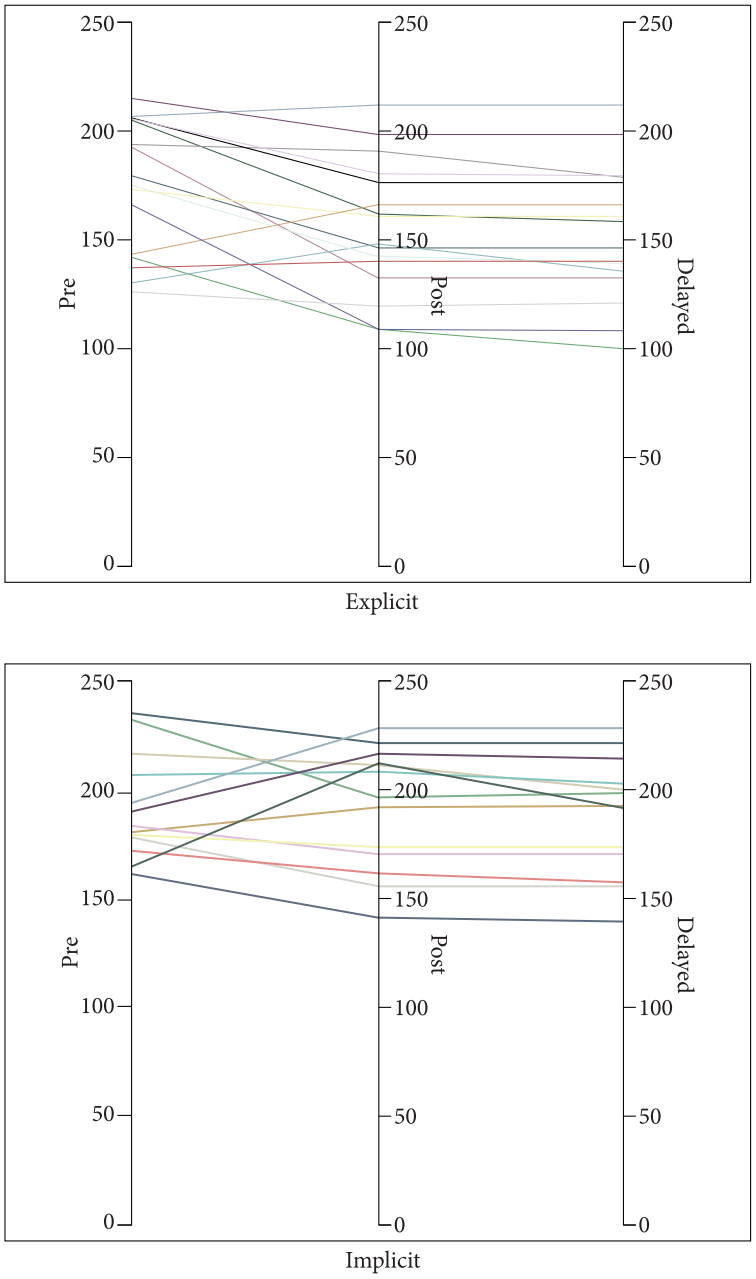

Figure 10. Parallel coordinate plot of individual mean differences in preceding vowel duration of underlyingly voiced word-final stops

However, by the post-test ($M = 11$ ms, $SD = 16$ ms), learners in the explicit condition reduced the duration of VOICING INTO CLOSURE for word-final stops by an average of 30 ms, with a large effect size of $d = 2.6$ ($CI = 1.63, 3.50$), suggesting that voicing of word-final stops was affected by the pedagogical intervention. In contrast, in the implicit model, only POSITION was significant $F(1, 711) = 17{,}158$, $p = .001$, suggesting there was a difference between word-final ($M = 39$ ms, $SD = 9$ ms) and non-word-final stops ($M = 42$ ms, $SD = 6$ ms) for VOICING INTO CLOSURE, but this did not change over time. A comparison of Figure 11 and Figure 12 illustrates the effect of the two learning conditions on VOICING INTO CLOSURE over time.

