Kruger, F. (2018). Processing indexical and dialectal variation in a second language. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. 129-140). Ames, IA: Iowa State University.

PROCESSING INDEXICAL AND DIALECTAL VARIATION IN A SECOND LANGUAGE

Franziska Kruger, Indiana University

To improve second-language (L2) learners' listening skills, it has been suggested to enhance classroom-input with variation (e.g. more speaker voices, different dialects) to simulate a realistic linguistic landscape in the classroom. However, previous studies reported that L2-learners struggle to distinguish dialects and voices, even at advanced levels. This study explored L2-classroom learners' ability to group words based on indexical variation (speaker voice and dialect). Twenty-seven learners of German (14 beginners, 13 intermediates) and five native speakers classified thirty tokens based on their perceived similarity of voice and dialect in a free classification task. All participants distinguished the stimuli to some degree, but classification accuracy for beginners was significantly less accurate than for native speakers. Intermediates presented with large variation, but accuracy did not differ significantly from native speakers. All groups relied on the same acoustic cues, but their perceptual spaces reveal that NS were more successful than both learner groups at using those cues to differentiate tokens. The findings suggest that L2-classroom learners process indexical variation less efficiently and that word-familiarity might influence their performance. Increasing input variability in classrooms without taking these observations into account could make listening tasks more difficult and hinder learning- and listening-skill development.

INTRODUCTION

The perceptual system for speech processing adapts very efficiently to speaking style, rates, and dialects (Pisoni, 1997), and studies on language variation find that listeners pay close attention to indexical information that helps them interpret the variation inherent in the speech stream (Clopper & Bradlow, 2009). *Indexical information* is speaker-specific information about their region of origin (dialectal variation), as well as idiosyncrasies (e.g. gender, voice quality, Labov, 1972). It is linked to linguistic information and facilitates word recognition and processing even in difficult listening conditions (Cooke, García Lecumberri, & Barker, 2008).

The perceptual system of second-language (L2) learners is reportedly less efficient and less flexible at processing linguistic—e.g. phonetic-phonological—information (Baker & Trofimovich, 2005), as well as indexical information in the L2 (Clopper & Bradlow, 2009; Sullivan & Schlichting, 2000). To date, few studies have explored how L2 classroom learners process indexical information and whether idiosyncratic or dialectal variation is salient enough to be noticed and interpreted accurately during the initial stages of the L2-acquisition process. This study explores L2-German learners' ability to classify words, based on two specific indexical properties (speaker voice and dialect) in relation to their proficiency level, as well as their linguistic background.

LITERATURE REVIEW

The L1-perceptual system is highly attuned to linguistic information, and previous research has shown a strong familiarity effect for standard and dialectal varieties (Clopper & Pisoni, 2004) as well as for voice identification (Goggin, Thompson, Strube, & Simental, 1991). Listeners with exposure to various dialects perform more accurately on categorization tasks than dialect-naïve listeners, i.e. those who have lived in one only region (Pisoni & Clopper, 2004). Nevertheless, naïve listeners do perceive acoustic-phonetic features relevant to distinguish regional varieties and perform above chance levels on categorization tasks (Clopper, Levi, & Pisoni, 2006; Clopper & Pisoni, 2004). For L2 learners, research focuses mainly on their acquisition of linguistic properties, often confirming that learners struggle to acquire the L2 phonological system (Baker & Trofimovich, 2005). These difficulties are mirrored in learners' performance when processing indexical information in their L2 (Eisenstein, 1982; Eisenstein & Verdi, 1985; Köster & Schiller, 1997; Köster, Schiller, & Künzel, 1995).

Clopper and Bradlow (2009) found that native speakers (NS) of American English (AE) were more accurate at sorting talkers into US regions than non-native listeners from various L1 backgrounds. Their L2 participants had been in the U.S. less than 12 months, which may not have sufficed to improve their processing skills for L2-indexical information. Eisenstein (1982) and Eisenstein and Verdi (1985) found varying degrees of accuracy for L2 learners. Beginning and intermediate learners with short and longer length of residence (LOR) in New York were less accurate than NSs when asked to distinguish AE dialects, but advanced learners' performance did not differ from the NSs. It remains unclear if proficiency or amount of input influenced learner performance; the authors provided little information on learner LOR and coded proficiency levels based on course placement at colleges. While these studies indicate that learners in immersion settings can improve, no studies have explored whether they can process dialectal variation during the initial stages of L2 learning and from classroom exposure.

