Individual differences in the perception of regional, nonnative, and disordered speech varieties[a)]

Tessa Bent[b)]
Department of Speech and Hearing Sciences, Indiana University, Bloomington, Indiana 47405, USA

Melissa Baese-Berk
Department of Linguistics, University of Oregon, Eugene, Oregon 97403, USA

Stephanie A. Borrie
Department of Communicative Disorders and Deaf Education, Utah State University, Logan, Utah 84332, USA

Megan McKee
Department of Speech and Hearing Sciences, Indiana University, Bloomington, Indiana 47405, USA

(Received 26 January 2016; revised 10 October 2016; accepted 19 October 2016; published online 17 November 2016)

Speech perception abilities vary substantially across listeners, particularly in adverse conditions including those stemming from environmental degradation (e.g., noise) or from talker-related challenges (e.g., nonnative or disordered speech). This study examined adult listeners' recognition of words in phrases produced by six talkers representing three speech varieties: a nonnative accent (Spanish-accented English), a regional dialect (Irish English), and a disordered variety (ataxic dysarthria). Semantically anomalous phrases from these talkers were presented in a transcription task and intelligibility scores, percent words correct, were compared across the three speech varieties. Three cognitive-linguistic areas—receptive vocabulary, cognitive flexibility, and inhibitory control of attention—were assessed as possible predictors of individual word recognition performance. Intelligibility scores for the Spanish accent were significantly correlated with scores for the Irish English and ataxic dysarthria. Scores for the Irish English and dysarthric speech, in contrast, were not correlated. Furthermore, receptive vocabulary was the only cognitive-linguistic assessment that significantly predicted intelligibility scores. These results suggest that, rather than a global skill of perceiving speech that deviates from native dialect norms, listeners may possess specific abilities to overcome particular types of acoustic-phonetic deviation. Furthermore, vocabulary size offers performance benefits for intelligibility of speech that deviates from one's typical dialect norms.
© *2016 Acoustical Society of America*. [http://dx.doi.org/10.1121/1.4966677]

[DB] Pages: 3775–3786

I. INTRODUCTION

Speech perception is generally a remarkably robust process. Listeners can successfully recognize words produced by a variety of talkers in a wide range of environmental listening conditions. However, the introduction of certain adverse listening conditions can substantially lower speech communication success (Mattys *et al.*, 2012). For example, the speech from talkers with unfamiliar regional dialects or nonnative accents can be more difficult to understand and slower to process (Adank *et al.*, 2009), as is the speech from individuals with pathological speech disorders (Borrie *et al.*, 2012a).[1] Similarly, challenging environmental listening conditions, such as high levels of noise or reverberation, can substantially hinder a listener's speech understanding (Sumby and Pollack, 1954). Although these types of listening adversity can cause substantial intelligibility decrements, listeners are also capable of significant learning and show improved abilities to map unfamiliar or disordered speech signals onto words in their mental lexicons following exposure and practice (e.g., Davis *et al.*, 2005; Bradlow and Bent, 2008; Sidaras *et al.*, 2009; Borrie *et al.*, 2012b; Baese-Berk *et al.*, 2013).

To date, models of speech perception and learning have been largely based upon group averages. However, there is evidence that significant individual variability in perceiving speech tends to emerge under adverse listening conditions. Even in groups of normal-hearing young-adult listeners, individuals vary substantially in their ability to recognize words in noise (Gilbert *et al.*, 2013) or to understand disordered speech (Borrie, 2015). These findings suggest that some listeners are more adept at navigating the complex cognitive-perceptual task of perceiving speech under challenging conditions (Borrie, 2015). However, whether listeners possess a global speech perception ability, which allows them to decipher speech in a broad range of adverse conditions, as opposed to specific expertise at understanding speech with certain types of acoustic-phonetic deviations, has yet to be addressed.

[a)]Portions of this work were previously presented in Baese-Berk *et al.* (2015) "Individual differences in perception of unfamiliar speech," in The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*, The University of Glasgow, Glasgow, United Kingdom, paper number 0460, pp. 1–5.

