2019-03-07

Using Electropalatography to Analyze Intra-Speaker Variability in German Second Language Fricative Production

Elizabeth D. Young
Brigham Young University

Honors Thesis

USING ELECTROPALATOGRAPHY TO ANALYZE INTRA-SPEAKER VARIABILITY IN GERMAN SECOND LANGUAGE FRICATIVE PRODUCTION

by
Elizabeth D. Young

Submitted to Brigham Young University in partial fulfillment
of graduation requirements for University Honors

Communication Disorders Department
Brigham Young University
April 2019

Advisor: Shawn L. Nissen

Honors Coordinator: David L. McPherson

ABSTRACT

USING ELECTROPALATOGRAPHY TO ANALYZE INTRA-SPEAKER VARIABILITY IN GERMAN SECOND LANGUAGE FRICATIVE PRODUCTION

Elizabeth D. Young

Communication Disorders Department

Bachelor of Science

Electropalatography (EPG) is a biofeedback system that tracks tongue contact with the palate during speech. The current study uses EPG data to examine the variability that occurs in speech production for individuals speaking a second language (L2). Five native German speakers and twelve native English speakers learning German as a second language were asked to produce the fricatives [ç], [x], and /ʃ/ in various linguistic contexts. Variability in center of maximal contact across the anterior-posterior dimension ("Center of Gravity" – COG) and duration for L2 production across sound type, L2 production across task type, L2 production across subject, and native language (L1) vs L2 productions was examined. The COG variability in L2 speakers was significant as a function of sound type, task type, subject, and sound type by task type. Duration variability in L2 speakers was significant as a factor of sound type and task type. L2 variability across task type generally increased as task type increased in complexity. The COG and duration variability were not found to be significant across language status. However, the COG variability was found to be significant between languages as a function of sound type. L1 speakers displayed a lower COG variability for the fricatives [ç] and [x], whereas L2 speakers displayed a lower COG variability for the fricative /ʃ/. Further research is needed to investigate the nature of L2 variability, but it is anticipated that the current thesis will contribute to a better understanding of the L2 learning process.

ACKNOWLEDGMENTS

I'd like to offer a special thanks to Dr. Nissen for all his guidance and support throughout the thesis process – this project would not have been possible without his help. I would also like to recognize the members of my thesis committee and the staff at the Honors office for providing me with the advice and support I needed to succeed. Finally, I'd like to say a special thanks to my husband – for being my number cruncher, my proofreader, my sounding board, and my endless source of encouragement.

TABLE OF CONTENTS

ABSTRACT II
ACKNOWLEDGMENTS IV
LIST OF TABLES VI
LIST OF FIGURES VII
I. Introduction 8
II. Methods 14
Participants 15
EPG Sensor 16
Stimuli 16
Native German (L1). 16
Non-native German learners (L2). 17
Procedures 17
Data Analysis 18
III. Results 19
Within-Subject Factors: L2 Learners 20
Sound type. 20
Task type. 22
Subject. 25
Between-Subject Factors: L1 vs L2 Speakers 27
IV. Discussion 30
REFERENCES 36

LIST OF TABLES

Table 1 *Reference studies for the complete methods of each stimulus type examined* ... 15

Table 2 *Number of tokens per stimulus type* ... 18

Table 3 *COG and duration measures by sound type for L2 speakers* ... 22

Table 4 *COG and duration measures by task type for L2 speakers* ... 24

Table 5 *COG and duration measures by fricative and task type for L2 speakers* ... 25

Table 6 *COG and duration measures by subject for L2 speakers* ... 26

Table 7 *COG and duration measures by fricative for L1 and L2 speakers* ... 28

LIST OF FIGURES

Figure 1 COG standard deviation by sound type for L2 speakers 21

Figure 2 Duration standard deviation by sound type for L2 speakers 21

Figure 3 COG standard deviation by task type for L2 speakers 23

Figure 4 COG standard deviation by sound and task type for L2 speakers..................... 23

Figure 5 Duration standard deviation by task type for L2 speakers 24

Figure 6 COG standard deviation by subject for L2 speakers 26

Figure 7 Overall COG standard deviation for L1 and L2 speakers 27

Figure 8 COG standard deviation by fricative for L1 and L2 speakers........................... 28

Figure 9 Overall duration standard deviation for L1 and L2 speakers 29

Figure 10 Duration standard deviation by fricative for L1 and L2 speakers 29

Introduction

Speech is, for the most part, an unseen process; many of the contributing mechanisms, such as the lungs and the soft palate, are difficult to observe with the naked eye. Speech scientists have therefore developed technologies designed to help gain insight into various aspects of the speech production process. One such technology is electropalatography (EPG), a biofeedback tool that records and displays linguapalatal contact. An EPG unit consists of a thin pseudopalate connected to an electronic processing device. The pseudopalate is molded to an individual's hard palate and maxillary teeth in order to interfere as little as possible with normal speech production (Fletcher, 1992). The palate is embedded with sensors which, when contacted by the tongue, complete an electrical circuit. The processing device of an EPG interprets the electronic signals to render a visual representation of linguapalatal contact.

EPG is a research tool that has been used to investigate both typical and disordered speech. Recently, researchers have turned to EPG as a tool to help understand and facilitate the process of second language (L2) acquisition. Second language learning requires a heavy cognitive load as individuals develop new or adjust existing articulatory placement in order to successfully produce L2 sounds. Difficulty with L2 articulation can occur due to the L2 sound being completely novel to the learner, or the L2 containing linguistically significant contrasts between sounds normally collapsed into a single phonetic category in a native language. In both situations, new motor patterns must be established to successfully articulate the L2 speech sound. In the second situation, however, the learner faces the added difficulty of learning to distinguish between the two L2 sounds and produce the sounds with sufficient contrast to be understood by native speakers. L2 learners therefore must learn a variety of new skills in order to successfully master a new language.

As a real-time visual biofeedback device, EPG can be an especially useful tool in helping students successfully acquire L2 language skills. Providing visual feedback for linguapalatal contact patterns can give L2 learners an additional resource during the language learning process. Although limited in number, studies have found that visual feedback can be helpful in aiding L2 students master linguistic contrasts not present in their native language. Gibbon, Hardcastle, and Suzuki (1991) showed that using visual EPG feedback over the course of two weeks was effective in aiding two native Japanese speakers establish a more consistent articulatory contrast between the English /l/ and /r/ sounds (allophonic sounds in the Japanese language). Several years later, Schmidt and Beamer (1998) were successful in replicating these results with native Thai speakers learning English. Schmidt and Beamer noted that sound contrasts that were not allophonic in Thai, such as /t - θ/, were much easier to establish than the allophonic contrast of /tʃ - ʃ/. Schmidt was again able to replicate these positive results with native Korean speakers learning English in a 2012 study.