Studies on voice identification have reported a native-language bias, where listeners are more accurate at discriminating voices that speak their L1 as opposed to voices speaking an unfamiliar language (Goggin et al., 1991; Köster et al., 1995). Winters, Levi, and Pisoni (2008) found support for this L1 bias, while Wester (2012) observed that language familiarity is not relevant when listeners are asked to determine if two sentences were spoken by the same voice.

Studies with L2-learners indicate varying degrees of identification accuracy, but previous observations may have had several confounding factors. Köster et al. (1995) examined voice recognition in L2-learner groups with varying degrees of German proficiency. Participants had been in Germany for several months at the time of testing and were more accurate at identifying a German voice in a German-voice lineup than NS-English listeners without German knowledge. German NSs' and learner performance did not differ. Concluding that L2 proficiency does not affect voice identification, the authors revised their observations when a follow-up with L1-Spanish and L1-Chinese learners of German revealed both groups were significantly less accurate than the learners and NSs from the previous study (Köster & Schiller, 1997). The Spanish and Chinese groups had studied exclusively in Spain and China, without any immersion experience, and received German input in a classroom setting. The difference of immersion vs. classroom learning may thus have been a confound in those studies. Sullivan & Schlichting (2000) investigated L1-English learners of Swedish specifically in a classroom setting. First-year learners were more accurate at identifying a familiar voice in a lineup than a control group with no knowledge of Swedish, but accuracy among advanced learners was not higher than 1st-year learners'. Advanced learners who had been abroad were also no more accurate than the lower-level learners, suggesting neither immersion nor higher proficiency are a relevant factor for improvement on voice identification. Here, task design may have also affected learner performance; the voice lineup was created using a voice imitator, deliberately increasing confusability.

THE PRESENT STUDY

In the above studies, L2 learners were generally less efficient at processing indexical cues. While some observed a learning curve for L2 learners, others found no indication of improvement. It is unclear how learner behavior in those studies was related to task effects (Sullivan & Schlichting, 2000), L2 immersion (Köster et al., 1995; Köster & Schiller, 1997) or L2 proficiency (Sullivan & Schlichting, 2000; Eisenstein, 1982). Those aspects are often underreported, limiting the conclusions to be drawn from previous investigations. If indexical information is found to be salient, learners might benefit from increased language variety and speaker variation in classroom material as previously suggested (Iverson, Hazan, & Bannister, 2005). If, however, learners learn to successfully interpret indexical information only after a certain amount of L2 exposure, classroom material should accommodate learners' capabilities.

The present study examined L2 learners' ability to distinguish dialect forms of German and German-speaking voices. In a cross-sectional investigation, L2 learners and German NSs completed a free classification task (FC). For learners, accuracy was predicted to increase with increasing L2 proficiency. Based on previous research, all groups were predicted to employ indexical cues to some degree (Clopper et al., 2006). However, for learners of German, less accurate processing skills were expected when compared to NS who have the L1 and familiarity advantage (Pisoni & Clopper, 2004; Winters et al., 2008). L2 learners also completed a language-background questionnaire, and a proficiency test (C-test) to allow for a more refined interpretation of the results. The results of this study contribute to our understanding of L2 learners' perception and processing of indexical information.

METHODOLOGY

Listeners

L1-American-English learners of German (*N* = 27) and German NS (*N* = 5) participated in this study. NS were recruited from regions where the investigated dialects were not spoken. L2 learners were recruited in German language classes at a Midwestern US university. Their parents were native speakers of English, but participants' residential history varied slightly. Eighteen had grown up in the Midwest, nine had moved to the Midwest region. All learners confirmed that their speech did not have salient dialect-specific characteristics. No participant reported a history of hearing or speech disorders, and all passed a hearing screening.

Talkers and Dialects

Talkers were five female German NS. Talker BD (age in years = 26) was from Southwest-Germany with Swabian as native dialect and had lived only in that region of Germany. The remaining four talkers had grown up and lived in the state of Saxony with Upper Saxon as their native dialect. Three (KN = 29, CX = 29, IK = 65) had never left the region for an extended period, while one talker (ES = 76) had permanently moved to the state of Brandenburg at age 25.