[b)]Electronic mail: tbent@indiana.edu

One of the central issues stemming from recent perceptual studies is the identification of listener-related variables that may explain why certain individuals are more successful at perceiving speech in adverse conditions than others. Several linguistic factors have been identified as possible contributors to individual differences in speech perception under adverse listening conditions. Specifically, listeners with larger vocabularies tend to more accurately perceive native speech in noise (Tamati *et al.*, 2013), disordered speech (McAuliffe *et al.*, 2013), and constructed unfamiliar accents (Banks *et al.*, 2015) [however, see Benichov *et al.* (2012) for conflicting results]. Further, an enhanced ability to use indexical properties (i.e., regional dialect identification and gender discrimination) is correlated with increased word identification in noise (Tamati *et al.*, 2013). Many cognitive abilities have also been observed to relate to individuals' perception of novel accents and speech in noise, including working memory, executive function, short-term memory, general cognitive function, making perceptual wholes from fragments, familiar sound recognition, and task switching (Watson *et al.*, 1996; Kidd *et al.*, 2007; Akeroyd, 2008; Adank and Janse, 2010; Benichov *et al.*, 2012; Janse and Adank, 2012; Tamati *et al.*, 2013; Banks *et al.*, 2015). An individual listener's ability to adapt to an unfamiliar constructed accent also is related to several cognitive-linguistic factors including inhibition, selective attention, and vocabulary size (Janse and Adank, 2012; Banks *et al.*, 2015). Together, these results highlight how higher level cognitive-linguistic functions may predict one's ability to perceive and adapt to ambiguity in the speech signal.

Although these studies have shed light on an interaction between cognitive-linguistic skills and speech perception abilities, a comprehensive model accounting for individual differences in speech perception, particularly for young normal-hearing listeners, does not exist. Furthermore, nearly all of the existing studies have focused on a single source of degradation and therefore, it is unknown whether the same cognitive-linguistic skills will be predictive of word recognition performance for multiple speech varieties when measured under the same experimental conditions. The investigation of this issue will provide information about whether skills used to recognize words produced by talkers with one type of unfamiliar speech will generalize to talkers from another unfamiliar variety. Therefore, the first aim of the present study is to investigate whether intelligibility scores across listeners for three speech varieties—nonnative-accented speech, a regional dialect, and a neurogenic speech disorder—are related. The second aim is to assess how several cognitive-linguistic variables contribute to an individual's perception of the three sources of speech degradation.

The three speech varieties employed in the current study are all examples of talker-related adverse conditions. Both the etiology of the production differences as well as the ultimate realization of segmental and suprasegmental features are distinct for each speech variety (see Table I for a summary). The talkers of the unfamiliar regional dialect (in this study, Dublin Irish English) are native speakers of English. American English—the dialect of the listeners—and Irish English share many acoustic-phonetic properties. For example, the two varieties share similar rhythmic patterns; Irish English, like standard American English, is traditionally described as stress-timed—meaning there is substantial durational variability across vocalic units (Dorn *et al.*, 2012). However, this dialect has a number of phonemic features that differ from American English. Many of these segmental-level differences reside in the vowels (Hickey, 2004; Hughes *et al.*, 2012). For example, the realization of the diphthong /aɪ/ has a raised onglide (i.e., /ɔɪ/) relative to American English so that "behind" is pronounced /bihɔɪnd/ (Hickey, 2004). /ʊ/ is found in some words where /ʌ/ or /ə/ would be used in American English, such as in "sunken" as /sʊŋkən/ (Hughes *et al.*, 2012). Similar to Received Pronunciation, a subset of words that are realized with /æ/ in American English have /ɑ/, as in "after" as /ɑftɚ/ (Hughes *et al.*, 2012). Although the two dialects differ primarily in vowel realization, several consonant differences also exist (Hickey, 2004; Hughes *et al.*, 2012). Specifically, /θ/ and /ð/ can be realized as dentalized stops (i.e., /t̪/ and /d̪/), so that "beneath" is pronounced /bənit̪/ and "then" is pronounced /d̪ɛn/. Unlike in American English, where intervocalic /t/ tends to be realized as a flap, the realization of intervocalic /t/ can vary including a fricative realization of /t/, tap, lenition into /h/ as in [mɑhɚ̬ɻ̊] for "matters," or /ʔ/.