Using EPG to establish an articulatory contrast can also aid in establishing a "perceptual" contrast between two non-native sounds. As stated above, L2 speakers tend to have particular difficulty learning contrasts that are allophonic in their own language; this is because they often cannot perceptually distinguish between the sounds that they are trying to produce. EPG feedback therefore provides a workaround for these learners, as they are able to use visual feedback rather than their auditory feedback to modify their production. Schmidt (2012) found that participants who trained with visual EPG feedback had an increased ability to perceptually distinguish between the sounds that they were being trained in after EPG intervention. This result occurred without deliberate perceptual training, indicating that visual feedback can aid individuals in training themselves to perceptually categorize non-native sounds correctly.

This assumption is further supported by research findings that support EPG as an effective tool in accent reduction. Bright (1999) used EPG as a tool to help three native Spanish speakers improve their production of English. The treatment course focused on producing a variety of minimal pairs that used contrasts not native to Spanish. After receiving EPG therapy three times per week for four weeks, all three participants showed a significant reduction in their accents both perceptually and as rated by a standardized test. Standardized test results indicated an average of a 50% improvement in the non-accented production of target sounds, and perceptual judges noted a rising trend in the overall "goodness" of production at the word level. These results indicate that despite the relatively artificial nature of EPG therapy, skills learned while using the EPG can be generalized to real-life speech in a manner that is significant to listeners.

During the L2 learning process, it is natural to assume that there will be differences between L1 and L2 production in terms of articulator speed. However, research has shown that even when a speaker develops proficiency in a second language, articulator speed continues to vary significantly between the individual’s first and second language. Nissen, Dromey and Wheeler (2007) revealed that there are multiple kinematic differences between the production of a first and second language. The study examined native Spanish and Korean speakers with relatively high proficiency in English as a second language. Each speaker’s tongue movements (recorded as “strokes”) were captured via a magnetic sensor attached near the tongue tip. It was found that while there was little difference between the overall amount of tongue movement in first and second languages, average stroke speed, peak stroke speed, and the number of strokes per second were all significantly higher in an individual's native language. The researchers also noted that L2 production included more pauses than L1 production. Perceptually, individuals with the lowest stroke speed and the greatest amount of pausing during speech were judged to have the greatest perceived accent. This study demonstrates that even when L2 proficiency is achieved, tongue movement will continue to vary between L1 and L2 production.

As demonstrated by Nissen et al. (2007), variability is a common and integral part of speech production for both L1 and L2 learners. While research exists on the topic of L1 variability, there is very little research examining L2 variability. Gaining a greater understanding of the variability of L2 learners, however, both as compared to other L2 learners and as compared to L1 speakers, could help shed light on the L2 learning process. In order to better understand articulatory variability in L2 learners, one must first consider the variability that exists in L1 speech production. Understanding the research on various forms of variability in speech-sound production allows for a better understanding of why and how L2 variability may exist.

Research has shown that linguapalatal contact becomes less variable as an individual ages. Cheng, Murdoch, Goozée, & Scott (2007) investigated the effect of aging on linguapalatal contact patterns by examining the production of several English sounds in children (aged 6-7 and 8-11 years), adolescents, and adults. While child productions of various sounds grossly resembled adult productions, articulatory placement grew more refined with each progressive age group. The children tended to display greater closure (and therefore higher linguapalatal contact) than their adult counterparts. Additionally, child productions of sounds had an overall higher rate of intra-speaker variability than adult productions. This suggests increased experience producing a particular sound can lead to greater refinement and lower variability in tongue placement while producing that sound. It can be assumed that a similar pattern might take place in second-language learning. Like children refining their L1 production, individuals learning an L2 may initially display large amounts of variability and then slowly refine their production over several years of practice.

Several studies have suggested that the articulatory variability of a speech sound is affected by the context in which the sound is produced. Vowel context was one of the first factors shown to have an effect on variability. Dagenais, Lorendo, & McCutcheon (1994) described /i/ as the vowel context in which consonants displayed the lowest variability. Evidence gathered by Dromey and Sanders (2009) also supports /i/ as the least variable vowel context. Linguistic contexts involving vowels produced in the back of the mouth have been shown to lead to greater linguapalatal variability; however, there is disagreement among studies of which of the back vowels lead to the greatest variability (Dagenais et al., 1994; Dromey et al., 2009). There is evidence to suggest that all sounds, whether they be vowel or consonant, tend to show higher variability as the place of articulation moves further back into the mouth (Dromey et al., 2009). Other factors, such as voicing and manner of articulation, have also been shown to affect variability (Dagenais et al., 1994; Dromey et al., 2009). These results demonstrate the complex and interrelated nature of speech sound production; such factors must be taken into consideration when considering the variability of speech sounds.

While the linguistic context can affect the variability of a speech sound, evidence suggests that some sounds are simply intrinsically more variable than others. McAuliffe, Ward, & Murdoch (2001) used EPG to examine the consonants /t, s, k, l/ in the /i/ vowel context. This study found that /l/ was produced with the most amount of variability between speakers, while /s/ exhibited the least amount of variability. These results were echoed and expanded upon by Dromey et al. (2009), who found that each of the fifteen consonants examined displayed differing levels of variability. Stop consonants, particularly nasal stops, were among the most variable, whereas fricatives tended to have a low degree of variability. Other sounds, such as /l/ and /r/, displayed wide ranges of variability among the different participants. These results make sense considering the articulatory demands placed on the tongue by various consonants. Stop consonants do not require extremely specialized placement, only the creation and subsequent release of an airway seal; fricatives, however, require precise placement in order to produce the proper sibilance. Similarly, sounds such as /l/ and /r/ can be produced accurately despite moderate variability in tongue placement.

One factor about variability that must be noted is that not all variability can be predicted. In 1989, Butcher conducted an EPG study in order to quantify normal variability of tongue contact caused by coarticulation. Butcher began by noting, however, that "variability is itself variable… in that in some instances its occurrence and extent is more predictable than in others" (Butcher, 1989, p.39). This statement speaks to the many variables that may impact variability in tongue motion – only some of which are noted above. Further complicating matters is the fact that some individuals seem to be inherently more variable in their articulation than others, despite having perceptually normal speech (Dromey et al., 2009).

EPG use in second language learning and variability in an L1 have been examined, but further research is needed on the articulatory variability that exists in L2 learners. The aims of this study are therefore to examine (1) the intra-speaker variability in linguapalatal contact for L2 German speakers for the fricatives [ç], [x], and /ʃ/, and (2) the extent to which intra-speaker variability differs between native and L2 German speakers. It is hypothesized that due to their relative lack of experience in producing L2 sounds, L2 learners will experience significantly more variance in the production of non-native speech sounds than native language speakers. It is hoped that the results of this study will lead to a better understanding of speech production during the second language learning process.

Methods

This project is part of a larger group of studies examining the efficacy of EPG use in second language learning. This study draws upon data collected and analyzed by Isaacson (2015), Lester (2017), and Cope (2018). It should be noted that while this study describes the general methods for data collection and analysis that each previous study held in common, small discrepancies in methods exist between each study. The author encourages readers to consult the original studies for an in-depth description of the methods used to gather each form of stimuli. A summary of the stimulus types and the study from which the data originated can be found in Table 1.