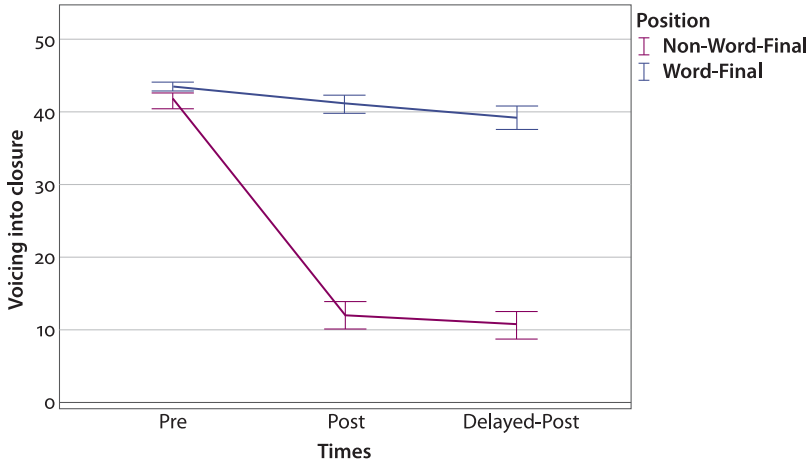

Figure 11. Duration of voicing into closure in the explicit condition

4. Discussion

This study examined the effects of implicit and explicit pronunciation instruction on the acquisition of final devoicing in an English-speaking L2 German classroom. Learners' word-final and non-word-final stops were analyzed using four temporal measures associated with consonant voicing. The composite analysis of all four measures indicated that over time, when compared to the implicit condition, learners in the explicit condition produced word-final stops that were significantly more characteristic of phonologically devoiced stops, suggesting that the effects of the explicit instruction on final devoicing were greater than the effects of implicit instruction. The results therefore corroborate previous findings on the positive effects of explicit pronunciation instruction. Because there are multiple cues to voicing, if one cue is not clearly realized, speakers can attend to other cues,

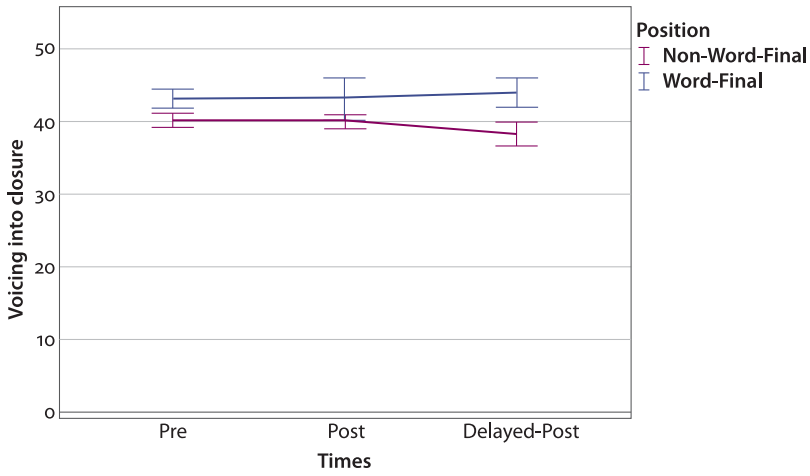

Figure 12. Duration of voicing into closure in the implicit condition

leading to efficient speech perception. The fact that the analyses suggest learners in the explicit condition were attending to all four cues indicates that learners were making efforts to signal durational differences through multiple parameters which are important for determining laryngeal status of word-final stops.

The findings from this study have direct implications for pronunciation instruction in the L2 German classroom. Although both learning conditions had six twenty-minute training sessions on pronunciation, in the explicit condition, training on final devoicing was limited to one twenty-minute training session and a follow-up review. In contrast, learners in the implicit condition carried out activities that exposed them to final devoicing in at least five of the six training sessions. Therefore, this study suggests that with as little as twenty minutes of explicit instruction on final devoicing, learners can significantly improve their production of word-final stops in German. Follow-up interviews with learners about the instruction they received also point to positive concomitants, such as a reduction in speaker and learner anxiety (Stratton, accepted). Some learners who received explicit instruction on final devoicing also suggested it helped circumvent cross-linguistic interference with the L1 (Stratton, accepted). These concomitant positive effects are in line with research that highlights additional benefits of explicit pronunciation instruction beyond the acquisition of L2 speech (Martin & Jackson, 2016).