The second speech variety employed in the current study is a nonnative variety (i.e., Spanish-accented English).

TABLE I. Examples of segmental and suprasegmental differences between standard American English and the three speech varieties included in the study.

Speech variety	Consonants	Vowels	Suprasegmentals
Irish English	• /θ/ → /t̪/ • /ð/ → /d̪/ • intervocalic /t/ → /ɾ/, fricative realization of /t/, /h/ or /ʔ/	• /aɪ/ → /ɔɪ/ • /ʌ/ or /ə/ → /ʊ/ (some words) • /æ/ → /ɑ/ (some words)	• stress-timed language, similar to American English
Spanish-accented English	• shortened VOT for voiceless stops • devoicing of word final obstruents • epenthesis in s-clusters	• duration and spectral differences for tense-lax pairs • unreduced vowels	• rhythmic differences; fewer differences between stressed and unstressed vowels • slower speaking rate
Ataxic dysarthria	• imprecise consonant articulation • irregular articulatory breakdowns	• imprecise vowel articulation • irregular articulatory breakdowns	• equal and excess stress • prolonged phonemes and intervals • reduced pitch variation • reduced loudness variation

From a listener's perspective, words produced by nonnative speakers also result in mismatches to the majority of stored exemplars and deviate from native dialect norms in both segmental and suprasegmental features (Sereno *et al.*, 2016). However, the underlying source of these differences arises from the interaction between the first and second languages of the speaker as well as from universal properties of nonnative speech. Because many of the features of nonnative accents are caused by interactions between the first and second language sound systems, the features of nonnative accents are systematic both within talkers as well as across talkers with the same first language (L1)–second language (L2) pairing. However, in some cases, nonnative speech can have greater intra-speaker variability both in the segmental (Wade *et al.*, 2007; Hanulikova and Weber, 2012) and the suprasegmental (Baese-Berk and Morrill, 2015) domains relative to native speech. For Spanish-accented English specifically, differences from standard American English have been observed in the mean realization of English consonants, including voice onset time for stops (Flege, 1991) and devoicing of word final obstruents (Imai *et al.*, 2005). Spanish language, relative to English, is also characterized by a restricted vowel inventory with only five vowels /i e a o u/. Additionally, there is less formant movement in Spanish vowels compared to English vowels and Spanish vowels do not show the type of duration contrasts found for tense-lax pairs in English (e.g., /i/ vs /ɪ/). These differences between the L1 and L2 vowel systems result in Spanish-speaking learners of English having difficulties producing both vowel duration and spectral differences for tense-lax vowel pairs (Flege *et al.*, 1997) as well as difficulties producing reduced vowels (Flege and Bohn, 1989). Differences in phonotactics across the two languages also cause deviations from native norms. For example, Spanish lacks s-clusters, which results in frequent epenthesis for these clusters in Spanish-accented English (Carlisle, 1991). In terms of suprasegmental features, Spanish is traditionally described as a syllable-timed language, with relatively low variability across vocalic segments (Grabe and Low, 2002). Due to influence from their first language, Spanish-accented talkers of English tend to show fewer temporal distinctions between stressed and unstressed vowels compared to standard American talkers (Shah, 2003).

The final form of variability included here is a disordered speech variety, dysarthria. Dysarthria is a motor speech disorder caused by neurologic impairment such as vascular disease, traumatic brain injury, degenerative disease, tumors, or infectious processes (Duffy, 2005). Such impairment can interfere with any or all of the subsystems involved in speech production including respiration, phonation, articulation, and resonance. Thus, the speech produced by individuals with dysarthria can differ from speech from typical speakers in many areas. Ataxic dysarthria—the specific form of dysarthria employed in the current study—reflects cerebellar damage which manifests in incoordination of motoric behavior. In regards to speech production, ataxic dysarthria is characterized predominately by breakdowns in both articulatory and prosodic aspects of speech. Ataxic dysarthria, therefore, deviates from native language norms in both segmental and suprasegmental speech features. Segmental deviations include imprecise articulation of consonants and vowels (perceived as slurry speech) and irregular articulatory breakdowns, whereas suprasegmental deviations include equal and excess speech stress (robotic-sounding speech with pauses after each syllable), prolonged phonemes and intervals (giving the impression of slow rate), and reduced variation in both pitch and loudness (Darley *et al.*, 1969a,b). Further, the speech deficits in dysarthria, and particularly ataxic dysarthria with the underlying pathophysiology of muscular incoordination, tend to manifest as varied and unpredictable, "…phonemes produced adequately in one context may be distorted or omitted in the next word, speech may deteriorate in a mumbled rush of speech at the end of a sentence, and voicing may break or cease intermittently" (Borrie *et al.*, 2012b). Due to these speech production deficits, individuals with the neurogenic speech disorder of dysarthria are frequently less intelligible than typical speakers (Kent *et al.*, 1989).