Table 1

Reference studies for the complete methods of each stimulus type examined

Native Language	Task Type	Reference Study
German	Real Words	Isaacson (2015)
German	Nonsense Words	Isaacson (2015)
English	Real Words	Lester (2017)
English	Nonsense Words	Cope (2018)
English	Sentences	Cope (2018)
English	Spontaneous Speech	Cope (2018)

Participants

This study examined two groups of participants: 6 native German speakers and 12 non-native German language learners. The native German speakers were young adults (aged 20-25) who were currently residing in the United States. Each participant identified themselves as native speakers of High German (Hochdeutsch) speakers who continue to use their native language regularly. Due to data collection errors, German participant 1 is not included in this study, resulting in a total of 5 native German-speaking subjects. The non-native German learning participants were young adults (aged 17-25) enrolled at Brigham Young University. Each L2 participant had completed an introductory-level university course in German and was currently taking a second German course. All participants, both L1 and L2, denied having any history of speech, language, or hearing impairments. All participants were judged to be anatomically typical in the development of their vocal tract, dentition, and hard palate. Each participant signed an IRB-approved informed consent document prior to their participation.

EPG Sensor

Each participant was required to visit a dentist's office and obtain an impression of their upper teeth and palate prior to participation in the study. The dental impressions were used by SmartPalate International© to create a customized EPG device for each participant. Each EPG unit consisted of a relatively thin (2mm) pseudopalate molded to fit the individual's dentition and palate, extending from the alveolar ridge to the back molars. Each pseudopalate contained 124 gold-plated electrodes, situated in a grid-like pattern across the pseudopalate surface. Each EPG unit connected to a microprocessor I/O device for electrode data processing, which was worn around the user's neck. The contact patterns were transferred via USB cord to a computer, where a specialized program displayed real-time linguapalatal contact patterns.

Stimuli

The native and non-native productions of the following German sounds were evaluated: [ç] ("*ich*-Laut"), [x] ("*ach*-Laut"), and /ʃ/. All productions were produced in the final position of various VC contexts. The vowel contexts examined were high-front (HF – the /I/ vowel), high-back (HB – the /u/ vowel) and low back (LB – the /ɑ/ vowel). Because [ç] and [x] are context-dependent sounds in German, [ç] was gathered only in front vowel contexts and [x] was gathered in back vowel contexts.

Native German (L1).

L1 speakers were asked to produce all three target sounds from stimuli lists of both real and nonsense words. Each word was embedded in the carrier phrase "*Ich sage…"* (I say). Table 2 describes the number of tokens gathered for each stimulus sound.

See Isaacson (2015) for additional details regarding data collection and complete stimuli lists.

Non-native German learners (L2).

L2 speakers were asked to produce the target sounds in a variety of contexts. Stimuli were collected embedded within real words, nonsense words, sentences, and spontaneous speech. The sounds [x] and [ç] were gathered in the appropriate vowel contexts for all task types. However, the /ʃ/ sound was only gathered in the context of real words. Stimuli embedded in real and nonsense words were presented in the carrier phrase "*Ich sage das wort…"* (I say the word). Target words in spontaneous samples were elicited through a serious of open-ended interview questions. See Lester (2017) for additional details regarding data collection and complete stimuli lists for real words and Cope (2018) for all other task types. See Table 2 for a description of the number of tokens gathered for each stimuli type.

Procedures

After signing an informed consent form, participants were fitted with their EPG unit. Each participant was allowed several minutes to adjust to the unit prior to data collection. Participants were familiarized with the stimuli before being given a stimuli list and asked to read them aloud. EPG data were recorded and saved to a PC in *.csv* format. Audio recordings of each data gathering session were also saved as *.wav* files. Please see Isaacson (2015), Lester (2017) and Cope (2018) for more complete details regarding procedures for the collection of each stimuli type.

Table 2

Number of tokens per stimulus type

Language and Task Type	Vowel	Sound [ç]	[x]	/ʃ/
L1 Real Words	HF	15	-	5
	HB	-	15	5
	LB	-	15	5
L1 Nonsense Words	HF	10	-	10
	HB	-	10	10
	LB	-	10	10
L2 Real Words	HF	9	-	3
	HB	-	9	3
	LB	-	9	3
L2 Nonsense Words	HF	5	-	-
	HB	-	5	-
	LB	-	5	-
L2 Sentences	HF	6	-	-
	HB	-	2	-
	LB	-	4	-
L2 Spontaneous	HF	varied	-	-
	HB	-	varied	-
	LB	-	varied	-

Data Analysis

High-quality sound recordings of each stimuli type were segmented in order to determine fricative onset and offset times to the nearest millisecond. Timepoints exhibiting a sudden increase or decrease of diffuse noise energy were used to determine the fricative onset and offset times, respectively. Tokens that were incorrectly produced or that displayed onsets or offsets that were difficult to pinpoint were discarded from further analysis. Finalized timepoints were run through a custom-designed MATLAB program to calculate the average electrode activation for each token. At this point in the analysis process, all three previous studies collapsed all tokens of each sound to create an average representation of the sound production. This study, however, uses the un-collapsed data to examine the differences that exist between each production of the target sounds.

Results

Variability in linguapalatal contact is reported using descriptive statistics for within-subject (subject, sound type, and task type) and between-subject (L1 vs L2) variables. Measurements for the center of gravity (COG) and duration (in ms) are provided for the fricatives [ç], [x], and /ʃ/ in their various vowel contexts. For the purposes of this study, each fricative is examined in relation to its various vowel contexts (with each consonant-vowel combination referred to as a "sound type"). Unless otherwise noted, figures and tables collapse all sound types into their respective fricative categories. In order to represent variability, standard deviation is used for the dependent variable in all calculations. Means are calculated using the overall mean, whereas standard deviation is presented as the average standard deviation by participant.

The COG index is designed to specify the area of greatest electrode activation, with increasing values indicating more anterior tongue contact. Following the descriptions made by Lester (2017) and Cope (2018):

> The COG measure was computed by taking the sum of the percentage of activation of each row multiplied by a relative power scale, divided by the total degree of activation across the EPG sensor. The power scale was dispersed from 1 to 14 across the posterior 14 rows of the EPG sensor (excluding the initial row of sensors that measure lip contact). Therefore, the power weight of the 2nd row was 14, the 3rd row was 13, the 4th row was 12, etc. The middle four sensors of rows

8, 13, 14, and 15, as well as the central region were excluded from the COG calculation. This was because of limited contact of the electrodes in these areas due to the shape of the dental arch and the nature of the fricative productions.

Thus, the computation of the COG measure is as follows:

$$COG = \sum [A_r (P)] / \sum(A_t)$$

A_r = percentage of electrode activation for each row
A_t = percentage of total electrode activation
P = relative power scale from 1 to 14 (Lester, 2017; Cope, 2018)

L2 Learners

Sound type.

The COG variability was found to differ significantly across sound type ($F(5,133) = 11.995$, $p < 0.001$, $\eta^2_{partial} = 0.218$). As seen in Figure 1, both vowel contexts of the [x] fricative displayed high variability (mean = 0.48), while all three vowel contexts of the /ʃ/ fricative displayed relatively low variability (mean = 0.17). The [x] fricative in the LB context displayed the highest variability (SD = 0.60), whereas the LB /ʃ/ displayed the lowest variability (SD = 0.15). The variability of duration was also found to be significant across sound type ($F(5, 133) = 3.341$, $p = 0.007$, $\eta^2_{partial} = 0.036$). Figure 2 displays a high duration variability for L2 production of the [ç] fricative (SD = 81 ms), with all other sounds demonstrating a variability level between 30 and 40 ms. See Table 3 for the COG and duration mean and standard deviation measures by sound type.