There are several possible explanations why learners in the explicit condition outperformed learners in the implicit condition. First, according to the Skill Acquisition Theory (DeKeyser, 2020), the acquisition of a skill starts with declarative knowledge, which is knowledge of information that can be verbalized, such as

/b, d, g/ become [p, t, k] word-finally. In the context of second language acquisition, declarative knowledge involves metalinguistic and metacognitive awareness of the L2. Assuming second language acquisition is comparable to skill acquisition, one interpretation of the results is that the explicit instruction provided learners in the explicit condition the opportunity to acquire declarative knowledge of the underlying phonological rule, which through practice, resulted in significantly greater learning than in the implicit condition. Follow-up interviews after the experiment suggest that many university language learners expect declarative knowledge of underlying phonological rules as a function of higher education instruction and are satisfied when the instruction reflects their expectations (Stratton, accepted).

Although final devoicing concerns not just stops, but also affricates and fricatives, for pedagogical simplicity the instruction the learners received focused only on the former. Given the simplicity and systematicity of this rule, it is unsurprising that learners in the explicit condition outperformed learners in the implicit condition, as categorical rules are particularly amenable to explicit instruction (DeKeyser, 1995). Learners in the explicit condition also applied this phonological knowledge to pseudowords (e.g., *Brob*), suggesting that learners were applying the underlying phonological rule to words they could not have seen before. In contrast, although learners in the implicit condition carried out tasks that provided opportunities to acquire final devoicing implicitly, and they received implicit pronunciation feedback, no significant adjustments in the cues of voicing were observed over time for their production of word-final stops. One explanation for this difference is the degree of noticing. According to the Noticing Hypothesis (Schmidt, 1990), there are three levels of awareness: "perception" (level 1), "noticing" (level 2), and "understanding" (level 3). According to this hypothesis, while understanding involves noticing, noticing does not necessarily involve understanding, suggesting that more noticing leads to more learning. If the implicit tasks were insufficient for awareness at the level of understanding to take place, the lower level of awareness could explain why no significant improvements in the implicit condition were observed. Moreover, the fact that recasts did not have a significant effect on the acquisition of final devoicing is in line with research that found recasts to be the least noticed by learners (Lyster & Ranta, 1997).

Orthographic interference may have also impacted acquisition. Because L2 learners already have a fully developed L1, and they usually encounter the L2 orthographic system before their L2 phonological system has fully developed, learners likely map their L2 orthography onto their L1 phonology, or vice versa. Orthography has long been shown to interfere with the acquisition of L2 phonology (Hayes-Harb & Barrios, 2021; Hayes-Harb, Brown & Smith, 2018). When acquiring German final devoicing, using the L1 phonological system as a template

can be problematic if underlyingly voiced stops are not devoiced word-finally in the L1. In the present study, none of the learners had L1s in which word-final stops were systematically devoiced, meaning that if learners map German orthography to their L1, they will fail to devoice word-final stops. This appears to have been the case for learners in the implicit condition who, according to the reported acoustic analysis, did not improve significantly over time. While orthography could have affected the explicit condition too, the effects were clearer in the implicit condition, likely because learners in the explicit condition received instruction that may have helped them circumvent this interference. The lack of an overt orthographic distinction between German word-final and non-word-final stops may have also misled learners in the implicit condition to believe German permits voiced stops word-finally. Although learners in the implicit condition received input in which German stops were devoiced during the intervention (e.g., *Bild* vs. *Bilder*), the meaning-focused input, plus the opaque orthography for final stops, appears to have resulted in a failure to notice this rule, at least at the level of understanding.

One of the aims of the present study was to promote the use of acoustic work when evaluating the effects of different pedagogical interventions on German L2 pronunciation. The goal was not to diminish the role of impressionistic ratings, but to emphasize that failure to improve in comprehensibility does not necessarily mean a failure to improve in pronunciation. The choice to use impressionistic ratings or an acoustic analysis should be dependent on the research question, which means conclusions drawn from a study must be constrained by the type of analysis. A side-by-side comparison with impressionistic ratings could have been insightful, but this was beyond the scope of one study. If the goal of a study is to assess improvements in comprehensibility, impressionistic analyses are appropriate, but if the goal is to assess improvements in L2 pronunciation as a result of instruction, an acoustic analysis can be an appropriate choice. Since learners can improve their pronunciation but such changes are not always noticeable through impressionistic ratings, the present study calls for additional acoustic work on the effects of different classroom-based instructional practices on direct changes in German L2 pronunciation.