Due to the deviations from typical, standard American English dialect norms, each of the speech varieties included in this study presents challenges for accurate word recognition for American English listeners. To accurately identify words produced by talkers with less familiar dialects or accents as well as those with speech disorders, listeners need to overcome mismatches between the majority of their experiences of a word and the in-coming acoustic-phonetic information. These mismatches may cause ambiguities in the interpretation of the speech signal, which some listeners are more adept at overcoming than others. Therefore, one of the purposes of the current study was to investigate three factors—cognitive flexibility, inhibitory control of attention, and vocabulary size—as possible skills underlying the individual differences in the perception of three less familiar speech varieties. For all the speech varieties included in this study, there are differences from standard American English in both consonant and vowel realizations and two of the varieties differ substantially in the suprasegmental domain as well.

To identify words produced by talkers from these speech varieties, listeners have to flexibly interpret the incoming acoustic-phonetic information, which may not be an ideal fit to any previously encountered productions of the words. Therefore, listeners who have a greater capacity for cognitive flexibility may be better equipped to interpret speech that deviates from the majority of their experiences. Likewise, as listeners encounter the novel pronunciations, there are likely to be acoustic-phonetic cues to phoneme and word identity may be ambiguous or misleading. Listeners will need to suppress attention to information that is misleading as to a word's identity. Therefore, listeners who are better able to tune their attention to target information while ignoring distracting information may be better able to accurately perceive words that deviate from typical dialect norms. Both cognitive flexibility and selective attention have been found to relate to accurate perception of or adaptation to an artificial constructed accent (Adank and Janse, 2010; Janse and Adank, 2012).

We extend this work to examine whether these two cognitive abilities are related to the perception of three naturally produced speech varieties that differ in various ways from typical standard American English. Last, we include one linguistic attribute as a possible predictor of individual differences, vocabulary size. Vocabulary size may benefit the perception of unfamiliar or degraded speech signals due to its relationship to linguistic experience. Listeners with large vocabulary sizes likely have a greater bank of experiences to draw upon when encountering speech that deviates from typical home dialect norms. Vocabulary size has been found to relate to accurate identification of words in anomalous phrases as produced by speakers with dysarthria (McAuliffe *et al.*, 2013) as well as to accuracy with artificially constructed accents in noise with meaningful sentences (Banks *et al.*, 2015). Here, we investigate the relationship between receptive vocabulary size and perception of three naturally produced speech varieties without the influence of environmental degradation (e.g., noise).

In addition to the investigation of whether performance can be predicted from individual cognitive-linguistic abilities, the second purpose of the current study was to investigate whether global speech perception skills exist: Is the ability to perceive one unfamiliar speech variety related to the ability to perceive another unfamiliar variety? That is, will listeners who are highly successful at perceiving one variety of unfamiliar speech also be adept at perceiving other varieties? It is possible that success in perceiving one type of unfamiliar speech would predict success at perceiving another type of unfamiliar speech, evidencing a global skill in the ability to recognize speech in adverse conditions. Some work has demonstrated a relationship between a listener's word recognition abilities across different adverse listening conditions (e.g., different signal-to-noise ratios or sentence types) (Gilbert *et al.*, 2013). However, it is also possible that listeners could excel at one unfamiliar speech type while not performing as well on another. Given that each of the three speech varieties included in the current study present listeners with only partially overlapping challenges for accurate word recognition, it may be the case that some listeners are more adept at overcoming certain types of deviations in the speech signal than others, reflecting sensitivity and/or preference for a particular strategy for processing speech in adverse listening conditions. For example, if a listener was particularly adept at understanding speech that differed from typical dialect norms in the suprasegmental domain, they may be better at perceiving the nonnative-accented variety and the disordered variety, but would not have as great an advantage for the non-home dialect, which does not differ as substantially from the home dialect on this dimension compared to the other two varieties.

