Journal of Phonetics 40 (2012) 129–140

Contents lists available at SciVerse ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/phonetics

Native English speakers learning German as a second language: Devoicing of final voiced stop targets

Bruce L. Smith *, Elizabeth A. Peterson

Department of Communication Sciences and Disorders, University of Utah, 390 S. 1530 E., Room 1216, Salt Lake City, UT 84112, USA

ARTICLE INFO

Article history:
Received 18 May 2011
Received in revised form
12 September 2011
Accepted 15 September 2011
Available online 11 October 2011

ABSTRACT

In contrast to German and other languages that devoice underlying word-final, voiced obstruent targets, English maintains a surface contrast between voiced and voiceless obstruents. The present study investigated the issue of what occurs when native speakers of American English, in an early stage of learning German as a second language (L2), produce word-final voiced and voiceless stop consonant targets in German versus English. The fact that the underlying voicing contrast in German is reflected orthographically (e.g., ‘‘Tod’’ versus ‘‘tot’’) might make it more difficult for native speakers of English to learn to devoice German word-final, voiced targets. The findings of this investigation indicate that many of the 12 native English learners of L2 German who were studied showed at least a tendency toward devoicing voiced targets in German relative to their productions of orthographically similar words in English (e.g., ‘‘toad’’ and ‘‘tote’’). Considerable inter-subject variability was observed, but in general, their partial devoicing in German (relative to their English productions) occurred as a result of producing somewhat shorter vowels before voiced consonant targets and/or less contrast between voiceless versus voiced consonant closure duration. Subjects who produced more characteristically ‘‘voiced’’ consonants when speaking English (i.e., with longer preceding vowel duration, etc.) also tended to devoice German final stops to a lesser extent.

1. Introduction

Because of their historical relationship, English and German share a number of common linguistic characteristics in terms of phonological patterns, lexical similarities, etc. There are, of course, many differences between the two languages as well. For instance, in terms of phonological patterns, one well-documented characteristic is that English has a word-final obstruent voicing contrast, whereas German (and a number of other languages such as Russian, Polish, Dutch, etc.) manifests a pattern of word-final obstruent devoicing (e.g., Charles-Luce, 1985; Fourakis & Iverson, 1984; Keating, Linker & Huffman, 1983). In general, this results in an underlying voicing contrast such as in ‘‘Rad/Rat, Bad/bat, Tod/tot’’ being neutralized in German speakers' productions (i.e., [rɑt/rɑt], [bɑt/bɑt], [tot/tot]).

However, it is also frequently observed that a variety of factors may be associated with less than total neutralization of the underlying voicing contrast in Standard German and other languages that have a pattern of phonological devoicing (e.g., Charles-Luce, 1985; Dmitrieva, Jongman, & Sereno, 2010; Piroth & Janker, 2004; Slowiaczek & Dinnsen, 1985; Smith, Hayes-Harb, Bruss, & Harker, 2009). This is referred to as ‘‘incomplete neutralization,’’ reflecting the fact that underlying voiced targets that are nominally devoiced when produced may show small differences in one or more acoustic properties when compared to underlying voiceless targets, e.g., preceding vowel duration, consonant closure duration, etc. Such differences may also sometimes be perceptually distinguishable for at least some native listeners (Port & Crawford, 1989; Port & O'Dell, 1985; Slowiaczek & Szymanska, 1989). For example, Port and O'Dell found that native German listeners were able to perceptually discriminate voiced and voiceless stops about 60% of the time, a rate that was significantly better than chance. Slowiaczek and Dinnsen (1985) also noted that despite the phonological pattern of word-final devoicing in Polish, the underlying voicing contrast was phonetically maintained in the productions of most of their native Polish subjects to some degree (but cf. Jassem & Richter, 1989).

Warner, Jongman, Sereno, and Kemps (2004) concluded that incomplete voicing neutralization in Dutch probably occurs due to various phonetic and non-phonetic reasons (e.g., orthography, word frequency, lexical factors, and individual speakers' familiarity with languages such as English that have a final voicing contrast), which may exert an influence on a speaker's L1 productions. Incomplete neutralization may also occur for certain acoustic parameters, for certain phonemes, in certain contexts or tasks, or for certain speakers, but not others (Charles-Luce, 1985;

* Corresponding author. Tel.: +1 801 585 6783.
E-mail address: bruce.smith@hsc.utah.edu (B.L. Smith).

0095-4470/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.wocn.2011.09.004

Fourakis & Iverson, 1984). Piroth and Janker (2004) found, for instance, that complete neutralization in German occurred for vowel duration and voicing during stop closure in syllable- and word-final position for all six speakers in their study, while two speakers produced a consonant duration contrast for stops but not fricatives. Such findings indicate that the circumstances associated with partial or total neutralization of the underlying final obstruent voicing contrast in German, Dutch, Polish, and other languages are very complex and may be affected by factors related to the dialect spoken, individual speaker differences, word frequency, phonotactic relationships, orthography, degree of familiarity with languages such as English that have a surface contrast, etc. (Charles-Luce, 1985; Kleber, John, & Harrington, 2010; Piroth & Janker, 2004; Port & Crawford, 1989; Warner, Good, Jongman, & Sereno, 2006; Warner et al., 2004).

In contrast to German, Russian, Polish, etc., English has an underlying voiced/voiceless contrast among word-final obstruents (e.g., "bag/back, had/hat, cab/cap") that is typically maintained in surface productions (i.e., [bæɡ/bæk], [hæd/hæt], [kæb/kæp]). However, although English generally manifests a word-final voicing contrast, many native speakers show a tendency for producing final voiced obstruent targets that are not fully voiced—particularly in terms of a reduction in the amount of glottal pulsing during the closure/constriction interval of final stops and fricatives. Depending on factors such as dialect and gender, many native adult speakers of English commonly produce glottal pulsing during only 60–70% of stop closure intervals (Holmes, 1996; Smith, 1979; Smith et al., 2009; Stevens, Blumstein, Glicksman, Burton, & Kurowski, 1992), and even less in final fricatives (Fullana & Mora, 2009; Smith, 1997). This tendency toward devoicing is likely a consequence of a decrease in the transglottal pressure differential that must be maintained for voicing to continue during the entire closure/constriction interval (Ernestus, Lahey, Verhees, & Baayen, 2006; Koenig, Fuchs, & Lucero, 2011; Kuzla, Cho, & Ernestus, 2007; Lousada, Jesus, & Hall, 2010; Rodgers & Fuchs, 2010; Westbury & Keating, 1986). Nevertheless, the amount of voicing present during voiced stop closure intervals, along with other acoustic cues such as preceding vowel length, etc. is routinely sufficient for listeners to recognize a voicing contrast (Fullana & Mora, 2009; Hillenbrand, Ingrisano, Smith, & Flege, 1984; Smith, 1997).

Smith et al. (2009) investigated what occurs when German speakers learning English as a second language attempt to develop a voicing contrast in their production of English word-final obstruents. One of the primary issues concerned whether L1 speakers of German could effectively acquire a word-final voicing contrast in English, and at the same time maintain a word-final devoicing pattern when speaking German. Acoustic analyses indicated that the native German speakers showed more evidence of a word-final voicing distinction when producing English words than orthographically similar German words (e.g., German "Tod/tot" versus English "toad/tote"). However, the acoustic cues that the native German talkers produced when speaking English were less robust than the cues produced by native English talkers. It was also found that not all of the acoustic cues that were investigated showed the same effects in the native German speakers' English productions. For example, they typically achieved a relatively native-like pattern in English of vowel lengthening preceding word-final voiced stops, whereas consonant closure duration differences and glottal pulsing during voiced stop targets were essentially "intermediate" between the acoustic patterns of native speakers of English and the patterns characteristic of their native German. There was also considerable individual variation among the native German subjects speaking English—something typical of L1 production patterns as well (Smith, 2000, 2002). However, inter-subject variability is likely even more common in the L2 acquisition process due to varying degrees of experience, proficiency, motivation, degree of concern for pronunciation accuracy, etc. (Flege, Frieda, & Nozawa, 1997; Moyer, 1999; Piske, MacKay, & Flege, 2001).