5. Conclusion

The present study highlighted the benefits of including explicit pronunciation instruction in the German L2 classroom. Using an acoustic analysis to measure changes in pronunciation, the present study showed that the effects of explicit pronunciation instruction on final devoicing, as measured by four correlates of voicing, were greater than the effects of implicit pronunciation instruction.

Learners who received explicit instruction made significantly greater gradient movements towards phonetic properties that are more characteristic of phonologically devoiced stops when producing underlyingly voiced word-final stops than learners in the implicit condition. This study therefore confirmed the benefits of explicit pronunciation instruction, it showed the applicability to German, particularly to a classroom-based setting, and it hopes to have provided an impetus for future acoustic work on German L2 pronunciation.

Funding

Open Access publication of this article was funded through a Transformative Agreement with University of British Columbia.

Acknowledgements

The author would like to thank Olga Dmitrieva for her tremendous support and guidance on this project. The author would also like to thank the three anonymous reviewers for their helpful comments.

References

Boersma, P., & Weenink, D. (2019). *Praat: Doing phonetics by computer* (Version 6.1.02) [Computer software]. http://www.fon.hum.uva.nl/praat

Charles-Luce, J. (1985). Word-final devoicing in German: effects of phonetic and sentential contexts. *Journal of Phonetics*, *13*(3), 309–324.

Charles-Luce, J., & Dinnsen, D.A. (1987). A reanalysis of Catalan devoicing. *Journal of Phonetics*, *15*(2), 187–190.

Colantoni, L., Steele, J., & Escudero, P. (2015). *Second language speech: Theory and practice*. Cambridge University Press.

DeKeyser, R. (1995). Learning second language grammar rules: An experiment with a miniature linguistic system. *Studies in Second Language Acquisition*, *17*(3), 379–410.

DeKeyser, R. (2020). Skill acquisition theory. In B. VanPatten, Keating, G.D., & Wulff, S. (Eds.), *Theories in second language acquisition: An introduction* (pp. 83–104) (3rd ed.). Routledge.

Derwing, T.M. & Munro, M. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. John Benjamins.

Dlaska, A., & Krekeler, C. (2013). The short-term effects of individual corrective feedback on L2 pronunciation. *System*, *41*(1), 25–37.

Dmitrieva, O., Jongman, A., & Sereno, J. (2010). Phonological neutralization by native and non-native speakers: The case of Russian final devoicing. *Journal of Phonetics*, *38*(3), 483–492.

Fischer-Jørgensen, E. (1964). Sound duration and place of articulation. *Zeitschrift für Phonetik und Allgemeine Sprachwissenschaft*, *17*(1–6), 175–207.

Gluszek, A., & Dovidio, J. (2010). The way they speak: A social psychological perspective on the stigma of nonnative accents in communication. *Personality and Social Psychology Review*, *14*(2), 214–237.

Hayes-Harb, R., & Barrios, S. (2021). The influence of orthography in second language phonological acquisition. *Language Teaching*, *54*(3), 297–326.

Hayes-Harb, R., Brown, K., & Smith, B. (2018). Orthographic input and the acquisition of German final devoicing by native speakers of English. *Language and Speech*, *61*(4), 547–564.

Hogan, J., & Rozsypal, A. (1980). Evaluation of vowel duration as a cue for the voicing distinction in the following word-final consonant. *Journal of the Acoustical Society of America*, *67*(5), 1764–1771.

House, A. S., & Fairbanks, G. (1953). The influence of consonantal environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of' America*, *25*(1), 105–113.

Huensch, A., & Thompson, A. (2017). Contextualizing attitudes toward pronunciation: Foreign language learners in the United States. *Foreign Language Annals*, *50*(2),410–432.

Hualde, J. & Eager, C. (2016). Final devoicing and deletion of /d/ in Castilian Spanish. *Studies in Hispanic and Lusophone Linguistics*, *9*(2), 329–353.

Lee, J., Jang, J., & Plonsky, L. (2015). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics*, *36*(3), 345–366.

Levis, J. (2016). Accent in second language research and teaching. *Journal of Second Language Pronunciation*, *2*(2), 153–159.