II. EXPERIMENT 1

Experiment 1 provides baseline intelligibility data for talkers from the three speech varieties—Irish English, Spanish-accented English, and ataxic dysarthria—when their productions were presented in a single variety condition (i.e., not interspersed with talkers representing the other varieties). Furthermore, these data were used to assess whether learning across the experiment differed for conditions with a single variety compared to conditions in which listeners were presented with multiple speech varieties (experiment 2).

A. Method

1. Participants

Sixty native, monolingual speakers of American English between the ages of 18 and 28 yrs [$M = 19.6$ yrs; standard deviation (SD) = 1.9] participated in experiment 1. Participants completed a language background and experiences questionnaire, which was administered prior to the onset of the speech perception test, and included questions about their experiences interacting with nonnative speakers, native speakers from other countries, and individuals with speech disorders. On this questionnaire, none of the participants reported significant prior exposure to any of the three speech varieties employed in the current study. No participants reported any history of speech, language, or cognitive impairment.

2. Stimuli

The stimuli consisted of 74 syntactically plausible but semantically anomalous phrases, created specifically for examining speech perception in adverse conditions (Liss *et al.*, 1998) (see the supplementary material[2]). These phrases have been used extensively in the study of perception of dysarthric speech (e.g., Borrie *et al.*, 2012b; Borrie, 2015). Semantically anomalous phrases (e.g., *award his drain away*) were used to reduce the influence of top-down semantic and contextual knowledge. Phrases were all six syllables in length and ranged from three to five words.

Two talkers (one male and one female) from each of the following three speech varieties produced the stimuli: Irish English, Spanish-accented English, and dysarthric speech. Talkers were carefully selected based on perceptual presentation of the cardinal speech features of their speech variety (see Table I). Accordingly, acoustic-phonetic deviations across the two talkers for each variety were relatively homogeneous. Irish talkers were from Dublin and spoke Irish in addition to English; the male talker was 73 yrs old and the female talker was 19 yrs old. The Spanish-accented talkers included a 24-yr-old male from Valencia, Venezuela, and a 27-yr-old female from Bogotá, Colombia. The Spanish-accented male had studied English in his home country since age four and the female since age seven. Both talkers with dysarthria presented with a mild-moderate ataxic dysarthria—as diagnosed by three independent Speech-Language Pathologists with expertise in differential diagnosis of motor speech disorders—resulting from cerebellar lesions post cerebral vascular accident (stroke). The male talker was 72 yrs old and 14 yrs post stroke. The female talker was 68 yrs old and 8 yrs post stroke.

3. Procedure

Participants completed a speech perception task, which was presented in EPrime (Psychology Software Tools,

Sharpsburg, PA) using Sennheiser HD 280 pro headphones (Sennheiser, Old Lyme, CT). Participants were randomly assigned to one of the three varieties of speech, hearing only phrases from the male and female talker of a single variety (i.e., only Irish English, Spanish-accented English, or dysarthric speech). The participants were first presented with six randomized practice trials. The practice trials included two phrases: the male talker from one variety produced one of the phrases and the female talker from the same speech variety produced the other. After the practice trials, listeners were presented with 72 test phrases (36 from the male talker and 36 from the female talker from 1 of the 3 varieties), which were presented in randomized order for each participant. The specific phrases assigned to each talker were randomly selected for each participant as well. Phrases were presented one at a time, and following each presentation, participants were instructed to use the keyboard to type out exactly what they thought was being said. Participants were encouraged to guess for any words they did not recognize. Once they had finished typing their response, participants were prompted to click a button on the screen to move on to the next phrase. They did not receive any feedback regarding the accuracy of their transcription attempt.

B. Analysis

Participant transcripts were analyzed for a standard measure of speech intelligibility, percent words correct. This analysis resulted in one intelligibility score for each listener. Words that included homophones or common misspellings were scored as correct. However, words with added or deleted morphemes were scored as incorrect (e.g., *supplied* for *supplies*). Two judges obtained 97% reliability on 10% of participant transcripts.