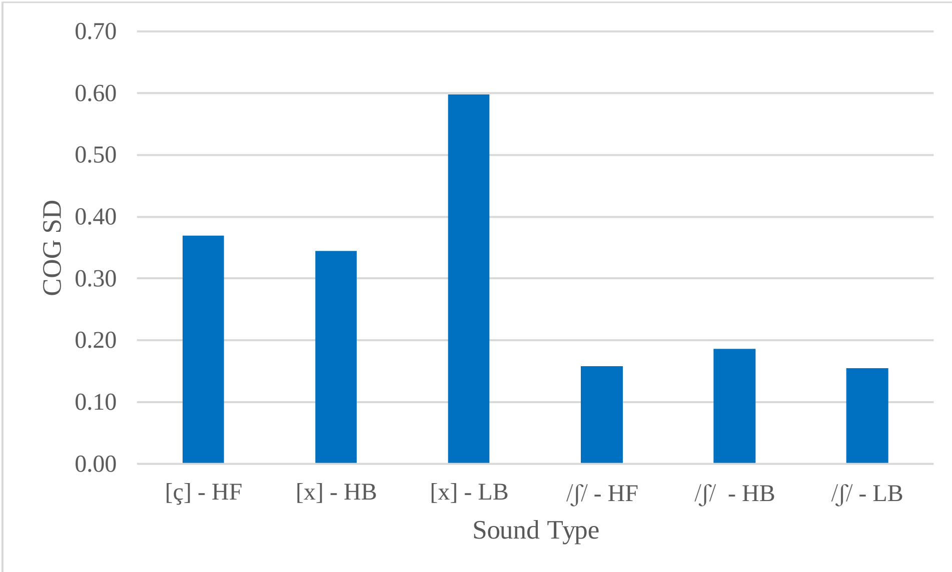

Figure 1. COG standard deviation by sound type for L2 speakers

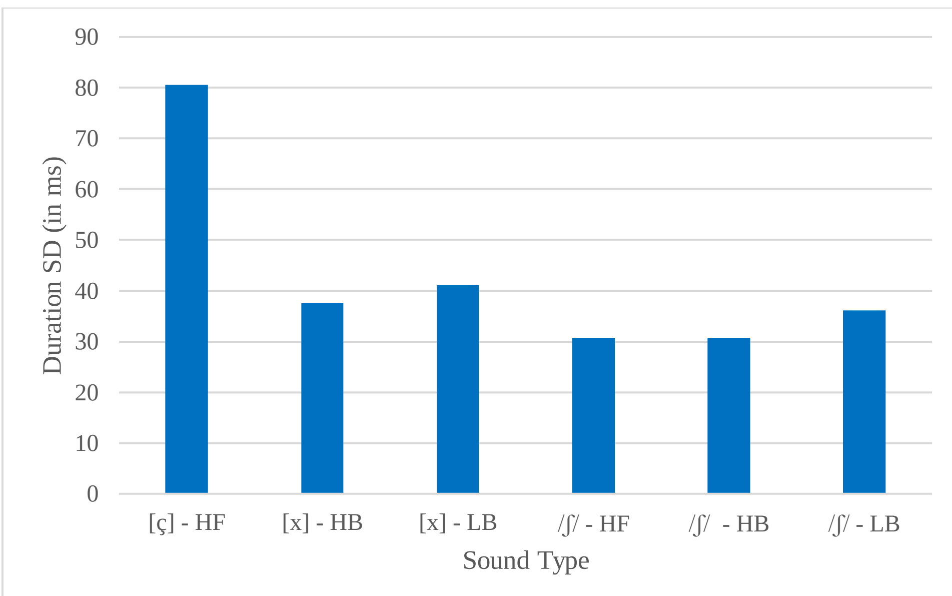

Figure 2. Duration standard deviation by sound type for L2 speakers

Table 3

COG and duration measures by sound type for L2 speakers

Measure	Vowel	Sound					
		[ç]		[x]		/ʃ/	
		Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*
COG	HF	4.73	0.37	-	-	5.51	0.16
	HB	-	-	3.34	0.34	5.39	0.19
	LB	-	-	3.31	0.60	5.50	0.15
Duration (ms)	HF	198.00	81.00	-	-	304.00	31.00
	HB	-	-	227.00	38.00	293.00	31.00
	LB	-	-	216.00	41.00	306.00	36.00

Task type.

There was a significant relationship found between the COG variability across task types ($F(3, 133) = 4.277$, $p < 0.001$, $\eta^2_{partial} = 0.088$). As seen in Figure 3, the COG values steadily increased with increasingly complex task types; real words displayed an average variability of 0.26, while words in sentences and spontaneous speech each averaged a variability of 0.53. Additionally, a significant difference was found between the COG variability in L2 sound type production as a function of task type ($F(5, 133) = 3.162$, $p = 0.001$, $\eta^2_{partial} = 0.106$). Figure 4 again displays a general rising trend in the COG measures as task type complexity increases.