Stimulus Materials

Thirty different German words were chosen as stimuli. Talkers read a list of forty-six words once in their native dialect, and once in Standard German. Recordings were made with a Zoom H2n recorder in a quiet environment. All recordings were spliced in Praat and saved as individual sound files. For each speaker, three dialect (D) and three Standard German (Std) tokens were chosen. Swabian and Upper Saxon exhibit several phonetic characteristics that differ from Standard German (See appendix). Both merge aspirated and unaspirated German stops in syllable-initial context. Furthermore, word-final, unstressed syllables ending in <-r> are produced with a near-open, unrounded central vowel [ɐ] in Standard, but pharyngealized [oˤ] in both dialects (Ud Dowla Khan & Weise, 2013; Russ, 1990). Standard-German stimuli were judged by the author not to exhibit any salient features of variation and to conform to the phonetic forms described in the literature (Mangold, 2000).

Procedure

For the FC task, participants saw a Power Point slide with a 16x16 grid on a computer screen. Next to the grid were 30 consecutively-numbered boxes, connected to a sound file (one per stimulus). Listeners were instructed to listen to each file and drag the items onto the grid grouping them according to talker dialect and talker voice. To ensure participants focused on both target forms (voice and dialect), they were told that the same talker might speak different dialects and if they thought that was the case, they were encouraged to separate their groups further. Participants could listen to and rearrange the objects as often as they liked. Listeners then completed a background questionnaire in Qualtrics including a word-familiarity task, where they listened to the thirty stimuli from the FC task and gave the German spelling and English translation. Finally, a 30-minute, timed C-test was administered online. The task randomly assigned listeners five short paragraphs with blanks that had to be filled out.

ANALYSIS AND RESULTS

Based on the C-test, learners were assigned to the beginner (BEG, $N = 14$) or intermediate (INT, $N = 13$) group. The test automatically calculated learners' scores and assigned a proficiency level based on the *Common European Framework of Reference for Languages*. Out of 100, beginners (BEG) scored between 8-40 and Intermediate (INT) learners scored 43-80. Proficiency scores were strongly correlated with learners' course enrollment ($r = .97$, $p < .01$). For word familiarity, the Scheffé post hoc test revealed the beginners knew the task items significantly less often (Std: $F(3, 34) = 77.16$, $p < .01$, D: $F(3, 34) = 53.93$, $p < .01$) than the intermediates, which in turn were significantly less accurate than the NS (Std: $F(3, 34) = 77.20$, $p < .01$; D: $F(3, 34) = 53.93$, $p < .01$). On the background questionnaire only one BEG listener indicated having spent 2 months in Germany, while seven INT learners had been to Austria or Germany for an average of 3.7 months (range 1.5-12 months). No participant had been to a region where the dialects investigated here were spoken.

Listeners' Perceptual Space

ALSCAL. To investigate the perceived similarities and dissimilarities of the FC stimuli, a multidimensional scaling (MDS) analysis was conducted using SPSS 24's ALSCAL function. A 30x30 pairwise dissimilarity matrix was created for each group. Each cell indicated how many times listeners did not group an item together with another item. Based on the elbow in the stress plot, a two-dimensional solution was deemed the best fit for all groups (BEG $r > .79$, stress = .27, INT $r > .83$, stress = .25, NS $r > .83$, stress = .25), suggesting that listeners perceived items as similar or dissimilar along two different dimensions (Figures 1, 2, 3).

Acoustic Analysis. To determine which perceptual dimensions were most salient for listeners, the fundamental frequency (f0) and the intensity (dB) of all stimuli were measured in Praat. Measurements were taken at 0%, 25%, 50%, and 75% of the stressed vowel in each stimulus. Dialect-specific dimensions, i.e. number of phonological processes in a D item as opposed to its StD form were tallied (0-3), and stimuli with frequently occurring processes were assigned a dichotomous value for absence or presence of the respective process (1=present, 0=absent). Speaker age was based on their biological age in years (1 = 25-35, 2 = 65, 3 = 76).