Although various studies have examined L1 speakers of languages that have a pattern of word-final obstruent devoicing (e.g., German, Russian, Dutch, etc.) and what occurs as they learn English, the opposite situation has received considerably less attention (e.g., Dmitrieva et al., 2010; Eckman, 1977; Moyer, 1999; Young-Scholten, 2002, 2004). That is, what occurs as native English talkers learn German or another language that does not have a surface, word-final voicing contrast is not well understood. It may seem that learning to neutralize an L1 contrast in the process of learning an L2 that has no such contrast might be easier to accomplish (e.g., for speakers of English learning German), compared to acquiring a final voicing contrast in learning an L2 when one does not exist in L1 (e.g., when native German speakers learn English). Young-Scholten (2004) indicated, for example, that "when an English speaker is acquiring a language such as German, where a final obstruent contrast present in the native language is neutralized in the target language, acquisition should be straightforward" (p. 69). However, because the influence of orthography might work against learning to devoice L2 words that are underlyingly voiced, the situation may not be so straightforward. Complex relationships exist between L1 and L2 orthographic and phonological knowledge and performance as a function age, experience, etc. (Abu-Rabia, 2001; Duyck, 2005; Lemhoefer & Dijkstra, 2004; Van Wijnendaele & Brysbaert, 2002; Wade-Woolley, 1999; Wang, Koda, & Perfettin, 2003). As a result, German word pairs such as "Rad/Rat" and "Bad/bat" reflect a voiced versus voiceless contrast in their spelling that could negatively impact an L2 learner's pronunciation of such pairs and the need to devoice underlying voiced targets. For instance, Young-Scholten (2002, 2004) studied three American teenagers on a year-long exchange program in Germany in terms of their progress toward acquiring the pattern of final consonant devoicing (based on perceptual judgments), and she obtained mixed results. The students' performance ranged from no evidence of final devoicing by one of them to limited evidence of occurrence in the speech of the other two. Young-Scholten attributed their differential progress, in part, to variations they each had in the amount of their exposure to written versus spoken input.

Similarly, Dmitrieva et al. (2010) examined the issue of word-final devoicing from several perspectives, including a group of native English speakers learning Russian as a second language. Utilizing acoustic analyses, Dmitrieva and her colleagues observed that the native speakers of American English learning Russian produced significant duration differences between members of Russian minimal pairs that were associated with underlying word-final voicing contrasts, whereas little or no difference would be expected among native speakers of Russian. For example, they found that the native speakers of English learning Russian produced differences in vowel duration before underlying voiced versus voiceless targets, differences in final stop closure or fricative duration, and voicing duration during the final stop closure/fricative interval. A significant correlation was also observed between the amount of experience L2 learners of Russian had speaking the language and the amount of devoicing that occurred in their speech.

Given the limited amount of information about what occurs when L1 speakers of a language that has a word-final voicing contrast attempt to learn a language such as German or Russian (e.g., Dmitrieva et al., 2010; Young-Scholten, 2002, 2004), the purpose of the present investigation was to examine a group of native speakers of American English at a relatively early stage in learning German as a second language to determine the extent to which they may have begun to develop the pattern of final stop consonant devoicing for underlying word-final, voiced targets. As Moyer (1999) notes, this is a pattern one might expect to be relatively easy to acquire. However, acquisition of word-final devoicing is likely to be dependent on various factors related to the amount of explicit training early learners receive, the nature of the speaking model(s) they may be exposed to (i.e., native or non-native speakers of German), individual learners' motivation, inherent phonological abilities, effects of orthography, etc. Rather than making perceptual-impressionistic judgments as to whether non-native speakers did or did not devoice underlying voiced targets (i.e., essentially a yes/no judgment), the approach for examining this issue in the present study was to measure several acoustic parameters known to be associated with the production of the voicing contrast for word-final consonants, including duration of the preceding vowel, duration of final consonant closure, duration of glottal pulsing during the consonant closure interval, and final release burst duration. The typical patterns associated with these production characteristics in English are that vowels tend to be longer before voiced versus voiceless stop targets, voiceless stop closure durations tend to be longer than voiced closure durations, there is a greater amount of glottal pulsing during voiced stop closure intervals than during voiceless stops, and burst duration tends to be longer upon release of voiceless versus voiced stop consonant closure (Lehiste, 1970). The primary issue of interest was whether beginning L2 German speakers whose native language was American English would neutralize the German underlying voiced/voiceless contrast for targets such as "Leid/leit" (or at least manifest an intermediate tendency toward devoicing) or whether they would maintain a voicing contrast comparable to their L1 pattern.

2. Method

2.1. Subjects

Twelve native American English-speaking undergraduate students (seven female, five male) enrolled in their second semester of German at the University of Utah served as subjects. They were within two to three weeks of the end of their second semester of studying German when they were recorded. They averaged approximately 24 years of age (range: 19–36 years). In providing self ratings of their knowledge of German on a scale of "basic/conversational/fluent," ten of the twelve subjects listed their level as "basic," one as "basic to conversational," and one as "conversational."

Instructors of the first and second semester courses the 12 subjects had previously taken or were enrolled in at the time of the study consisted of two graduate teaching assistants. They were both native speakers of American English, but they were very proficient in German as a result of having undergraduate degrees in German, and they were also both currently enrolled in a Master's degree graduate program in German. One of them had previously participated in a one-semester study abroad experience, and she had also visited family in Germany annually for more than 20 years because her spouse is German. The other instructor had lived in Germany for almost two years during his early twenties and had spoken German regularly during that time. He also reported that his parents had spoken German in their home when he was a child, but he never spoke German or studied it until high school.

Both instructors indicated that they did not provide any explicit instruction in the first and second semester classes they taught regarding the devoicing pattern of German, nor did they specifically correct students' pronunciation "errors" related to devoicing (since it did not tend to interfere with comprehension). They both reported that the textbook they used did not contain any information regarding this pronunciation pattern, but an online workbook associated with their classes did have some relevant activities. They both were of the opinion that the primary exposure their students most likely had to the German devoicing pattern occurred indirectly, i.e., based on the extent they paid attention to their instructors pronouncing various words that involved word-final devoicing of underlying voiced targets. Both instructors indicated that they themselves were aware of the Standard German pattern of devoicing and incorporated it into their own speech; however, no recordings or acoustic measurements were made of the instructors' speech to determine how their devoicing productions may have been manifested in the context of teaching their courses. (Neither instructor indicated that they had any awareness of any regional or other dialectal variant of German that might have influenced their devoicing pattern in a different manner from Standard German, e.g., North German vowel differences, with a short [a] in "Rad, Bad" and a long [aː] in "Rat, bat.") Given all these factors, it is not possible to know with any certainty whether, or to what extent, the 12 individual subjects may (or may not) have been aware of the general devoicing pattern of Standard German. Given the differences in linguistic abilities and learning styles that these different students undoubtedly had, it is possible that some may have observed their instructors' devoicing (to a greater or lesser extent), whereas other students may not have noticed the pattern at all.

2.2. Procedures

2.2.1. Stimuli

The 12 subjects completed the same tasks as those employed in a previous study of native speakers of German learning English as a second language (Smith et al., 2009). Specifically, they produced 10 repetitions of 26 different German words that were embedded in non-final (five repetitions) and final (five repetitions) position of a carrier sentence (with the restriction that within any repetition, the same target word did not occur twice, nor was a minimal pair cognate included). The German carrier phrase in which the different words were embedded was: "Ich habe schon oft ____ gesagt; jetzt sage ich nur noch ___" (I have said ___ often; now I say ___). Each subject thus produced a total of 130 German sentences with two different target words per sentence. To provide some variety within the word list and to help minimize awareness of the stimuli of interest, approximately 25% of the target words in the carrier sentences on each page of the list were "filler" words. Several additional filler sentences with stimulus words that were not measured were also included to avoid possible "beginning-/end-of-list" prosodic effects. After completing the recordings in German and having a brief rest, each of the subjects also read a list of English carrier sentences, producing 20 different target words with five repetitions of each randomly interspersed in both non-final and final position of the phrase (with the same restriction of no repetition of a word or its minimal pair cognate within the same sentence). The English carrier phrase in which the different words were embedded was: "I like to say ____ some of the time, but today I say ___". Each subject produced a total of 100 English sentences with two target words per sentence. All subjects read the German sentence list first to reduce the likelihood they would recognize or be affected by any similarities among some of the German and English minimal pairs (see below).