C. Results and discussion

Descriptive data are presented in Table II. Scores for each speech type reflect performance of 20 listeners. Of note, comparable average intelligibility scores were found for the dialect and dysarthric speech varieties, whereas the average intelligibility score of the nonnative speech variety was substantially lower.

In addition to measuring intelligibility performance for each of the speech varieties, we also assessed learning, or adaptation, during the course of the speech perception task by calculating performance across quartiles of phrases—participants' intelligibility scores for 4 blocks of 18 phrases each were calculated. Participants who were exposed to Irish English did not demonstrate significant improvement from the first to the fourth quartile [$t(19) = 0.037$, $p = 0.97$]. However, participants who were exposed to Spanish-accented English and dysarthric speech demonstrated significant improvement from the first to fourth quartile of the experiment [$t(19) = -2.31$, $p < 0.04$, and, $t(19) = -4.06$, $p < 0.001$, respectively]. These results suggest that perceptual adaptation occurred for two of the three speech varieties when listeners were exposed to talkers representing a single speech variety during the course of the experiment. These data are illustrated in Fig. 1 below. Although the Irish talkers and talkers with dysarthria had similar overall levels of intelligibility, performance in the first quartile of the experiment was much lower for the dysarthric talkers than the Irish talkers. This lower level of performance in the initial quartile for the dysarthric talkers may have provided more room for improvement in subsequent quartiles. Furthermore, the dysarthric speech variety may have been even less familiar to listeners than Irish English and thus, with even a relatively short amount of experience, listeners were able to overcome some of their initial challenges in mapping the degraded acoustic-phonetic properties to words in their lexicons.

TABLE II. Mean, minimum, and maximum percent words correct scores for the three speech varieties in experiments 1 and 2. Standard deviations are shown in parentheses after the mean scores.

	Nonnative		Dialect		Dysarthria	
Comparison	Exp.1	Exp. 2	Exp. 1	Exp. 2	Exp. 1	Exp. 2
Mean	63 (6)	64 (6)	78 (6)	79 (7)	78 (7)	78 (7)
Minimum	50	44	63	56	72	55
Maximum	71	74	87	91	83	89

III. EXPERIMENT 2

The goal of this experiment was to assess whether listeners who are particularly proficient at perceiving the speech from one variety that deviates from typical, home dialect norms are also more accurate at identifying words from talkers representing another speech variety. To address this question, listeners were presented with all three varieties within the same speech perception task. Furthermore, we assessed three cognitive-linguistic skills—inhibitory control of attention, cognitive flexibility, and receptive vocabulary—to investigate whether these skills relate to accurate perception of any or all of the speech varieties tested. Last, we investigated adaptation during the course of mixed variety blocks and compared it to performance in the single variety blocks from experiment 1.

A. Method

1. Participants

Fifty-one monolingual, native speakers of American English (38 female) between the ages of 18 and 33 yrs ($M = 21$ yrs; SD = 2.5) participated. All participants passed a pure-tone hearing screen at octave interval frequencies from 500 to 8000 Hz at 20 dB and 250 Hz at 25 dB. Two additional participants were tested, but were excluded from the data analysis due to bilingual status. Participants completed a language background questionnaire prior to the onset of the experiment tasks, as described above for experiment 1. None of the participants reported significant prior exposure to any of the three speech varieties included in the study. None of the participants reported any history of speech, language, or cognitive impairment.

2. Stimuli

The stimuli were the same as those used in experiment 1.