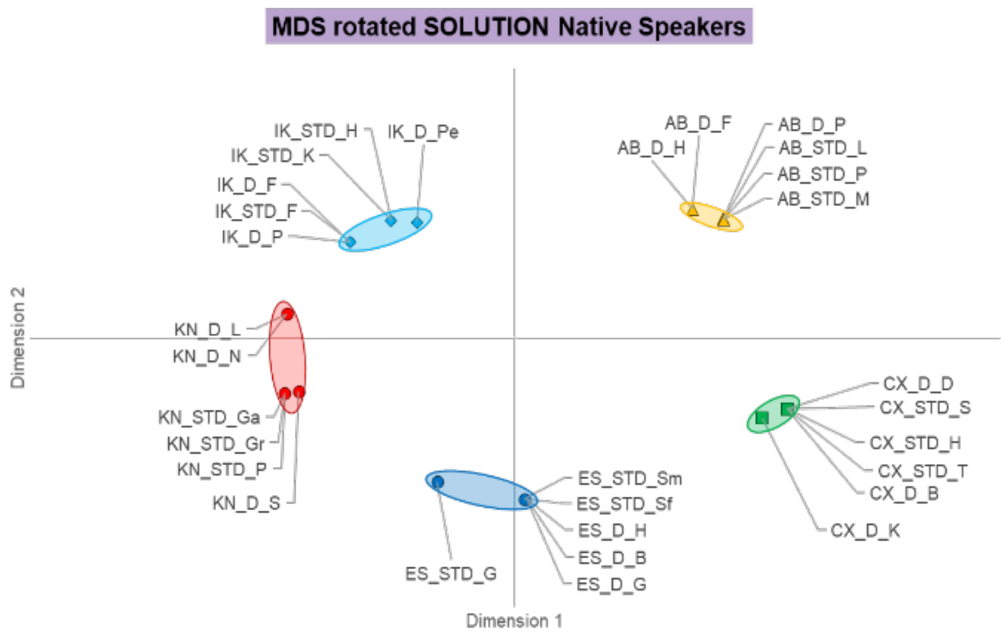

Figure 1. High variability training in the lab and in the language classroom.

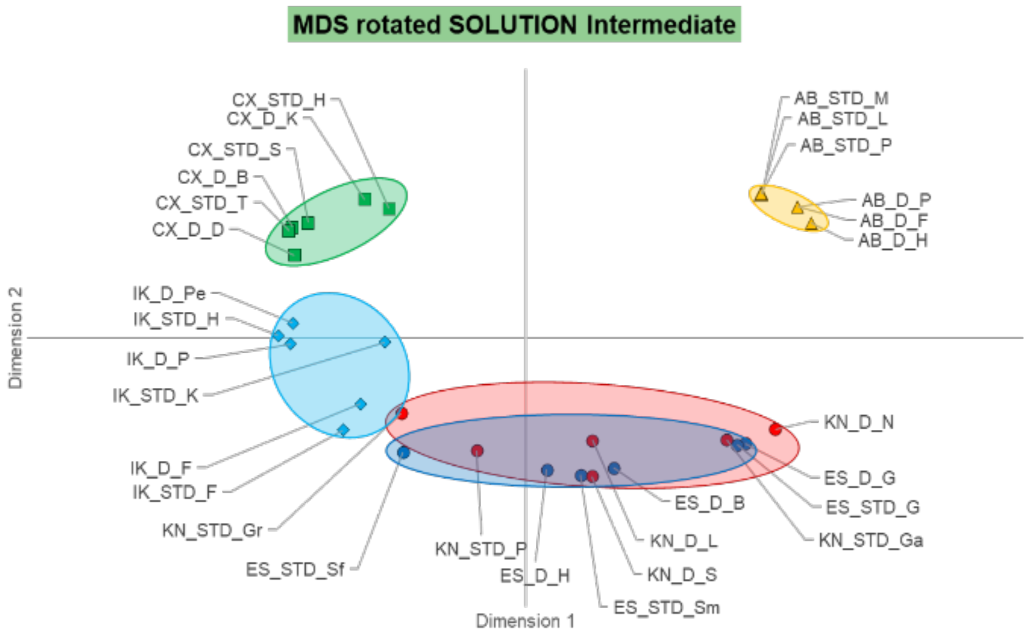

Figure 2. MDS rotated solution illustrating the perceptual space for intermediate level learners of German.