The target words of primary interest included six minimal pairs that are orthographically similar in German and English;

they were the same targets words employed in the study by Smith et al. (2009). The specific target words included the six German pairs: "Bad/bat, Leid/leit, Log/lock, Rad/Rat, seid/seit, Tod/tot," which are quite similar to the English pairs: "bad/bat, lied/light, log/lock, rod/rot, side/sight, toad/tote." (Table 1 presents a list of the German and English target words, including glosses for the German words.) All of these German and English words ended in voiceless or voiced stop consonant targets, as did the additional stimulus words in both languages. Given the cross-linguistic nature of these orthographic pairings, it was not possible to control for factors such as lexical frequency within or across the two languages, morphological complexity within or across the two languages (e.g., one morpheme versus two morphemes—such as, respectively, German "Leid" versus English "lied"), whether some words with underlying voiced targets have surface alternations (e.g., "Tod/Tode" [tot]/[todə]) whereas others do not (e.g., "seid"), place of consonant articulation, vowel quality or length, or other possible parameters that have been suggested as potential factors to consider when analyzing minimal pair contrasts pertaining to the issue of neutralization (Piroth & Janker, 2004; Slowiaczek & Dinnsen, 1985; Warner et al., 2004).

2.2.2. Recording

All subjects were recorded reading the German and English sentence lists in a Whisper Room SE 2000 sound booth, speaking into a Samson QV head-mounted microphone positioned at a distance of approximately 5 cm to the side of their lips. The recordings were made on a Marantz PMD660 Solid State recorder at a sampling rate of 44.1 kHz. The files were then downloaded to an Apple iMac computer, where they were analyzed using waveforms and spectrograms displayed utilizing Praat (5.2.1.5) software (Boersma & Weenink, 2006).

Table 1
German and English word lists organized by orthographic similarities. The six word pairs of primary interest in each language are indicated with asterisks. (Some words are capitalized and others are not because nouns are routinely capitalized in German orthography.)

German	English
Bad * (n., bath)	bad *
bat * (v., asked)	bat *
	bought
Bett (n., bed)	bed
	bet
Bild (n., picture)	build
	built
Dieb (n., thief)	deep
gab (v., gave)	cob
kapp (v., imper., cut the top off)	cop
	cub
	cup
Leid * (n., pain/sorrow)	lied *
leit * (v., imper., lead)	light *
lock * (v., imper., entice)	lock *
Log * (n. – loanword, log)	log *
Rad * (n., wheel)	rod *
Rat * (n., advice)	rot *
Ried (n., marsh)	reed
riet (v., advised/counseled)	
rott (v., rot, imperative)	
Sack (n., sack)	sack
sag (v., imper., say)	sag
satt (adj., full)	sod
	sought
seid * (v., are)	side *
seit * (prep., since)	sight *
sog (v., sucked)	
Sud (n., stock)	sued
	suit
Tag (n., day)	talk
Tod * (n., death)	toad *
tot * (adj., dead)	tote *

2.2.3. Acoustic analyses

The temporal measurements made of the 12 speakers' productions included: (1) vowel duration preceding voiced and voiceless final stop consonant targets, (2) final stop consonant closure duration, (3) duration of glottal pulsing during final stop consonants, and (4) final stop consonant release burst duration. Segmentation of the consonants and vowels was based on commonly utilized acoustic characteristics associated with substantial changes in waveform shape and/or amplitude, the occurrence of consonant release bursts, and other salient acoustic events related to frequency and amplitude characteristics of the waveform and spectrographic displays (Smith, Hillenbrand, & Ingrisano, 1986).[1]

It is important to note that the results reported below include both absolute duration values and relative measures. The reason for the relative measures is to facilitate comparisons among different vowels and consonants, different measurement intervals across languages, etc.[2] In addition to individual segmental measures, a "composite voicing" measure was calculated by averaging the relative durational patterns of the four individual acoustic measures. The rationale for calculating this composite measure was to determine how the various individual acoustic measures that were made might collectively function in relation to the contrast between voiced and voiceless target consonants for the subjects' English versus German productions. Because there was no preconceived reason to assume that any one of these parameters should be weighted more than any of the others, the composite values that were calculated averaged in all four parameters equally (i.e., relative vowel lengthening before voiced stops, voiceless relative to voiced stop closure, percentage of glottal pulsing during voiced stop closure, and voiceless relative to voiced burst duration were averaged together). Although subjects produced the target words in both non-final and final position, the data reported are for sentence-final target words only. This was done to limit potential assimilatory effects that might have occurred in non-final word position due to the subsequent phonetic context.

Intra- and inter-judge measurement reliability for the various acoustic parameters were assessed by: (1) having the same person re-measure data from one subject's English and German productions a second time after a period of approximately two months, and (2) by having a second person measure the English and German data for that same speaker. Intra-judge reliability measurements were obtained by comparing absolute differences between means across the multiple repetitions of the various target words. Intra-judge reliability was found to be approximately the same for each of the acoustic parameter for the English, word-final voiced versus voiceless target words. Differences averaged less than 1 ms (< .5%) for vowel duration

[1] A figure showing the specific segmentation patterns can be found in Smith et al. (2009).

[2] The relative values were calculated by comparing the durations for two measures of interest, e.g., vowel duration preceding voiced versus voiceless consonants, voiceless versus voiced consonant closure duration, etc. If the two (absolute) duration values were equal, a relative value of 1.00 was assigned to represent their relationship. When the duration was longer for one of the measures than the other (e.g., vowels before voiced stops were longer than those before voiceless stops), results that were in the "expected" direction were assigned a relative value greater than 1.00 (i.e., the longer vowel duration was divided by the shorter vowel duration). When durations were longer for segments that were in the unexpected direction, relative values were less than 1.00—indicating the opposite of the typical English pattern; that is, the shorter duration was divided by the longer duration.

preceding both voiced and voiceless targets, 2–4 ms (4–8%) for voiced versus voiceless consonant closure duration, 3 ms (8%) for glottal pulsing during voiced stop closure, and 10–11 ms (7–8%) for both voiced and voiceless burst duration. Intra-judge reliability for the German target words was also about 1–2 ms (.5–1%) for vowel duration before voiced versus voiceless stop targets, .4 ms (< 1%) for closure duration of voiced versus voiceless stop targets, 6 ms (12%) for glottal pulsing during voiced target consonants, and 15–28 ms (9–17%) for burst duration of voiced versus voiceless stop targets. There did not appear to be any systematic differences in intra-judge reliability for voiced versus voiceless target words.

Inter-judge reliability was also obtained by comparing absolute differences between means across the multiple repetitions of the various target words, as measured by the two investigators. Inter-judge reliability averaged 1–8 ms (.4–5%) for vowel durations preceding voiced versus voiceless stop targets in both English and German, 1–8 ms (2–15%) for voiced versus voiceless closure duration in both languages, 5–8 ms for glottal pulsing during consonant closure of voiced targets in both languages, and 14–24 ms (11–16%) for burst duration in English versus .2–14 ms (< .5–14%) in German. As with intra-judge reliability, there did not appear to be any systematic differences in inter-judge reliability associated with voiced versus voiceless target words. The reliability findings suggest that differences reported below concerning vowel duration, consonant duration, or glottal pulsing during closure that are within 3–10 ms of one another (or 15–25 ms for release burst duration) are in the typically observed realm of 5–10% measurement and/or speaker variability and are likely to be of little consequence in making any comparisons.

3. Results

A one-way repeated measures ANOVA was conducted to compare the effect of the language spoken (English versus German) on the various temporal measures that were made for the 12 subjects. It was determined that there was a significant effect of the language spoken ($F=70.163$, df = 13, 143, $p < .0001$). Post-hoc comparisons were calculated using Tukey–Kramer Multiple Comparisons Test and are reported below for the individual measures of interest.

3.1. *Vowel lengthening before voiced versus voiceless stop targets*

As shown by the left-most pair of solid black and striped bars in Fig. 1, the 12 subjects' vowels preceding voiced stop targets averaged 216 ms in English (black bar) versus 177 ms in the orthographically similar words in German (striped bar): Tukey–Kramer $q=5.949$, $p < .01$. The second pair of solid black and striped bars from the left shows that before voiceless consonants, vowels were also somewhat longer in English (153 ms) versus the orthographically similar words in German (141 ms); however, this difference was not statistically significant (Tukey–Kramer $q=1.694$, $p > .05$). On average, the ratio of vowel durations preceding voiced versus voiceless stop targets was 1.42 when speaking English (i.e., longer vowels before voiced relative to voiceless stop targets) and 1.26 when speaking German. Related to these values, vowel duration preceding voiced stops was longer than vowel duration preceding voiceless stops in both their English (i.e., the black bars in the first two pairs: Tukey–Kramer $q=9.551$, $p < .001$) and their German productions (i.e., the striped bars in the first two pairs: Tukey–Kramer $q=5.297$, $p < .05$).