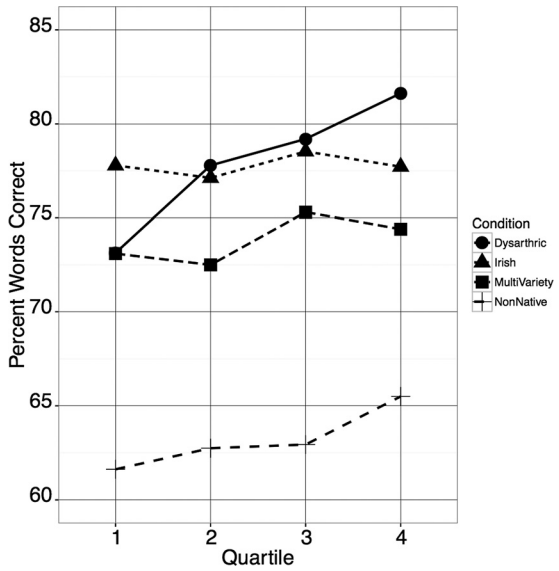

FIG. 1. Percent words correct scores across quartiles are shown for experiment 1 (single varieties) and experiment 2 (multiple varieties). Scores are shown for listeners who were presented only with the talkers with ataxic dysarthria (circles with a solid line), only with the Irish English talkers (triangles with a short-dashed line), only with the nonnative talkers (plusses with a long-dashed line), or with all three varieties (squares with a long-dashed line).

3. Procedure

Participants were tested individually in one experimental session lasting approximately 1 h. During the session, participants completed two cognitive assessments, a vocabulary assessment, and a speech perception task. The cognitive tests included the Intra-Extra-Dimensional Set Shift Task (IED) from the Cambridge Neuropsychological Test Automated Battery (Cambridge Cognition, 2014) and the Flanker test of cognitive inhibition, which were both administered on a Dell Optiplex 760 computer (Dell, Round Rock, TX) with an Elo touch screen monitor (Elo, Milpitas, CA). The IED is a test of rule acquisition and reversal that measures cognitive flexibility. The test uses two dimensions of visual stimuli, white lines and pink shapes. In a series of stages, participants learn through trial and error to select the correct line or shape. In subsequent stages, the "rule" for the correct response changes and participants must adapt. The two critical stages of the test measure the participants' abilities with intra- and extra-dimensional shifting. In the intra-dimensional shift stage, the correct response within a category (i.e., lines or shapes) changes. In the extra-dimensional shift stage, the correct response switches from lines to shapes (or shapes to lines). The Flanker task was used to measure inhibitory control of attention. This task was run through Pebl (Mueller, 2014). Participants clicked the appropriate key on the keyboard indicating the direction of a central arrow while ignoring surrounding arrows. Trials included congruent trials in which the flanking arrows were pointing in the same direction as the target arrow (e.g., $<<<<<$), incongruent trials in which the flanking arrows pointed in the opposite direction (e.g., $<<>><<$), and neutral trials without any flanking arrows (e.g., $<$). Accuracy and reaction time were measured. Visual tasks for assessing cognitive flexibility and inhibitory control of attention were selected to determine whether there was a modality-independent contribution of these cognitive skills to the perception of less familiar speech varieties.

The Peabody Picture Vocabulary Test (PPVT-4; Dunn and Dunn, 2007), a standardized assessment tool, was administered to evaluate receptive vocabulary. The PPVT-4, administered in hard copy, requires participants to work through a series of picture sets, identifying a picture (out of four possible options) that best illustrates a spoken word (e.g., identify "lifting").

The speech perception task was presented by custom-designed software written in Python and presented on a Mac mini (Apple, Cupertino, CA). Participants were individually tested in a sound-attenuated booth, and the speech stimuli were presented through Sennheiser HD 280 pro headphones. The participants were first presented with six randomized practice trials. The practice trials included two phrases: the male speaker from each variety produced one of the phrases and the female speaker from each speech variety produced the other, such that listeners heard each phrase three times. After the practice trials, participants were presented with 72 phrases (12 from each talker). Although the participants heard the same set of 72 phrases, the phrases assigned to each talker were randomly selected for each participant. Likewise, stimuli were presented in a randomized order for each participant with all the talkers and varieties randomized (i.e., talkers or varieties were not blocked). All other aspects of the speech perception task were the same as in experiment 1. Participants were not informed about the specific speech varieties included in the speech perception test.

B. Analysis

As in experiment 1, participant transcripts were analyzed for a standard measure of speech intelligibility, percent words correct, using the same scoring criteria. This analysis resulted in three percent correct scores for each participant in experiment 2, one for each variety. Two judges obtained 99% reliability on 50% of the participant transcripts. The cognitive and vocabulary assessments were scored using standardized scoring protocols.