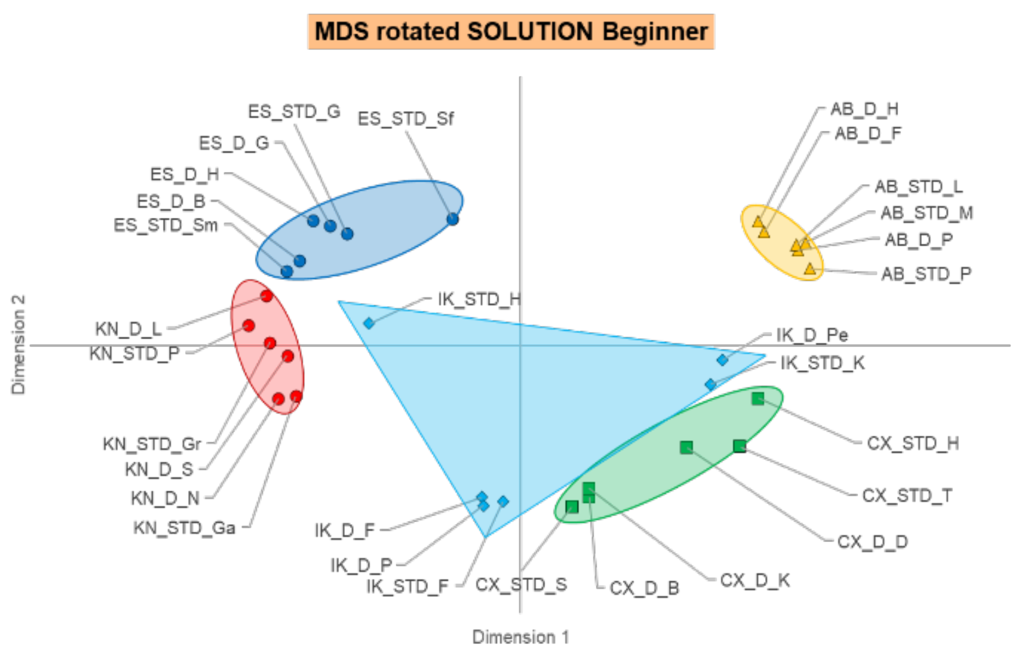

Figure 3. MDS rotated solution illustrating the perceptual space for beginning level learners of German.

Narrow groupings in the MDS solutions indicate that items were consistently grouped together, while wide-spread groupings indicate confusion about item group-membership. The nearly equidistant, narrow groupings in the NS-MDS solution reveal that listeners perceived items in other groups as very different. The wide-spread groupings in the INT- and BEG-MDS solutions show more overlap between stimuli, i.e. a larger number of stimuli was perceived as similar by learners where the NS made a clear distinction. After rotating the MDS solutions, a correlational analysis for acoustic or dialect-specific cues in Excel revealed the most salient dimension for all groups was *voice* (NS $r = -.93$, BEG $r = .81$, INT $r = -.80$) while neither group based their decisions on dialect-specific information (Table 1 below).

Table 1

Pearson correlation coefficient (R^2) for dialect-specific and acoustic cues calculated for all groups' rotated MDS solutions

Variable	**BEG**		**INT**		**NS**	
	Dim1	Dim2	Dim1	Dim2	Dim1	Dim2
dialect-specific						
variation (STD vs. D)	.07	-.01	.10	-.03	.0	-.04
acoustic measurements						
voice	**.81**	.27	-.27	**-.80**	**-.93**	.16
f0	**-.44**	.06	.05	.21	**.40**	.02
dB	-.27	**.55**	**.53**	.04	.35	**.44**
age	.19	**.50**	-.06	**-.61**	-.27	**.44**

For NS and BEG, *voice* was best interpreted as f0, while the INT group relied on speaker *age*. In the second dimension each group relied on either *age*, dB or a combination thereof. A series of linear regressions was carried out in SPSS to investigate if f0 or dB could predict talker or talker age. For f0, a significant effect was found for talker KN whose f0 values were significantly lower than the other talkers' ($\mu=165.6$, $r = .59$, $F(4,25) = 3.4$, $p < .05$). Therefore, a certain f0 range corresponded strongly to KN tokens. Furthermore, dB was a predictor for the middle-aged talker (IK), the quietest ($\mu_{dB}=48.2$, $r = .78$, $F(2,27) = 20.74$, $p < .01$) and for the oldest (ES), the loudest talker ($\mu_{dB}=60.5$, $r = .78$, $F(2,27) = 20.74$, $p < .05$).