Fig. 2 indicates that there was considerable variability among the 12 individual subjects in the extent to which they showed the pattern of relative vowel lengthening before voiced stop targets in both English (black diamonds) and German (open squares). A value of 1.00 represents equal durations for the two measures being compared, with values greater than 1.00 indicating longer vowels preceding voiced versus voiceless stop consonant targets, and values less than 1.00 indicating the opposite of the typical English pattern. Each of the 12 subjects showed the vowel lengthening pattern to a greater or lesser extent in both languages, i.e., *all* squares and diamonds are above 1.00. Based on a rank order of their English productions (black diamonds), relative lengthening values ranged from subject 9 (who had a ratio of 1.68

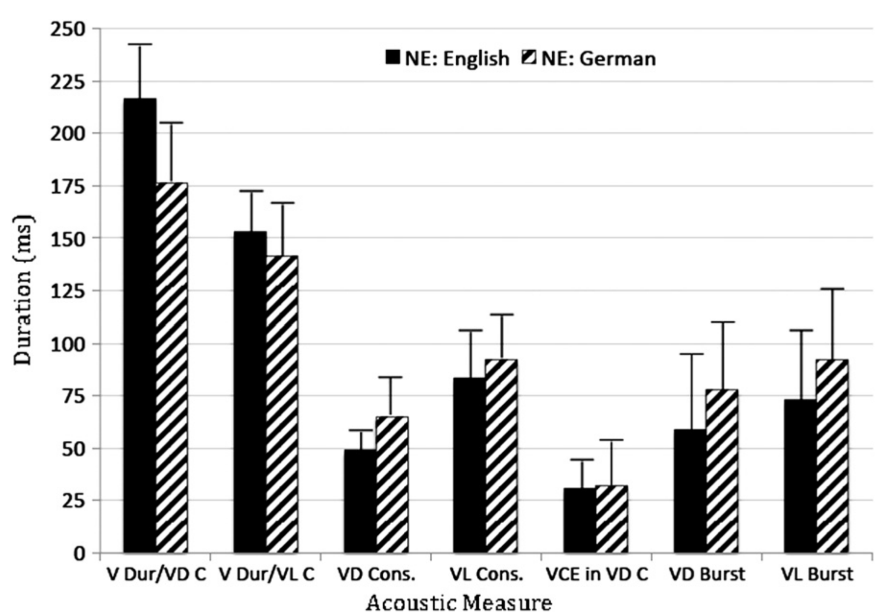

Fig. 1. Group means and standard deviations of various acoustic measurements for 12 native English (NE) subjects speaking English (black bars) versus German (striped bars). From left to right, the abbreviations represent: Vowel duration before voiced consonants, Vowel duration before voiceless consonants, Voiced consonant closure duration, Voiceless consonant closure duration, Duration of voicing during voiced consonant targets, Duration of voiced consonant release bursts, Duration of voiceless consonant release bursts.

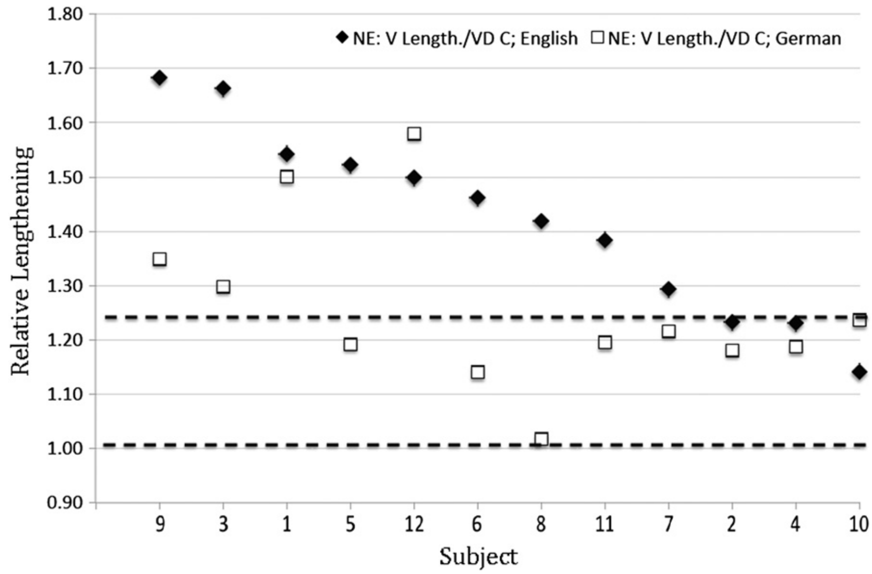

Fig. 2. Rank ordering for 12 native English (NE) speakers' relative measures of vowel lengthening preceding voiced (versus voiceless) stop consonants in English (black diamonds) and paired values for the same speakers' German (white squares) productions. A value of 1.00 indicates that vowel durations were equal before voiced and voiceless stops. Values greater than 1.00 indicate that vowels were longer before voiced than voiceless stops, whereas values less than 1.00 indicate vowels were longer before voiceless than voiced stops. The two horizontal dashed lines show the range of values among 13 native German speakers' relative vowel lengthening before voiced stop targets in German (from Smith et al., 2009).

for vowels before voiced versus voiceless stops in English and a ratio 1.35 for the vowels of German words with final voiced versus voiceless stop targets) to subject 10 with a ratio of 1.14 for vowels before voiced versus voiceless stops in English and 1.24 for vowels before underlying voiced versus voiceless targets in German.[3] For six of the native English-speaking subjects (9, 3, 5, 6, 8, 11), relative vowel lengthening was considerably greater in English versus German (on the order of .20–.35). The other six subjects (1, 12, 7, 2, 4, 10) showed very little difference (.10 or less) between their measures of relative vowel lengthening in English and German. Disregarding the *specific* amount of difference between the two languages, 10 of the 12 subjects showed somewhat greater relative vowel lengthening in English than in German, whereas two (12 and 10) showed slightly greater relative vowel lengthening in German.

Because similar measures as those for these 12 native, American English-speaking talkers were previously obtained for native German subjects (Smith et al., 2009), the *range* of values for those 13 native German talkers (speaking German) is shown in Fig. 2 by the two horizontal dashed lines. These data thus show how the 12 native English-speaking subjects[4] of the present study compare with the native German speakers of the earlier study. As can be seen, the native German subjects evidenced a range of approximately 1.00–1.23 for relative vowel lengthening when speaking German. When comparing the 12 native English talkers speaking German (open squares) with the range shown by the 13 native German subjects, it can be seen that there was considerable overlap between the two groups; that is, eight of the 12 native English talkers (5, 6, 8, 11, 7, 2, 4, and 10) were within the range shown by the native Germans. (It is worth mentioning that values above 1.00 represent a form of incomplete neutralization by the native German talkers.) It should also be noted that three of the native English talkers who fell within the German range when speaking German (2, 4, and 10) showed very similar amounts of relative vowel lengthening when speaking English; therefore, at least in this regard, they did not seem to do anything especially different when speaking German versus English. In contrast, subjects 1 and 12 also showed very little difference for this measure in English and German, but their values were much more English-like.

3.2. Voiceless versus voiced stop consonant closure duration

The third pair of bars from the left in Fig. 1 indicates that the subjects' closure duration for voiced stop targets averaged 49 ms in English (black bar) versus 65 ms in German (striped bar): Tukey–Kramer $q=2.473$, $p>.05$. For voiceless stop targets (fourth pair of bars), closure duration averaged 84 ms in English versus 92 ms in German (Tukey–Kramer $q=1.343$, $p>.05$). On average, the ratio of their voiceless versus voiced closure intervals was 1.71 in English versus 1.42 in the orthographically similar words in German, with their voiceless closure durations being significantly longer than their voiced closure durations in English (i.e., the black bars in the third and fourth pairs: Tukey–Kramer $q=5.209$, $p<.05$), but not in German (i.e., the striped bars in the third and fourth pairs: Tukey–Kramer $q=4.079$, $p>.05$).