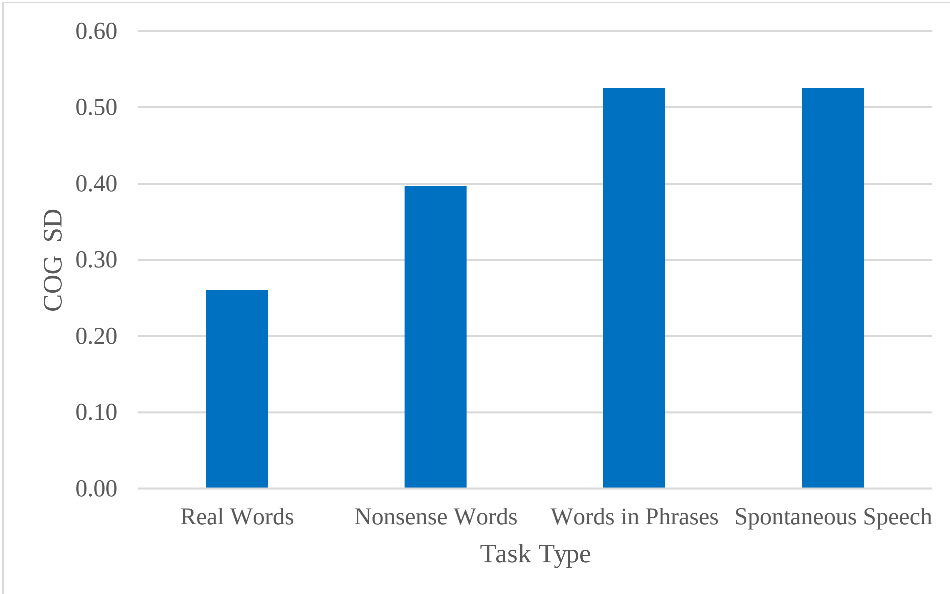

Figure 3. COG standard deviation by task type for L2 speakers

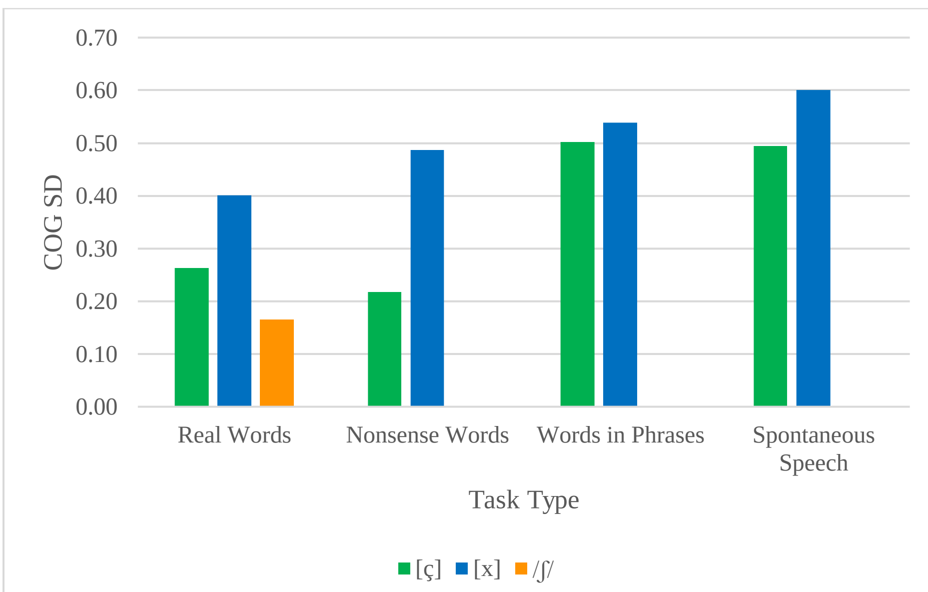

Figure 4. COG standard deviation by sound and task type for L2 speakers

Duration variability as a function of task type was found be significant, $F(3,133) = 9.729$, $p < 0.001$, $\eta^2_{partial} = 0.180$. As seen in Figure 5, spontaneous speech displays a high duration variability (SD = 136 ms) as compared to the other task types. Duration variability in sound type as a function of task type did not vary significantly, although it did near significance with a p-value of 0.061. Post hoc analysis of the data showed that the [ç] fricative displayed very high variability in the spontaneous setting. When the spontaneous utterances of the [ç] fricative were removed from analysis, duration variability by task type was no longer significant; variability by task type as a function of sound type, however, was found to be significant ($F(4, 122) = 2.728$, $p = 0.032$, $\eta^2_{partial} = 0.082$). Tables 4 and 5 can be seen for the COG and duration variability measures across task type and across task type as a function of fricative, respectively.

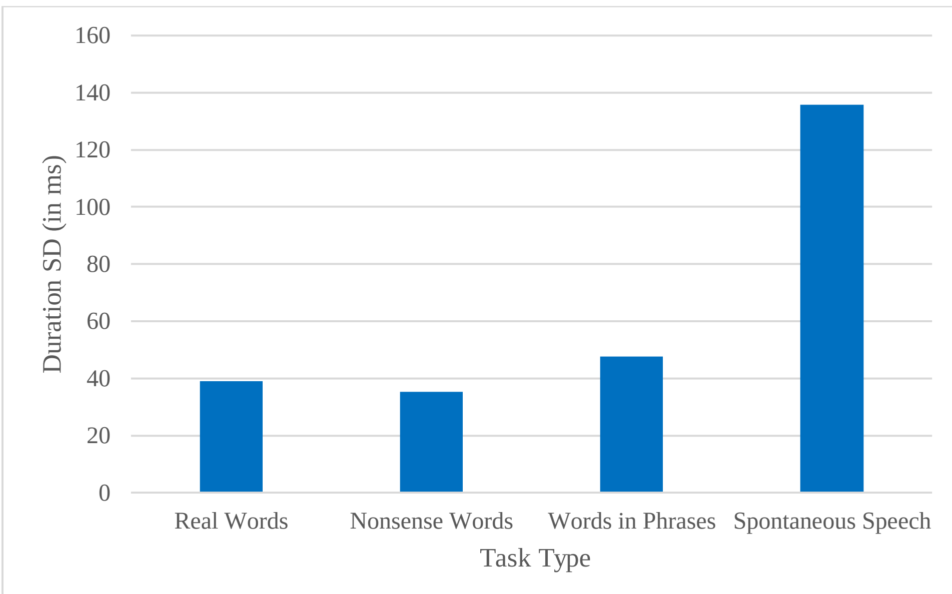

Figure 5. Duration standard deviation by task type for L2 speakers

Table 4

COG and duration measures by task type for L2 speakers

	COG		Duration	
Task Type	*Mean*	*SD*	*Mean*	*SD*
Real Words	4.24	0.26	260	39
Nonsense Words	3.68	0.40	229	35
Words in Phrases	4.22	0.53	162	48
Spontaneous Speech	4.60	0.53	179	136

Table 5

COG and duration measures by fricative and task type for L2 speakers

Task Type	Sound											
	[ç]				[x]				/ʃ/			
	COG		Duration		COG		Duration		COG		Duration	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
Real	4.93	0.26	254	46	3.27	0.40	243	45	5.47	0.17	301	33
Nons.	4.66	0.22	227	40	3.19	0.49	230	33	-	-	-	-
Phrase	4.77	0.50	167	62	3.63	0.54	157	40	-	-	-	-
Spont.	4.66	0.49	180	174	3.73	0.60	151	43	-	-	-	-

Subject.

The COG variability was shown to be significant across participants ($F(11, 133) = 2.177$, $p= 0.019$, $\eta^2_{partial} = 0.151$). Subjects 9 and 12 displayed the highest levels of variability (SD = 0.55 and 0.59, respectively), whereas subjects 3 and 8 displayed relatively low variability levels (SD = 0.26 and 0.24), as displayed in Figure 6. Duration did not vary significantly across subjects. When post hoc analysis of duration variability was completed by removing the spontaneous utterances of [ç], however, duration varied significantly across subjects ($F(11, 122) = 5.468$, $p < 0.001$, $\eta^2_{partial} = 0.326$). The COG and duration mean and standard deviation measures can be found in Table 6.