Listeners' Accuracy

The FC task also provides for objective accuracy measures, thus a difference score (DiffS) was calculated for each listener:

1) DiffS = (percent correct pairings) – (percent error pairings)

Percent correct and incorrect pairings are similar to 'hits' and 'false alarms' in signal detection theory (Clopper & Bradlow, 2009). Stimuli pairs were counted as correct pairing if both items were spoken by the same talker and in the same variant (e.g. both D or both StD). Following Clopper & Bradlow (2009), the proportion of correct pairings was calculated out of the total possible number of correct talker pairings. Similarly, the proportion of mispairings was calculated out of the total possible number of incorrect pairings. A pair was counted as error if it had two different talkers, or the same talker but different variants grouped together. In Table 2, a low DiffS indicates that listeners frequently grouped items incorrectly, while a higher DiffS reflects more correct pairings. The maximal score on this task was 30.

Table 2

Between group comparisons for voice and dialect difference scores.

	BEG N = 14		INT N = 13		NS N = 5	
	M	SD	M	SD	M	SD
DiffS	-28.35**	20.05	-17.38	20.79	5	12.59

Note: *p < .05, **p < .01

An analysis of variance in SPSS revealed that the BEGs' DiffS was significantly lower than the NS groups' ($F(4, 34) = 3.59$, $p < .01$), while there was no significant difference in group means between the INT and NS groups (see Figure 4). In a correlational analysis, word familiarity was found to correlate with DiffS ($r = .73$), while LOR in the immersion setting did not ($r = .4$).

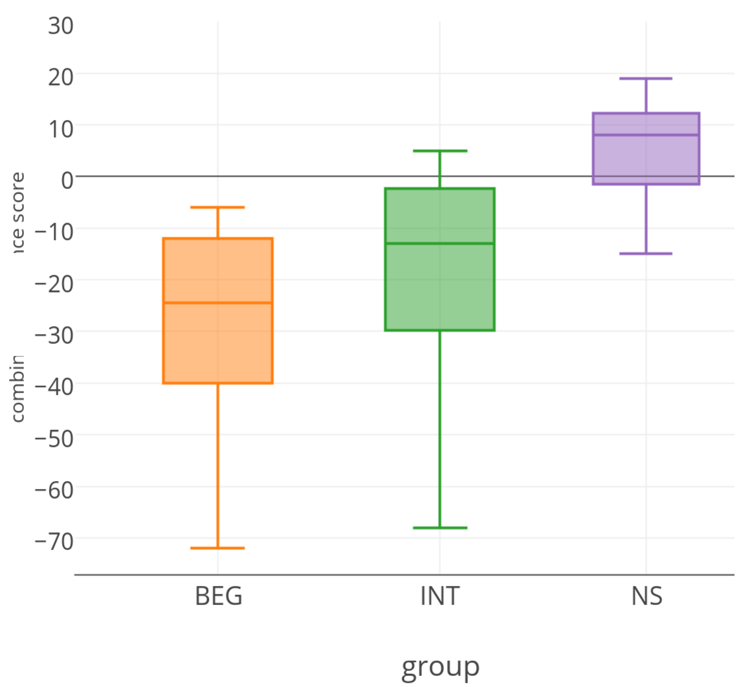

Figure 4. Boxplot for difference scores. This figure illustrates average DiffS for all groups with standard error bars.

DISCUSSION

This study investigated L2 learners' abilities to classify stimuli based on two indexical cues: talker voice and talker dialect. The classification results indicate that all groups could discriminate the five talkers. The NS performed more accurately than the learner groups, as expected from previous studies that reported an L1 advantage (Goggin et al., 1991; Winters et al., 2008). Classification accuracy, determined by difference score (DiffS), differed significantly for BEG and NS listeners. INT had a higher DiffS, but as much variation as BEG. Their MDS solution also revealed an overlap for talkers which was not observed in the NS group. Nevertheless, the INTs were not significantly different from NSs indicating a general trend for improvement with learner proficiency. However, this non-significance might be due to low statistical power from the small sample sizes. Larger samples from these populations are needed to draw more generalized conclusions.

The perceptual maps of the three groups revealed that listeners determined group membership based on f0 and dB, rather than on dialectal-specific information. All groups also relied on age. As the analysis revealed, age could be predicted based on dB level, where the middle-aged talker (IK) would have been perceived as the quietest speaker and the oldest talker (ES) as the loudest. f0 predicted which tokens were produced by KN. MDS solutions and DiffSs indicate that the groups used those cues with differing rates of success. BEG learners relied on f0 and reliably distinguished KN from the other speakers. They also used dB to classify tokens, which should have allowed them to identify IK from the other speakers. Yet, she was grouped with ES and CX and thus dispersed across the BEG perceptual map (Fig 3). The INT listeners relied only on dB, ignoring f0, which could explain KN and ES overlapping in their perceptual space (Fig 2).