As shown in Fig. 3, all 12 native English speakers produced longer voiceless than voiced stop closure intervals when speaking English (i.e., all black diamonds are above 1.00), ranging from 2.37 for subject 9 to 1.05 for subject 8. Thus, subject 9's voiceless closure intervals were more than twice as long as her voiced closure intervals in English productions, whereas subject 8's voiceless closure intervals were only slightly longer than his voiced closure intervals. It is also worth observing that in addition to having the greatest relative amount of voiceless versus voiced consonant closure "lengthening," subject 9 also had the greatest relative vowel lengthening before voiced consonants in Fig. 2. For seven of the 12 subjects (9, 7, 11, 10, 6, 3, 5), relative lengthening

[3] The relative values in this and other figures are rank ordered from greatest to least (left to right) on the basis of the subjects' English production patterns. Each subject's value for English (black diamond) is paired with his/her value in German (open square) in each figure. Subject numbers are consistent from one figure to the next in order to allow comparisons of subjects across different figures.

[4] None of the 12 native American English-speaking subjects in the present study was in the control group of 13 native English-speaking subjects in the earlier study by Smith et al. (2009).

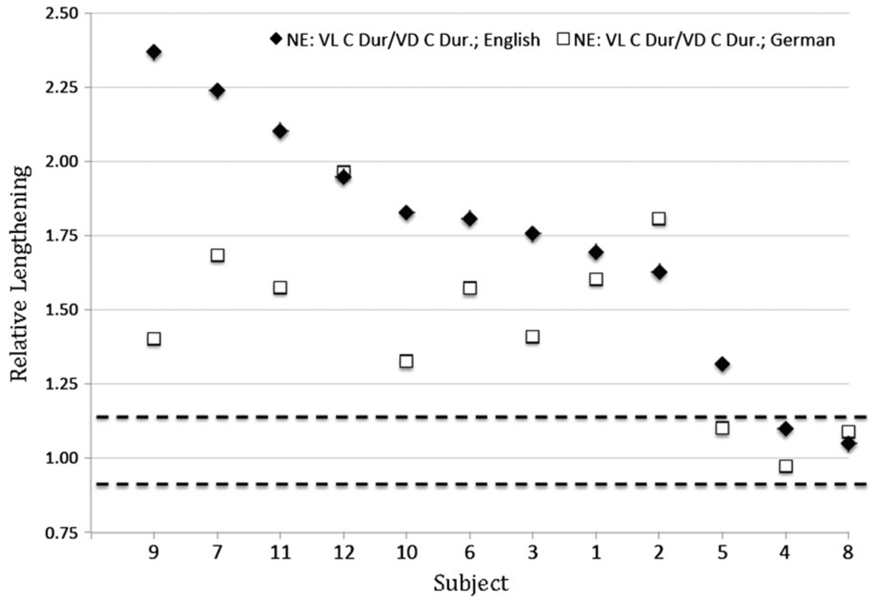

Fig. 3. Rank ordering for 12 native English (NE) speakers' voiceless stop consonant "lengthening" relative to voiced stops in English (black diamonds) and paired values for the same speakers' German (white squares) productions. A value of 1.00 indicates that voiced and voiceless stop closure durations were equal. Values greater than 1.00 indicate that voiceless closure was longer for voiceless than voiced stops, whereas values less than 1.00 indicate voiceless stop closure was shorter than voiced stop closure. The two horizontal dashed lines show the range of values among 13 native German speakers' relative voiceless stop lengthening compared to underlying voiced stop targets in German (from Smith et al., 2009).

of voiceless stop closure intervals compared to voiced stop closure intervals was .20 or greater in English versus German (i.e., black diamonds are higher than open squares). The range of values shown by the native German talkers speaking German (Smith et al., 2009) is again indicated by the two horizontal dashed lines. As can be seen, when speaking German only three of the native English talkers in the present study (5, 4, 8) fell within the range of the native German talkers for this measure. However, their relative values were also quite similar when they spoke English, i.e., despite the fact that their German productions were within the range shown by the native Germans for this parameter, they were not very different from their English productions. As with vowel duration, subjects 1 and 12 also showed very similar productions in the two languages for this measure (as did subject 2), but their productions were much more English-like.

3.3. *Glottal pulsing during stop closure*

As seen in the fifth pair of bars from the left-hand side of Fig. 1, there was not a significant difference between the *absolute* amount of glottal pulsing the subjects produced during voiced stop targets in English (31 ms) versus German (32 ms): Tukey–Kramer $q=.151$, $p>.05$. Thus, they were producing essentially identical amounts of glottal pulsing during closure in both languages. However, due to differences in absolute closure duration for voiced consonant targets mentioned above (English: 49 ms; German: 65 ms), the *relative* amount of glottal pulsing during voiced stop targets differed between the two languages. For this relative comparison, the measure of interest is the proportion of the consonant closure interval that showed glottal pulsing in English versus German voiced stop targets, with 1.00 representing "fully voiced," i.e., glottal pulsing during the entire closure interval. As can be seen in Fig. 4, their relative amount of glottal pulsing when speaking English (i.e., black diamonds) ranged from .87 of the closure interval for subject 1 to .13 for subject 11, with a group average of .61. When speaking German (i.e., open squares), the proportion of glottal pulsing was .48 of the closure interval on average—ranging from .90 for subject 4 to .07 for subject 11 (who also showed the smallest relative amount of voicing in English). The group differences between English and German constituted a significant decrease in the proportion of voicing during closure for L1 versus L2 productions (paired $t=2.318$; df$=11$; $p<.05$), but as noted, this was a consequence of differences in consonant closure duration rather than a difference in the absolute amount of voicing in German versus English. Although the native English talkers' proportion of glottal pulsing was less during voiced consonant targets in German (.48) than English (.61), their relative amount of glottal pulsing in German still exceeded the native German speakers' average ($M=.25$; range$=.15$–.49, as shown by the two horizontal dashed lines) when they were speaking German (Smith et al., 2009). It is noteworthy that among the 12 native English talkers, subject 8 was the only one who showed a substantial decrease in the *absolute* amount of glottal pulsing during closure from English (50 ms) to German (26 ms). The other seven subjects who had decreases in absolute voicing from English to German showed differences on the order of 1–8 ms (which is potentially within the range of measurement error). Of the four subjects who had more absolute glottal pulsing in German rather than English, only subject 9 showed a rather sizeable contrast (English$=16$ ms; German$=47$ ms).

3.4. *Final stop release burst duration*

As shown by the two pairs of bars on the right-hand side of Fig. 1, final burst duration was longer when the native English talkers were speaking German versus English for both voiced (78 versus 59 ms) and voiceless (92 ms versus 72 ms) stop targets, but neither of these differences was statistically significant (Tukey–Kramer $q=2.862$ and 2.950, respectively; $p>.05$ in both cases). In relative terms (i.e., voiceless/voiced burst durations), the subjects' bursts were longer when speaking English (1.37) versus German (1.23), but this difference was not significant (paired $t=1.32$; df$=11$; $p>.05$).

3.5. *Composite measure of voicing*

Because each of the individual acoustic parameters examined above is associated to some extent with the phonological voiced

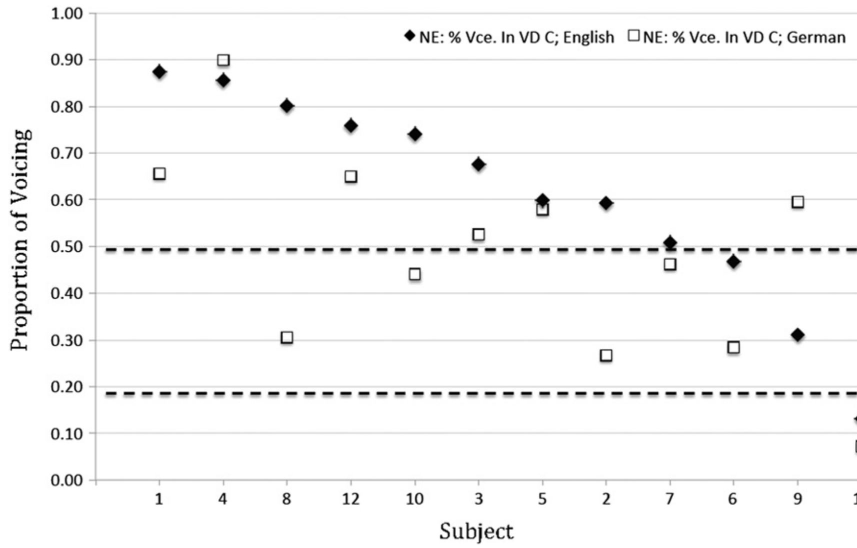

Fig. 4. Rank ordering for 12 native English (NE) speakers' relative measures of glottal pulsing during voiced stops in English (black diamonds) and paired values for the same speakers' voiced stop targets in German (white squares). A value of 1.00 indicates "fully voiced," whereas .00 represents no glottal pulsing during closure. The two horizontal dashed lines show the range of values for 13 native German speakers' proportion of glottal pulsing during voiced stop targets in German (from Smith et al., 2009).