Levis, J. (2019). Pronunciation research in recent dissertations. *Journal of Second Language Pronunciation*, *5*(2), 181–194.

Lousada, M., Jesus, L. M., Hall, A. (2010). Temporal acoustic correlates of the voicing contrast in European Portuguese stops. *Journal of the International Phonetic Association*, *40*(3) 261–275.

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, *19*(1), 37–66.

Mack, M. (1982). Voicing-dependent vowel duration in English and French: Monolingual and bilingual production. *The Journal of the Acoustical Society of America*, *71*(1), 173–187.

Martin, I. (2018). Bridging the gap between L2 pronunciation research and teaching: Using iCPRS to improve novice German learners' pronunciation in distance and face-to-face classrooms. (Doctoral dissertation, Pennsylvania State University).

Martin, I., & Jackson, C. (2016). Pronunciation training facilitates the learning and retention of L2 grammatical structures. *Foreign Language Annals*, *49*(4), 658–676.

McCandless, P., & Winitz, H. (1986). Test of pronunciation following one year of comprehension instruction in college German. *The Modern Language Journal*, *70*(4), 355–362.

Munro, M. (2003). A primer on accent discrimination in the Canadian context. *TESL Canada Journal*, *20*(2), 38–51.

Munro, M., & Derwing, T. M. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System*, *34*(4), 520–531.

Olson, D., & Offerman, H. (2021). Maximizing the effect of visual feedback for pronunciation instruction: A comparative analysis of three approaches. *Journal of Second Language Pronunciation*, *7*(1), 89–115.

Peltekov, P. (2020). The effectiveness of implicit and explicit instruction on German L2 learners' pronunciation. *Die Unterrichtspraxis*, *53*(1), 1–22.

Piroth, H. & Janker, P. (2004). Speaker-dependent differences in voicing and devoicing of German obstruents. *Journal of Phonetics*, *32*(1), 81–109.

Plonsky, L. & Oswald, F. L. (2014). How big is 'big'? Interpreting effects sizes in L2 research. *Language Learning*, *64*(4), 878–891.

Port, R., & O'Dell, M. (1985). Neutralization of syllable-final voicing in German. *Journal of Phonetics*, *13*(4), 455–471.

Port, R., & Crawford, P. (1989). Incomplete neutralization and pragmatics in German. *Journal of Phonetics*, *17*(4), 257–282.

Roccamo, A. (2015). Teaching pronunciation in just ten minutes a day: A method for pronunciation instruction in first-semester German language classrooms. *Die Unterrichtspraxis*, *48*(1), 59–83.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*(2), 129–158.

Simon, E. (2010). Phonological transfer of voicing and devoicing rules: evidence from L1 Dutch and L2 English conversational speech. *Language Sciences*, *32*(1), 63–86.

Slowiaczek, L. M., & Dinnsen, D. (1985). On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics*, *13*, 325–341.

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.

Stratton, J. M. (accepted). Implicit and explicit pronunciation instruction: A study of learner preferences in higher education. *Unterrichtspraxis*.

Stratton, J. M. (2022). Explicit historical, phonetic, and phonological instruction in second language acquisition. (Doctoral dissertation, Purdue University).

Sturm, J., Miyamoto, M., & Suzuki, N. (2019). Pronunciation in the L2 French classroom: Student and teacher attitudes. *The French Review*, *92*(3), 60–78.

Thomson, R. I., & Derwing, T. M. (2015). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics*, *36*(3), 326–344.

Turk, A., & Shattuck-Hufnagel, S. (2007). Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics*, *35*(4), 445–472.