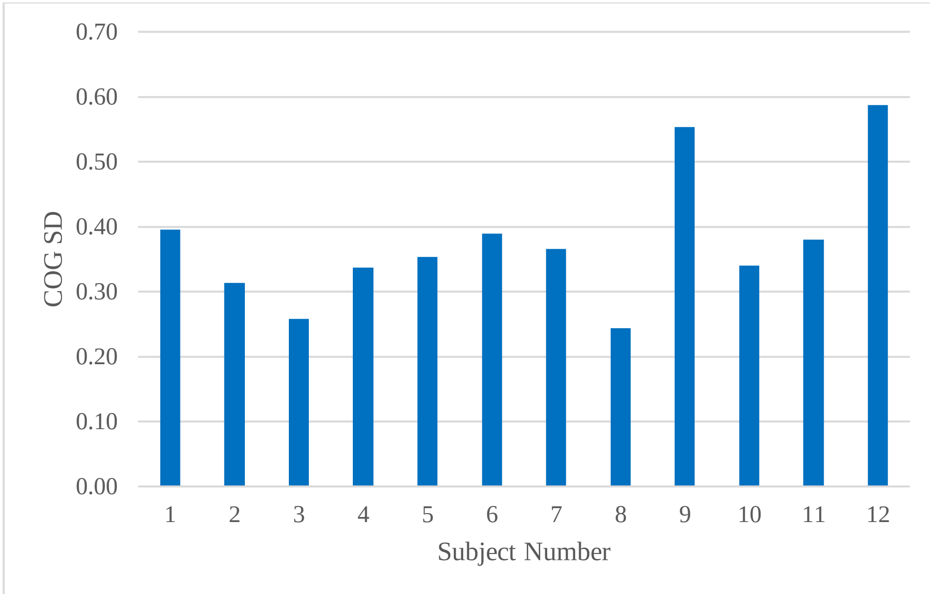

Figure 6. COG standard deviation by subject for L2 speakers

Table 6

COG and duration measures by subject for L2 speakers

	COG		Duration	
Subject	*Mean*	*SD*	*Mean*	*SD*
1	4.38	0.40	218	34
2	4.16	0.31	269	72
3	4.28	0.26	187	62
4	4.21	0.34	197	33
5	4.58	0.35	202	38
6	4.11	0.39	226	53
7	4.66	0.37	231	40
8	4.01	0.24	206	39
9	3.95	0.55	242	38
10	3.61	0.34	247	97
11	4.19	0.38	184	35
12	4.86	0.59	201	72

L1 vs L2 Speakers

The intra-speaker variability in the COG measure was not found to be significant as a factor of language. Figure 7 displays the overall COG variability levels for L1 and L2 speakers. The levels of variability were similar, with native German speakers displaying slightly higher variability as compared to L2 speakers. The COG variability was found to be significant across sound types as a function of language ($F(5, 189) = 3.383$, $p = 0.006$, $\eta^2_{partial} = 0.082$). As seen in Figure 8, L2 speakers demonstrated a generally higher variability rate for [ç] and [x], while L1 speakers demonstrated markedly more variability for /ʃ/. Examination of the data revealed an unusually high COG variability level for the [x] fricative in the HB position in German speakers; when this sound was removed from analysis, the COG variability as a factor of language grew closer to significance ($p = 0.150$) while sound types as a function of language remained significant, $F(4, 147) = 2.629$, $p = 0.037$, $\eta^2_{partial} = 0.067$).

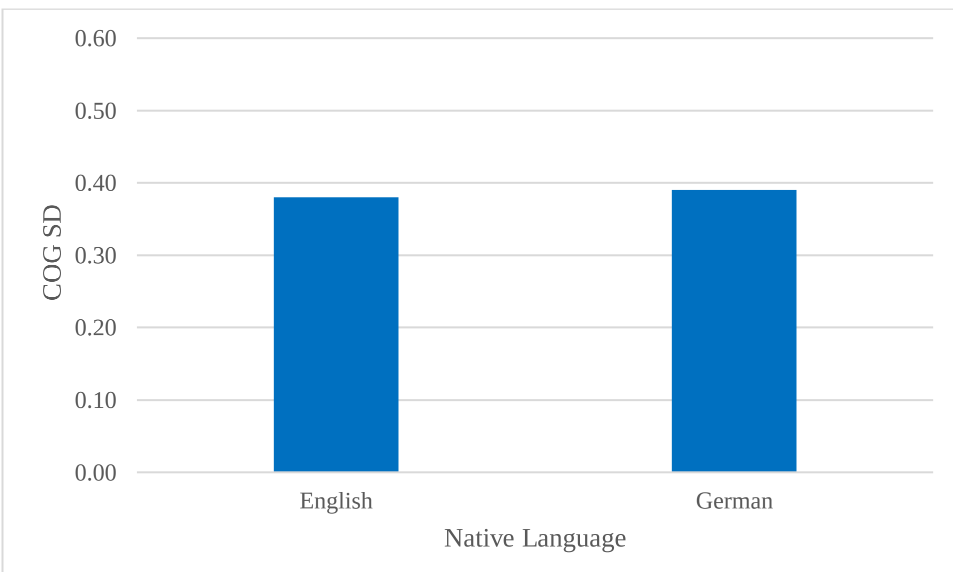

Figure 7. Overall COG standard deviation for L1 and L2 speakers

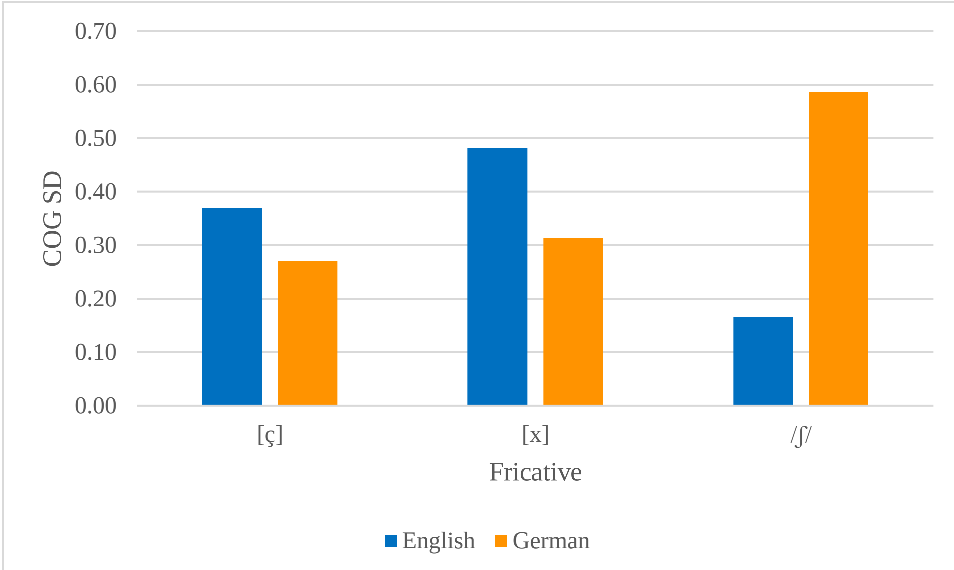

Figure 8. COG standard deviation by fricative for L1 and L2 speakers

Variability in duration across L1 and L2 was not significant, although on average, L1 speakers displayed more duration variability than L2 speakers (see Figure 9). Variability in duration by sound type was found to be significant ($F(16, 202) = 2.736$, $p = 0.020$, $\eta^2_{partial} = 0.061$), but not as a function of language. Despite this, it did appear that L1 variability for duration stayed relatively constant across each fricative (mean = 62 ms), whereas L2 speakers displayed a wide range of duration variability (see Figure 10). COG and duration values by fricative type and language are listed in Table 7.