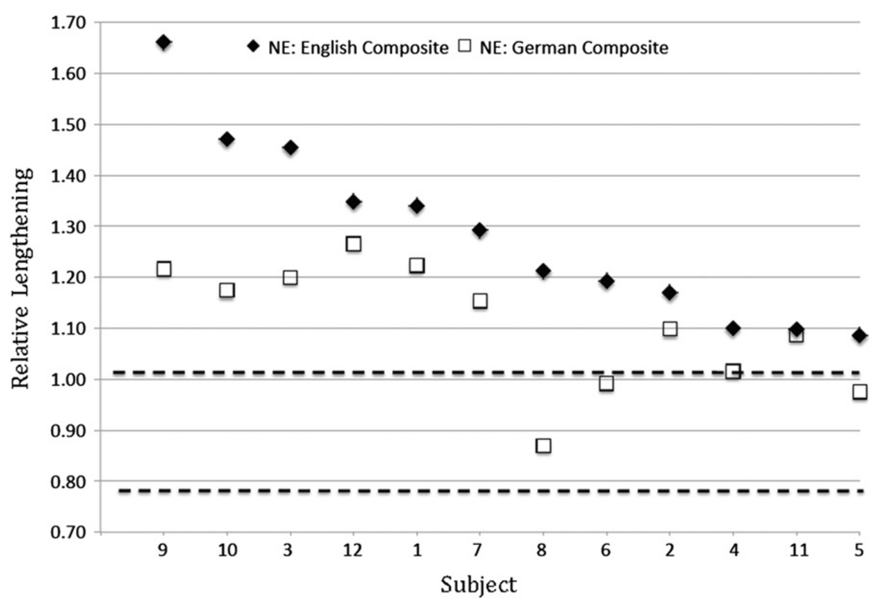

Fig. 5. Rank ordering for 12 native English (NE) speakers' voicing composite in English (black diamonds), which is paired with each individual's composite German value (white squares). A value of 1.00 indicates that for the four acoustic parameters, collectively, no lengthening occurred. Values greater than 1.00 indicate that for the four acoustic parameters, collectively, some amount of lengthening occurred, whereas values less than 1.00 indicate that "shortening" occurred for the four parameters, collectively. The two horizontal dashed lines show the range of values for 13 native German speakers' composite voicing for voiced stop targets in German (from Smith et al., 2009).

versus voiceless contrast for final stops in English, a "composite" relative voicing value was calculated in an attempt to represent how these four measures *collectively* reflected the contrast between voiced and voiceless target consonants when the subjects were speaking English versus German. Since it is not known whether some of these parameters should be weighted more than any of the others and because some subjects could implement them differently from others (Bohn, 1995; Escudero & Boersma, 2004; Flege, Bohn, & Jang, 1997; Kondaurova & Francis, 2008; Morrison, 2005), the composite relative values that were calculated factored in all four parameters equally. As shown in Fig. 5, all 12 native English speakers had at least a somewhat greater composite relative voicing value when speaking English versus German (i.e., black diamonds are above open squares), but the differences across languages were not very large for several of them (e.g., 11, 2, 12, 4). On average, the native English talkers' composite voicing value was 1.29 when speaking English compared to 1.11 when they were speaking German (paired $t=4.777$; df=11; $p<.001$). Although the 12 native English-speaking subjects had a smaller composite voicing value when speaking German, their German composite value was nonetheless greater than that of the native German talkers (.91) studied by Smith et al. (2009)—unpaired, Welch-corrected $t=5.016$; df=13; $p<.0001$.

It can also be seen in Fig. 5 that four of the native English talkers (8, 6, 4, 5) had composite voicing values when speaking German that were within the range of the native German speakers (i.e., between the two horizontal lines). However, subjects 4 and 5 showed reasonably small decreases in German relative to their L1 production patterns (on the order of .10), whereas subjects 8 and 6 showed decreases on the order of .20–.35, suggesting that they may have implemented intentional changes in one or more of the acoustic parameters in an attempt to devoice their final stops. Although the other eight native English speakers did not devoice as much as the native German speakers, several of them (e.g., 9, 10, 3) also showed quite large decreases (.25–.45) in their composite voicing measures when comparing L1 and L2.

3.6. Relationships among the acoustic measures

In terms of the individual acoustic parameters measured for the orthographically similar words in the two languages, three of the four parameters showed statistically significant differences when comparing the native English speakers' productions of English versus German underlying voiced stop targets. Relative vowel duration preceding final voiced stops, relative consonant closure duration of voiceless versus voiced stops, and the relative amount of glottal pulsing during voiced stop targets were all less, on average, when the subjects spoke German versus English. Final consonant closure duration showed the largest relative contrast between the two languages, i.e., the voiceless/voiced closure duration lengthening ratio was .29 less in German than English. By comparison, the relative vowel lengthening ratio was .16 less in German than English, and relative glottal pulsing during stop closure was .13 less in German. In addition, "composite voicing" values averaged .18 less in German than English.

When comparing the native English speaking subjects' individual acoustic parameters for their English versus German productions, significant correlations between the two languages were found for relative closure duration of voiceless versus voiced stops (r=.65; $p<.05$), percentage of glottal pulsing in underlying voiced targets (r=.62; $p<.05$), and relative burst duration of voiceless versus voiced stop targets (r=.60; $p<.05$). The composite voicing measure also showed a significant correlation for the subjects' English versus German productions (see Fig. 6; r=.68; $p<.01$). That is, subjects who tended to produce characteristically "voiced" stops when speaking English (e.g., more glottal pulsing during voiced targets, etc.) devoiced German voiced stop targets to a lesser extent, and vice versa.

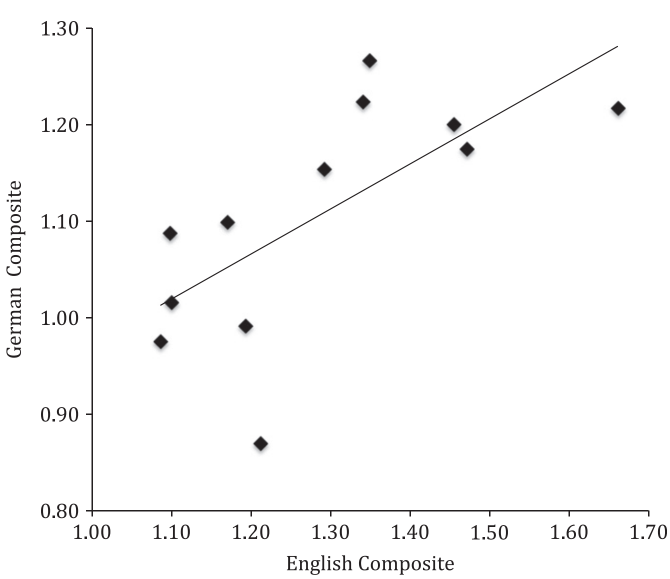

Fig. 6. Scatterplot of the 12 native English talkers' composite voicing measures in English versus German, i.e., based on relative vowel lengthening before voiced stops, relative voiceless versus voiced stop closure duration, percentage of glottal pulsing during underlying voiced stop consonant targets, and relative voiceless versus voiced burst duration.