The NS relied on the same cues as the learners yet were more successful at classifying stimuli. Previous studies that have investigated perceived correlates of age found that f0 and intensity are relevant, but other characteristics, e.g. laryngeal tension, air loss, and preciseness of articulation might be more important to distinguish voices from one another (Ryan & Burk, 1974). Thus, NS might have focused on a cluster of variables that the learners did not use. Since the analysis was limited to measurable acoustic cues, it may not have captured all voice-related features that listeners may have perceived as relevant. Thus, future investigations could benefit from a more precise voice assessment and post-task inquiries asking listeners to describe characteristics in the stimuli they deemed relevant for their decisions.

Despite more accurate groupings by the NS group, their DiffS did not approach ceiling levels and listeners only inconsistently grouped dialect vs. standard token. On the one hand, this might be because no NS had had significant exposure to the relevant dialects, a criterion found to increase dialect discrimination (Pisoni, 2004a). In that case, the NS results would support previous studies that found lower performance for unfamiliar language variation. On the other hand, the study design required listeners to focus on both voice and dialect for their decisions. The instructions for participants to focus on voice and dialect at the same time conflate several acoustic and perceived dimensions and complicate the interpretation of listeners' perceptual space. Future investigations should take care to separate them to obtain discrete perceptual spaces for voice and dialect classifications and allow for a more controlled investigation of acoustic dimensions that are salient to the NS and L2 learners.

While LOR in the L2 environment was not a good predictor for classification accuracy, word familiarity was shown to have influenced learners' results. BEG knew significantly fewer words than INT who recognized significantly fewer words than NS. Stable lexical representations connected to a specific phonological form may have allowed NS listeners to compare the stimuli with the familiar form and make more consistent decisions.

Despite some limitations, this study sheds some light on classroom learner's abilities to process socio-indexical cues. The results indicate a learning curve for learners, illustrating that indexical processing might increase with proficiency. The interlanguage seems to also require stable representations for learners to reliably process indexical cues in the L2. Suggestions to increase the language input with highly variable stimuli, i.e. different voices, L2 variations, should take into consideration that learners' might initially not be equipped with the tools to process all types of information with equal precision and that exposure and practice could be necessary to allocate resources towards indexical information.

ABOUT THE AUTHOR

The author is a PhD candidate at Indiana University, Bloomington. Her research interest lies in second language acquisition and therein primarily in second language phonology. Previous and current works examine the perception and production of non-native segmental categories by second language learners and naïve listeners, as well as the establishing of phonological forms for lexical representations in second language learners. Her research also branches out into sociolinguistics and dialectology. Focusing on variation in German-speaking regions, she investigates the processing of dialectal variants by second language learners and second dialect acquirers, and she studies the potentials of incorporating indexical variation into the second-language classroom.

Correspondence concerning this article should be addressed to Franziska Kruger, Department of Germanic Studies, Global and International Studies Building 3103, 355 North Jordan Ave. Bloomington, IN 47405-1105, telephone: 812-855-1553, email: fkruger@indiana.edu

REFERENCES

Baker, W., & Trofimovich, P. (2005). Interaction of native- and second-language vowel system(s) in early and late bilinguals. *Language and Speech, 48*, 1-27.

Clopper, C. G., & Bradlow, A. R. (2008). Perception of dialect variation in noise: Intelligibility and classification. *Language and Speech, 51*, 175–198.

Clopper, C. G. & Bradlow, A. R. (2009). Free Classification of American English dialects by native and non-native listeners. *Journal of Phonetics, 37*, 436-451.

Clopper, C. G., Levi, S. V., & Pisoni, D. B. (2006). Perceptual similarity of regional dialects of American English. *JASA, 119,* 566-574.

Clopper, C. G. & Pisoni, D. B. (2004). Homebodies and army brats: Some effects of early linguistic experience and residential history on dialect categorization. *Language Variation and Change, 16*, 31-48.

Cooke, M., García Lecumberri, M. L., Barker, J. (2008). The foreign language cocktail party problem: Energetic and informational masking effects in non-native speech perception. *JASA, 123*, 414-427.