Table 7

COG and duration measures by fricative for L1 and L2 speakers

	Sound											
	[ç]				[x]				/ʃ/			
Language	COG		Duration		COG		Duration		COG		Duration	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
English	4.73	0.37	198	81	3.33	0.48	221	39	5.47	0.17	301	33
German	5.00	0.27	346	61	2.36	0.59	331	60	4.96	0.31	378	64

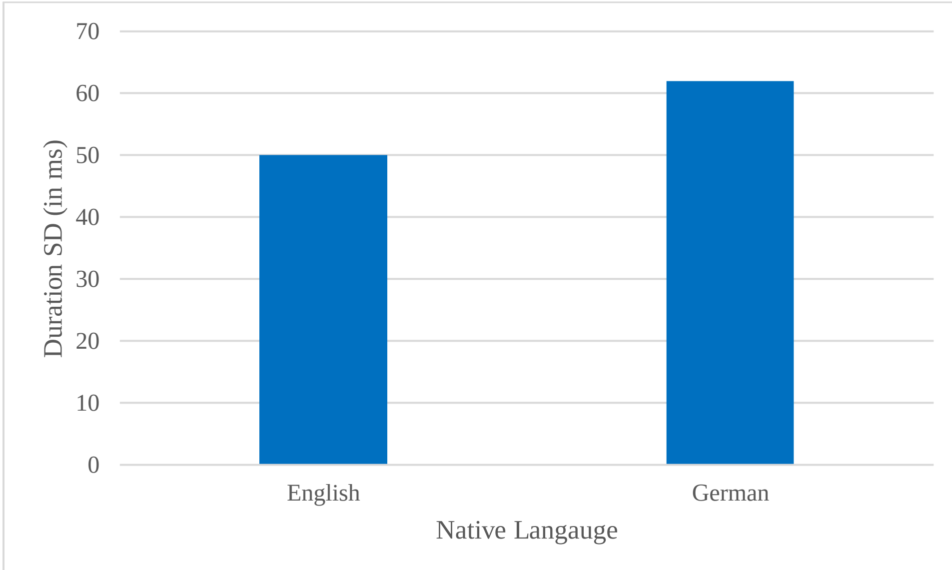

Figure 9. Overall duration standard deviation for L1 and L2 speakers

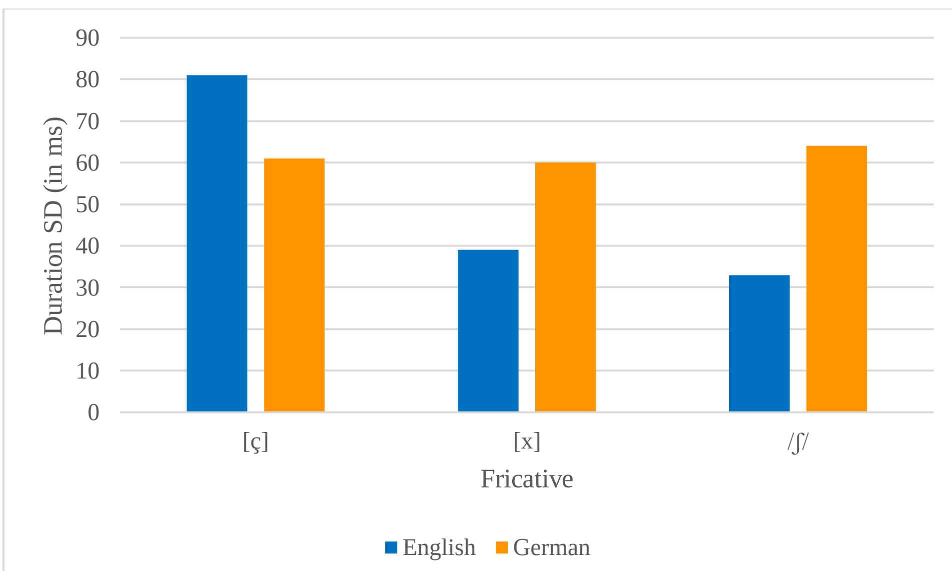

Figure 10. Duration standard deviation by fricative for L1 and L2 speakers

Discussion

The purpose of this study was to examine the intra-speaker variability of native German speakers and native English speakers learning German as a second language. When using COG and duration standard deviation as dependent variables, L2 speakers displayed significant COG variability across subjects, sound types, and task types. Additionally, duration variability was found to be significant in L2 speakers for sound and task type. The variability between L1 (German) and L2 (English) speakers was only significant across sound type.

Among English speakers learning German as a second language, several factors were found to have a significant effect on variability. Sound type in particular had a strong effect on both the COG and the duration standard deviation. This supports previous research done by Dromey et al. (2009) and McAuliffe et al. (2001), as the data presented in the current study shows that different sounds have differing levels of variability. Similarly, results provide further evidence that vowel context has an effect upon variability, as noted in Dagenais et al. (1994). The results of the current study were mixed as to which vowel context displayed the greatest variability; however, this could be due partially to the fact that certain vowel contexts were sampled far less frequently than others. There was a particular discrepancy in the number of tokens analyzed for the HB vowel context, due to its infrequent appearance in the participant's spontaneous speech.

Chen et al. (2007) demonstrated that individuals reduce variability in their sound productions as they grow older and therefore gain more experience in speaking a language. The current study supports Chen et al.'s finding because the L2 speakers showed lower variability in the phonemes with which they had greater familiarity. This relationship existed despite the fact that the familiar sounds were being produced in an unfamiliar language setting. This suggests that individuals learning a second language do not "start from scratch" when it comes to non-native sound acquisition, but rather build upon foundations of previously learned motor patterns. Further research in this area involving a broader range of phonemes could potentially reveal the extent to which L2 learners rely on established motor patterns when learning non-native sounds.

Cope (2018) noted that task type affected the electrode activation levels, duration, and the COG of L2 productions of German fricatives. The current study adds to this finding by noting that variability in the COG and duration also have a significant relationship with task type. Similar to what was noted by Cope, increasing complexity in task type resulted in increasing variability in the COG measurements. This is likely due to the increasing effects of coarticulation with increasing task type complexity. In tasks with a low degree of expected coarticulation, such as speaking real or nonsense words, the COG and duration stayed more consistent; tasks with higher levels of expected coarticulation, however, lead to greater variability in the COG and duration measurements.

Variability differences across task type must be considered in the context of several factors. The first is that the number of tokens varied widely across each task type. In real words, nonsense words, and words in phrases, the number of tokens was under the control of the experimenters. The number of tokens in spontaneous speech, however, was entirely dependent on the participant. Because of this, certain stimuli were much more common than others in the spontaneous speech samples. The HF [ç] fricative, or “ich”, is a German word meaning “I” and was therefore a very commonly used word in the personal narratives subjects were asked to relay for the spontaneous speech samples. Tokens of the HB context of [x], on the other hand, was so rare in the spontaneous speech samples that it was forced to be excluded from the analysis completely. Additionally, tokens of the /ʃ/ fricative were only available for analysis in the context of real words. In order to confirm the relationship of task type on variability levels, additional studies could be designed to address these discrepancies.