4. Discussion

Although a number of acoustic studies have examined various issues pertaining to final obstruent devoicing in different languages (e.g., whether "complete neutralization" of an underlying voicing contrast does or does not occur), one line of research that has received relatively little attention – especially from an acoustic perspective – concerns how readily native speakers of English who are learning a second language that manifests final obstruent devoicing can acquire this L2 pattern. To this end, the present study investigated a relatively homogeneous group of native American English speakers at an early stage in the process of learning German, i.e., they were all in their second semester of undergraduate German language courses. Based on the auditory models their instructors provided, as well as the extent to which they attended to those models, and/or depending on whatever information they may have obtained from their course workbooks regarding the devoicing pattern of German, it was anticipated that these L2 learners would have varying, but reasonably limited awareness of the Standard German pattern of final obstruent devoicing. In addition, it was expected that any information about devoicing that they may have acquired through their classroom experiences might be offset to some extent by the orthographic representations of German words they were exposed to, i.e., because the German spelling system manifests the underlying, word-final voicing contrast in word pairs such as "To*d*/to*t*, Ra*d*/Ra*t*, Ba*d*/ba*t*", etc. In this regard, Young-Scholten (2002) noted that orthographic input "serves to strengthen the learners' initial assumption that the phonemes /b, d, g/ are phonetically realized as [b,d, g] in German in all positions, just as they are in English, because they are spelled the same regardless of their position" (p. 271). Thus, learning to devoice underlying voiced targets may not be as simple or as straightforward as might be assumed, making it reasonable to expect that learners are likely to exhibit considerable individual variation in their acquisition of the German final devoicing pattern.

To examine these issues, the present study assessed the extent to which American English learners of German devoiced final, voiced stop targets and how their German productions compared to their English productions of orthographically similar target words (e.g., German "Tod/tot" and English "toad/tote"). Although the results of a repeated measures ANOVA indicated that, overall, there was a significant difference between the various temporal measures when the subjects spoke English versus German, many of the individual measures of specific interest were not significantly different when comparing the two languages. Vowel duration preceding voiced consonants was the primary absolute measure that differed between the two languages, with vowels being significantly shorter when the subjects spoke German. Otherwise, they tended to utilize very similar temporal patterns in speaking both languages; that is, their temporal patterns in English largely seemed to carry over into German. Because they were quite early in the process of learning German as a second language and were presumably largely focused on learning vocabulary items and various grammatical principles, the fact that their pronunciation continued to reflect English-like patterns is not surprising.

Nevertheless, despite having quite comparable amounts of formal experience learning German, considerable variation was observed in the performance of the 12 individual subjects—as seen in Figs. 2–6. As a group, they did show a modest tendency toward devoicing German underlying, word-final voiced stop targets when compared with orthographically similar, English word-final voiced stops. However, they did not neutralize the German underlying voicing contrast to the same extent as native German talkers speaking German (Smith et al., 2009). This finding is consistent with observations by Dmitrieva et al. (2010) for native speakers of American English who were learning Russian as a second language. Their native English speakers also produced significant duration differences between members of Russian minimal pairs that had underlying word-final voicing contrasts. Similar to the present investigation, Dmitrieva and colleagues observed differences in vowel duration preceding underlying voiced versus voiceless targets, final stop closure or fricative duration differences, and the amount of voicing present during final stop closure/fricative intervals, as opposed to the (near) neutralization shown by native talkers.

Given that all the subjects in the present study were at a relatively similar stage of learning German, they constituted a reasonably homogeneous group in that respect, but it is not possible to account for differences in individual language learning abilities, motivation, attention to their instructors' models, class attendance, etc., and how such factors may have impacted any person's tendency to devoice word-final underlying voiced stop targets in German (Piske et al., 2001). As is commonly the case for both native and non-native speakers (e.g., Smith, 2000, 2002; Smith & Hayes-Harb, 2011), there was considerable variation among the individual subjects in the present study, with some producing little or no difference between underlying voiced stop targets in German versus final voiced stops in English. That is, not everyone showed a tendency to devoice. This variation was shown in Fig. 5, for example, as several of these English speakers' (8, 6, 4, 5) acoustic parameters for the relative "composite voicing" value when speaking German were within the range of values shown by the native German talkers speaking German in a study by Smith et al. (2009). In contrast, the other native English speakers had values above the high end of the range of performance shown by the native German speakers. Although no perceptual assessment of the native English speakers' German productions was conducted, one might speculate that, all else being equal, subjects 8, 6, 4, and 5 would be perceived by native German listeners as being less accented than the other native English talkers. However, in terms of what the findings of this investigation may mean for native speakers of English attempting to learn German, it is important to recognize that pronunciation and issues pertaining to accent are only one component of learning a second language.

While the present study determined that there were a number of significantly different absolute and relative acoustic measures when comparing the orthographically similar English versus German words, it was also of interest to note in Fig. 6 that those subjects who had lower *English* composite voicing values (i.e., they tended to devoice voiced stop targets to some degree in English) were often the ones who produced somewhat less voicing in underlying voiced stop targets in German. That is, subjects 8, 6, 2, 4, 11, 5 (in Fig. 5) produced voiced stop targets in English that were "weaker" in terms of what would be considered typical parameters of voicing, and they also had a tendency toward a greater amount of devoicing in German. Therefore, it seems that the inclination to devoice (or not devoice) word-final voiced stop targets may not necessarily have been entirely a result of having learned something about this German pattern from their instructors and/or other sources; rather it could have been at least partly related to whatever proclivities these subjects had toward some amount of devoicing in their English productions. That is, there were certain similarities in their production patterns when speaking English (in which partial devoicing commonly occurs due to aerodynamic, and/or possibly dialect factors) and when they were speaking German (in which devoicing is expected for phonological reasons).

Regarding the possible impact orthography may have had on the subjects learning to devoice word-final stops, Bassetti (2008) observed that orthography can have both positive and negative consequences in learning the phonology of a second language. In terms of negative consequences, she noted (referring to Young-Scholten, 2002) that "L2 orthographic input can also lead learners to produce contrasts that do not exist in the L2 acoustic input. In German, word-final obstruents are always devoiced. Although L2 acoustic input contains no voiced obstruents in word-final position, English learners of German pronounce some word-final obstruents as voiced, presumably because they are spelled as voiced obstruents, for instance pronouncing [d] instead of [t] in ⟨Bund⟩" (p. 197). Another example she provided does not relate specifically to German, but does demonstrate the effect that orthography can have on L2 phonological acquisition—in this instance pertaining to consonant length. She noted that "In some languages consonant length is contrastive; for instance, in Italian /kɔpia/ and /kɔpːia/ mean 'copy' and 'couple' respectively…In the Italian orthography, these geminates are represented by double consonant letters, e.g. ⟨p⟩ vs. ⟨pp⟩ in ⟨copia⟩ and ⟨coppia⟩. In English phonology there is no contrast between short and long consonants, but English orthographic words can contain double consonant letters. There is evidence that Italian ESL learners pronounce long consonants in English words that are spelled with double consonant letters. For instance, all the Italian children in Browning's study…pronounced the [p] in 'apple' with a closure that was 50% longer than the average closure in /p/" (p. 197).

Bassetti (2006) also suggested that "orthographic input results in non-target-like pronunciations that cannot be explained in terms of L1 influence or L2 phonological input, and to differences between the early phonologies of children and L2 learners. *The most likely explanation is that the orthographic input is interfering with the phonological input in the creation of L2 learners' mental representations of the L2 phonology*, at both the segmental and suprasegmental level" (p. 97, our emphasis). In addition, in a study of L2 acquisition of Spanish /b, d, g/, Zampini (1994) observed that "native language transfer plays an important role in hampering the acquisition of the voiced spirants" among native speakers of English, in that "students largely fail to spirantize the voiced stops in L2 speech and incorrectly transfer the phonemic status of English…to Spanish." She also noted that "the presence of orthographic 'v'…interferes with acquisition of Spanish [b]…" (p. 470). Finally, Young-Scholten (1995) explained that "when the L2 learner confronts the graphemes of a new language at the start of acquisition, prior to the establishment of the L2 phonology, the learner will be compelled to search for the phonological constituents which these graphemes might represent. Because the L2 phonology has not yet been established, the learner will only be able to access L1 phonology" (p. 113).