Appendix A. The 24 slides used in this study

1. Es wir**d** kühl
2. klicht
3. klücht
4. Lo**b** mich, warum? Ich habe Muetter geholfen
5. Ich fahre gern mit dem Ra**d**. Ah ja, Fahrrä**d**er sind schön. Sie sind echt cool finde ich, aber nicht, wenn ich über die Brücke fahren muss, weil das gefährlich sein kann
6. Uber
7. Pie**g**
8. Mütter lo**b**en oft ihre Kin**d**er unter Freun**d**en
9. klacht
10. Tei**g**e
11. Gi**b** mir den Tei**g**. Ich mache mit. Ich ma**g** chinesisches Essen nicht. Echt? Wollen wir andere Sachen kochen?
12. unter
13. Sie machen sich lustig über mich! Ist es Krie**g** oder was! Ich hasse Krie**g**e
14. Ich weiß nicht, ob ich dich nach Hause bringen kann
15. Pin**d**
16. erminkel**d**
17. Ich komme aus China, nicht aus Chemnitz, also nicht aus Deutschlan**d**! Ah schön. Übrigens, kennst du Brecht? Ja, er ist super freundlich
18. Guten Ta**g**! Möchten Sie etwas Brot? Ja bitte, das ist lie**b** von dir. Ich lie**b**e Brot! Aber gibt es einen Grun**d** dafür, dass das Brot gel**b** ist?
19. Ich komme aus einem komischen Lan**d**, wo es keine Dächer gibt, aber viele Län**d**er haben auch kein Dach
20. Pin**d**e
21. Bro**b**e
22. Bro**b**
23. Pie**g**e
24. Ich sah einen Hun**d** mit einem Hem**d**. Normalerweise finden Kin**d**er Hun**d**e schön, aber dieses Kin**d** fan**d** ihn komisch

Appendix B. Pronunciation instruction summary for explicit condition

Session	Theme	Content
Session 1	Introduction	Introduction to Phonetics: – Place of articulation (e.g., stopping the airflow using the lips creates a labial sound, such as b and p) – Tongue position Learners were instructed to produce sounds without moving their tongue (such as /l/ and /t/), drawing explicit awareness to tongue position. Review Learners produced English sounds and described their tongue position to a partner.
Session 2	Consonants vs. Vowels	Place and manner of articulation: – Obstruction difference between consonants and vowels (i.e., level of obstruction) – Stops Students were instructed to produce six stops (bilabial, alveolar, and velar stops), paying attention to voicing by placing their fingers on the larynx. – Final Devoicing Students were informed that German has a rule that states when [b], [d] and [g] are at the end of a word they are devoiced. Therefore, [b] is pronounced [p], [d] is pronounced [t], and [g] is pronounced [k]. Several examples were given. "For instance, how do you say 'dog' in German? *Hund* [t] not *Hund* [d]. However, if there are multiple dogs, you say *Hunde*. The <d> is pronounced [d] because it is not word-final." Example contrasts were written on the board (e.g., *Tag* vs. *Tage*). Students practiced pronouncing them (group activity)
Session 3	Fricatives	Review – Review Five-minute review of final devoicing, pronouncing target words on the board. – Introduction to IPA chart Fricatives Palatal fricative
Sessions 4–6	Review	– Review

Appendix C. Pronunciation instruction summary for implicit condition

Session	Theme	Content
Session 1	Listening-Speaking	Heads-up activity: – A PowerPoint was projected. Only learners could see it. – Every thirty seconds a picture appeared on the screen. Learners provided clues to the instructor about the word, without saying it (e.g., for *Hund* 'dog' → *ein Haustier mit vier Beinen. Es ist nicht eine Katze* 'a pet with four legs. It is not a cat'). – Target words were included (excluding pseudowords). – After ten minutes, learners replaced instructor.
Session 2	Listening-Speaking	Heads-up activity: – Same activity as in session 1, with additional stimuli. – Learners were given list containing target words (excluding pseudowords) and completed activity with a partner. They received implicit pronunciation feedback (i.e., recasts).
Session 3	Listening-Speaking	Bingo: – Learners were given a three-by-three table containing words (some distractors, some target containing word-final and non-word-final stops). – The instructor read aloud the word and learners circled it on their sheet.
Session 4	Listening-Speaking	Communicative activity: – Learners completed a communicative-based roleplay. They had a list of words to use, some affected by final devoicing.
Session 5	Listening-Speaking	Speed Dating: – Learners had a list of target words. They had a two-minute conversation with a partner about these words (e.g., *Hast du einen Hund*?), switching partners every two minutes.
Session 6	Listening-Speaking	Coffee Conversation: – Learners spent twenty minutes conversing with native speakers.

Address for correspondence

James M. Stratton
Department of English Language and Literatures
University of British Columbia
397-1873 East Mall
Vancouver, British Columbia V6T 1Z1
Canada

james.stratton@ubc.ca
https://orcid.org/0000-0002-7255-5545

Publication history

Date received: 9 September 2022
Date accepted: 12 June 2023
Published online: 20 July 2023