A second factor that may have affected task type results was the pragmatic context in which the spontaneous HF context of [ç] was used by speakers. As the L2 speakers delivered their spontaneous sample, it was common for subjects to use the word “ich” as a filler word, drawing out the word’s duration in order to give themselves time to think about the remainder of their sentence. In addition, it was not uncommon for L2 speakers to reduce the [ç] fricative to an English stop such as /k/, especially in spontaneous speech. The combination of these two factors, one of which lengthened and the other of which shortened the duration of [ç], may have led to the high average duration standard deviation for the [ç] fricative in spontaneous speech. When this factor was considered in the analysis, the results changed quite dramatically. It is therefore very likely that the high duration variability of spontaneous productions of [ç] presents a confounding variable in this study.

Dromey et al. (2009) noted that certain speakers tended to be more variable in their sound production than others. Results of the current study support this finding. Certain participants were highly variable in their COG and duration measures, while others were fairly consistent. It is interesting to note, however, that often those subjects that displayed greater variability in one measure (COG or duration) generally were not the same subjects who displayed high variability in the other measure. Subject 9, for example, displayed the highest duration variability but one of the lowest COG variability levels. This finding lends additional support to Butcher's (1989) declaration that "variability in itself is variable".

The between-subject analysis revealed no significant relationship between language and the COG variability. This contrasted the author's hypothesis that L2 learners would display greater linguapalatal variability than L1 speakers due to the unfamiliar nature of articulatory placement involved in L2 production. While overall the COG variability was essentially comparable between the two languages, COG variability for each sound was significantly different between the two languages. It is possible that this relationship is due to the specific sound types selected for analysis. The [ç] and [x] sounds are both non-native to English speakers, whereas the /ʃ/ sound exists in the English language. It is interesting to note that English speakers generally displayed higher COG variability for [ç] and [x] compared to German speakers, but much lower COG and duration variability for /ʃ/. This suggests that the English speakers in this study do indeed have increased variability for non-native sounds, but when presented with a native sound such as /ʃ/ will display comparable or even lower variability than German speakers.

Duration variability was not found to be different between languages or across sound types, although L1 speaker did demonstrate an overall higher level of variability as compared to L2 speakers. This could potentially be due to L2 speakers taking greater care with their articulation than L1 speakers, leading to reduced co-articulation and therefore a greater uniformity in the duration of their speech sounds. It could also be that since L2 speakers have to dedicate much more attention to semantics and syntax, they have fewer cognitive resources to dedicate to varying prosodic factors such as speech rate. This idea is supported by the findings of Nissen et al. (2007), which noted that even with a relatively strong grasp of a second language, L2 speakers have reduced tongue speed and speech rate.

Many factors make it difficult to assess whether the L2 speakers in this study are more or less variable than the L1 speakers. Without normative data about German linguapalatal contact variability, assumptions cannot be made about whether L2 speaker variability is higher or lower than the average variability of German speakers; it could be that German is inherently a more variable language as compared to English. Additionally, the sounds chosen for this analysis could have skewed the results; Isaacson (2015) noted that the [x] fricative is produced so far posteriorly in native German speakers that the EPG unit may not be able to adequately capture linguapalatal contact information. This fact might have contributed to the high COG variability of the L1 production of [x] in the HB context as compared to all other CV contexts. One must also consider that the task types, number of tokens, and number of subjects varied between the two data-sets, and each of these factors could have had confounding effects upon the data presented in this study.

The scope of this study was limited in that it compared only two languages, had relatively few subjects, examined only a small selection of speech sounds, and had no normative data to which results could be compared. Future research could therefore be conducted to expand upon any of these factors. Examining variability between L1 and L2 productions across several languages, for instance, could generate a clearer picture of whether variability is affected by an individual's language learning status.

Although further research is needed, this study presents a solid foundation for gaining an understanding of variability in L2 language production. The study added further substantiation to previous research demonstrating the effect of sound type, task type, and individual differences upon variability in speech sound production. While these results did not indicate a relationship between language status and variability levels, there was evidence to suggest that L2 speakers demonstrate lower variability in speech sounds with which they are already familiar. The quantification and examination of variability can help lead to a greater understanding of the normal speech process. The results of the current study also present an important step towards gaining an understanding of how individuals learn and adapt their motor speech patterns when learning an L2. It is hoped that this added understanding will improve teaching and learning strategies for L2 acquisition and other related fields.

REFERENCES

Bright, A. L. (1999). *The palatometer as an instrument for accent reduction therapy with three native ESL Spanish speakers*. Brigham Young University, Provo, UT.

Butcher, A. (1989). Measuring coarticulation and variability in tongue contact patterns. *Clinical Linguistics & Phonetics*, *3*(1), 39–47.

Cheng, H. Y., Murdoch, B. E., Goozée, J. V., & Scott, D. (2007). Electropalatographic assessment of tongue-to-palate contact patterns and variability in children, adolescents, and adults. *Journal of Speech Language and Hearing Research*, *50*(2), 375.

Cope, E. (2018). *The influence of task type on speech production by second language learners of German: An electropalatographic study*. (Master's thesis). Retrieved from http://etd.byu.edu/

Dagenais, P. A., Lorendo, L. C., & McCutcheon, M. J. (1994). A study of voicing and context effects upon consonant linguapalatal contact patterns. *Journal of Phonetics*, *22*(3), 225–238.

Dromey, C., & Sanders, M. (2009). Intra-speaker variability in palatometric measures of consonant articulation. *Journal of Communication Disorders*, *42*(6), 397–407.

Fletcher, S. G. (1992). *Articulation: A physiological approach.* San Diego, CA: Singular Publishing Group, Inc.

Gibbon, F. E., Hardcastle, W. J., & Suzuki, H. (1991). An electropalatographic study of the (r), (l) distinction for Japanese learners of English. *Computer Assisted Language Learning*, *4*(3), 153–171.

Isaacson, L. D. (2015). *Establishing normative data for contact patterns of fricative production by native German speakers: An electropalatography study*. (Master's thesis). Retrieved from http://etd.byu.edu/

Lester, K. E. (2017). *Quantifying lingua-palatal contact patterns of fricative productions by non-native students enrolled in a university German language course: An electropalatography study*. (Master's thesis). Retrieved from http://etd.byu.edu/

McAuliffe, M. J., Ward, E. C., & Murdoch, B. E. (2001). Tongue-to-palate contact patterns and variability of four English consonants in an /i/ vowel environment. *Logopedics Phoniatrics Vocology*, *26*(4), 165–178.

Nissen, S. L., Dromey, C., & Wheeler, C. (2007). First and second language tongue movements in Spanish and Korean bilingual speakers. *Phonetica*, *64*(4), 201–216.

Schmidt, A. M. (2012). Effects of EPG treatment for English consonant contrasts on L2 perception and production. *Clinical Linguistics & Phonetics*, *26*(11–12), 909–925.

Schmidt, A. M., & Beamer, J. (1998). Electropalatography treatment for training Thai speakers of English. *Clinical Linguistics & Phonetics*, *12*(5), 389–403.