5. Conclusion

The present study found a number of significantly different absolute and relative acoustic measures when comparing orthographically similar English versus German words, including vowel duration preceding final consonants, final consonant closure duration, glottal pulsing during stop closure, and a composite voicing measure obtained by averaging relative patterns of all four acoustic parameters that were examined. The fact that the L2 German learners in this investigation did not perform the same in German as in English in their production of word-final voiced stop targets suggests that at least some of them may have learned the German word-final devoicing pattern to some degree and were able to incorporate it into their speech by modifying one or more of the relevant acoustic parameters. Thus, although German reflects the voiced/voiceless contrast in its orthographic system, a number of these German learners appear to have been able to overcome the influence of orthography to some extent due to the pronunciation models provided by their course instructors and/or workbook exercises they may have completed (Simon, 2010). Nevertheless, given the many lexical, grammatical, and other phonetic and phonological factors they were also attempting to learn in the early months/semesters of studying German as a second language, most of them showed only limited progress in acquiring the pattern of word-final devoicing in German. Moreover, to the extent they did manifest this pattern, it is possible that at least part of the reason for any success in this regard may have been related to the degree to which they also tended to manifest some degree of phonetic (as opposed to phonological) devoicing in their native *English* productions of word-final, voiced obstruents. That is, students who, because of their dialect or other phonetic/phonological patterns, tend to have a somewhat less salient final voicing contrast in English may appear to make greater progress toward this phonological pattern in German than students who exhibit a "stronger" voicing contrast in English. However, it is possible that they have at least, in part, transferred their English (L1) production pattern to German (L2), and may have little if any awareness of what they are producing in this regard in either English or German.

Acknowledgments

We would like to thank Trevor Perry for his assistance with data collection and Rachel Hayes-Harb for helpful comments on an earlier version of this manuscript. We also greatly appreciate a number of helpful suggestions provided by three anonymous reviewers and the Associate Editor.

References

Abu-Rabia, S. (2001). Testing the interdependence hypothesis among native adult bilingual Russian-English students. *Journal of Psycholinguistic Research*, *30*, 437–455.

Bassetti, B. (2006). Orthographic input and phonological representations in learners of Chinese as a Foreign Language. *Written Language and Literacy*, *9*, 95–114.

Bassetti, B. (2008). Orthographic input and second language phonology. in: T. Piske, & M. Young-Scholten (Eds.), *Input Matters in SLA* (pp. 191–206). Clevedon, UK: Multilingual Matters.

Boersma, P. & Weenink, D. (2006). Praat: doing phonetics by computer (Version 5.2.1.5) Computer program from 〈http://www.praat.org/〉.

Bohn, O.-S. (1995). Cross language speech perception in adults: First language transfer doesn't tell it all. in: W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 279–304). Baltimore: York Press.

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

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*, 483–492.

Duyck, W. (2005). Translation and associative priming with cross-lingual pseudo-homophones: Evidence for nonselective phonological acquisition activation in bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 1340–1359.

Eckman, F. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning*, *27*, 315–330.

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

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition*, *26*, 551–585.

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*, 437–470.

Flege, J. E., Frieda, E. M., & Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics*, *25*, 169–186.

Fourakis, M., & Iverson, G. K. (1984). On the 'incomplete neutralization' of German final obstruents. *Phonetica*, *41*, 140–149.

Fullana, N., & Mora, J. C. (2009). Production and perception of voicing contrasts in English word-final obstruents: Assessing the effects of experience and starting age. in: M. A. Watkins, A. S. Rauber, & B. O. Baptista (Eds.), *Recent research in second language phonetics/phonology* (pp. 97–117). Newcastle upon Tyne, UK: Cambridge Scholars Publishing.

Hillenbrand, J., Ingrisano, D. R., Smith, B. L., & Flege, J. E. (1984). Perception of the voiced–voiceless contrast in syllable-final stops. *Journal of the Acoustical Society of America*, *76*, 18–26.

Holmes, J. (1996). Losing voice: Is final /z/ devoicing a feature of Maori English?. *World Englishes*, *15*, 193–205.

Jassem, W., & Richter, L. (1989). Neutralization of voicing in Polish obstruents. *Journal of Phonetics*, *17*, 317–325.

Keating, P., Linker, W., & Huffman, M. (1983). Patterns in allophone distribution for voiced and voiceless stops. *Journal of Phonetics*, *11*, 277–290.

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.

Koenig, L. L., Fuchs, S., & Lucero, J. C. (2011). Effects of consonant manner and vowel height on intraoral pressure and articulatory contact at voicing offset and onset for voiceless obstruents. *Journal of the Acoustical Society of America*, *129*, 3233–3244.

Kondaurova, M. V., & Francis, A. L. (2008). The relationship between native allophonic experience with vowel duration and perception of the English tense/lax vowel contrast by Spanish and Russian listeners. *Journal of the Acoustical Society of America*, *124*, 3959–3971.

Kuzla, C., Cho, T., & Ernestus, M. (2007). Prosodic strengthening of German fricatives in duration and assimilatory devoicing. *Journal of Phonetics*, *35*, 301–320.

Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.

Lemhoefer, K., & Dijkstra, T. (2004). Recognizing cognates and interlingual homographs: Effects of code similarity in language-specific and generalized lexical decision. *Memory and Cognition*, *32*, 533–550.

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

Morrison, G. S. (2005). An appropriate metric for cue weighting in L2 speech perception. *Studies in Second Language Acquisition*, *27*, 597–606.

Moyer, A. (1999). Ultimate attainment in L2 phonology: The critical factors of age, motivation, and instruction. *Studies in Second Language Acquisition*, *21*, 81–108.

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

Piske, T., MacKay, I. R., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, *29*, 191–215.

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

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

Rodgers, B., & Fuchs, S. (2010). How intraoral pressure shapes the voicing contrast in American English and German. *Zentrum für Allgemeine Sprachwissenschaft Papers in Linguistics*, *52*, 63–82.

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

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

Slowiaczek, L. M., & Szymanska, H. (1989). Perception of word-final devoicing in Polish. *Journal of Phonetics*, *17*, 205–212.

Smith, B. L. (1979). A phonetic analysis of consonantal devoicing in children's speech. *Journal of Child Language*, *6*, 19–28.

Smith, B. L. (2000). *Variations in temporal patterns of speech production among speakers of English*, *108*, 2438–2442*Journal of the Acoustical Society of America*, *108*, 2438–2442.

Smith, B. L. (2002). Effects of speaking rate on temporal patterns of English. *Phonetica*, *59*, 232–244.

Smith, B. L., & Hayes-Harb, R. (2011). Individual differences in the perception of final consonant voicing among native and non-native speakers of English. *Journal of Phonetics*, *39*, 115–120.

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*, 257–275.

Smith, B. L., Hillenbrand, J., & Ingrisano, D. (1986). A comparison of temporal measures of speech using spectrograms and digital oscillograms. *Journal of Speech and Hearing Research*, *29*, 270–274.

Smith, C. L. (1997). The devoicing of /z/ in American English: Effects of local and prosodic context. *Journal of Phonetics*, *25*, 471–500.

Stevens, K. N., Blumstein, S. E., Glicksman, L., Burton, M., & Kurowski, K. (1992). Acoustic and perceptual characteristics of voicing in fricatives and fricative clusters. *Journal of the Acoustical Society of America*, *91*, 2979–3000.

Van Wijnendaele, I., & Brysbaert, M. (2002). Visual word recognition in bilinguals: Phonological priming from the second to the first language. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 616–627.

Wade-Woolley, L. (1999). First language influences on second language word reading: All roads lead to Rome. *Language Learning*, *49*, 447–471.

Wang, M., Koda, K., & Perfettin, C. A. (2003). Alphabetic and nonalphabetic L1 effects in English word identification: A comparison of Korean and Chinese English L2 learners. *Cognition*, *87*, 129–149.

Warner, N., Good, E., Jongman, A., & Sereno, J. (2006). Orthographic versus morphological incomplete neutralization effects. *Journal of Phonetics*, *34*, 285–293.
Warner, N., Jongman, A., Sereno, J., & Kemps, R. (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics*, *32*, 251–276.
Westbury, J. R., & Keating, P. A. (1986). On the naturalness of stop consonant voicing. *Journal of Linguistics*, *22*, 145–166.
Young-Scholten, M. (1995). The negative effects of 'positive' evidence on L2 phonology. in: L. Eubank, L. Selinker, & M. S. Smith (Eds.), *The current state of interlanguage* (pp. 107–121). Amsterdam: John Benjamins Publishing Company.
Young-Scholten, M. (2002). Orthographic input in L2 phonological development. in: P. Burmeister, T. Piske, & A. Rohde (Eds.), *An integrated view of language development: Papers in honor of Henning Wode* (pp. 263–279). Trier, Germany: Wissenschaftlicher Verlag.
Young-Scholten, M. (2004). Prosodic constraints on allophonic distribution in adult L2 acquisition. *International Journal of Bilingualism*, *8*, 67–77.
Zampini, M. L. (1994). The role of native language transfer and task formality in the acquisition of Spanish spirantization. *Hispania*, *77*, 470–481.