Eisenstein, M. (1982). A study of social variation in adult second language acquisition. *Language Learning, 32*, 367–391.

Eisenstein, M., & Verdi, G. (1985). The intelligibility of social dialects for working-class adult learners of English. *Language Learning, 35*, 287–298.

Goggin, J.P., Thompson, C.P., Strube, G., & Simental, L.R. (1991). The role of language familiarity in voice identification. *Memory & Cognition*, *19*, 448-458.

Iverson, P., Hazan, V., & Bannister, K. (2005). Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/-/l/ to Japanese adults. *JASA, 118*, 3267-3278.

Köster, O., & Schiller, N. O. (1997). Different influences of the native language of a listener on speaker recognition. *Forensic Linguistics, 4*, 18-28.

Köster, O., Schiller, N. O., & Künzel, H. J. (1995). The influence of native-language background on speaker recognition. In K. Elenius, & P. Branderud (Eds.), *Proceedings of the Thirteenth International Congress of Phonetic Sciences*, *4*, (pp. 306-309). Stockholm.

Labov, W. (1972). *Sociolinguistic patterns.* Philadelphia: University of Pennsylvania Press.

Mangold, M. (2000). Duden – Das Ausssprachewörterbuch. Mannheim: Brockhaus.

Pisoni, D. B. (1997). Some thoughts on “Normalization” in speech perception. In K. Johnson, & J. H. Mullennix. *Talker variability in speech processing* (pp. 9-66). San Diego: Academic Press, Inc.

Russ, C. V. J., (1990). Swabian. In: C. V. J Russ (Ed.), *The dialects of Modern German. A linguistics survey* (pp. 337-363). Routledge: London.

Ryan, W. J., & Burk, K. W. (1974). Perceptual and acoustic correlates of aging in the speech of males. *Journal of Communication Disorder, 7*, 181-192.

Sullivan, K. P. H., & Schlichting, F. (2000). Speaker discrimination in a foreign language: first language environment, second language learners. *Forensic Linguistics, 7*(1), 95-111.

Ud Dowla Khan, S., & Weise, C. (2013). Upper Saxon (Chemnitz dialect). *Journal of the International Phonetic Association, 43*, 231–241, doi: 10.1017/S0025100313000145.

Wester, M. (2012). Talker discrimination across languages. *Speech Communication, 54*, 781-790.

Winters, S. J., Levi, S. V., & Pisoni, D. B. (2008). Identification and discrimination of bilingual talkers across languages. *JASA*, *123*, 4525-4538.

Appendix Table A1 - FC- and familiarity- task token in Standard and dialect

speaker	item	standard phonetic form	dialect phonetic form
BD	feather	[fe:tɐ]	[fæ:toˤ]
BD	home	[haɪmatʰ]	[hɔɪmətʰ]
BD	party	[pʰa:tʰi]	[pa:ti]
CX	cabin	[bu:də]	[bʉ:də]
CX	poet	[tiçtʰɐ]	[tiʃtoˤ]
CX	pillow	[kʰɪsn̩]	[kɪzn̩]
ES	blister	[pla:zə]	[pla:sə]
ES	belt	[gʏʁtʰl̩]	[gʏʁtl̩]
ES	hunger	[hʊŋɐ]	[hʊŋoˤ]
IK	fan	[fɛçɐ]	[fɛʃoˤ]
IK	whip	[pʰa͜ɪt͡ʃə]	[pa͜ɪdʃə]
IK	powder	[pʰu:dɐ]	[pu:doˤ]
KN	lion	[lø:və]	[le:ʋə]
KN	needle	[na:dl̩]	[noˤ:dl̩]
KN	sleigh	[ʃlɪtn̩]	[ʃlɪdn̩]
BD	life	[le:bən]	
BD	coat	[mantl̩]	
BD	test	[pʰry:fʊŋ]	
CX	Satan	[za:tʰan]	
CX	cabin	[hʏtʰə]	
CX	trauma	[tʰra͜ʊma]	
ES	grate	[kɪtʰɐ]	
ES	conductor	[ʃafnɐ]	
ES	mold	[ʃɪm̩l]	
IK	string	[fa:dn̩]	
IK	cave	[hø:lə]	
IK	cat	[kʰa͡tsə]	
KN	beating	[pʰry:gl̩]	
KN	garden	[kaʁtn̩]	
KN	graphic	[kra:fɪkʰ]	