C. Results and discussion

1. Intelligibility scores

Intelligibility data, expressed by the mean percent words correct for each speech variety—Irish English, Spanish-accented English, and ataxia dysarthria—were collected for each of the 51 participants. Descriptive data are presented in Table II. Descriptive data from the listeners in experiment 2, in which listeners were exposed to all three linguistic varieties demonstrated remarkable consistency with the descriptive data from experiment 1, in which listeners only heard one linguistic variety. Specifically, the mean and standard deviations were comparable.

As in experiment 1, listeners' adaptation during the course of the experiment was analyzed (Fig. 1). A paired *t*-test comparing the first and fourth quartile revealed that

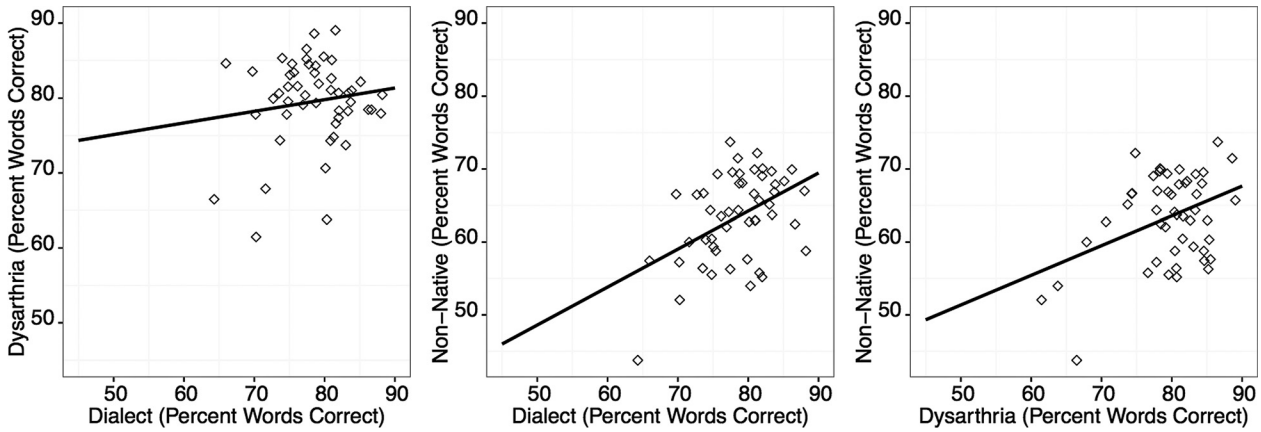

FIG. 2. Scatter plots showing individual listeners' percent words correct scores. The panels show the relationships between scores for the unfamiliar dialect and ataxic dysarthria (left panel), for the unfamiliar dialect and nonnative accent (middle panel), and for the ataxic dysarthria and nonnative accent (right panel).

performance was stable when listeners were presented with all three speech varieties [$t(50) = -1.09$, $p = 0.28$] with scores averaging from 73% correct in the first quartile and 74% correct in the fourth quartile. This result suggests that listeners did not learn to better recognize multiple varieties of unfamiliar speech during the course of the single speech perception task. This result differs from the experiment 1 results, in which listeners' perception of both the nonnative and dysarthric varieties improved over the course of the experiment.

Correlations between intelligibility data for each of the three speech varieties were conducted to determine whether intelligibility scores for each speech type would correlate with intelligibility for the other two speech types. That is, are listeners who are proficient at deciphering the speech of one variety also successful at deciphering the speech of other varieties? The intelligibility scores entered into the correlations were averaged across all items and across the two talkers representing each speech variety. Results revealed that intelligibility scores for the nonnative-accented speech were correlated with the intelligibility scores of both the dysarthric speech ($r = 0.388$, $p = 0.005$) and the regional dialect ($r = 0.455$, $p = 0.0009$). However, the intelligibility scores for the regional dialect and the dysarthric speech were not significantly correlated ($r = 0.142$, $p = 0.323$). Thus, listeners who were successful at recognizing the nonnative-accented speech were also more successful with both the regional dialect and the dysarthric speech. However, the ability to recognize words produced by talkers with dysarthria did not relate to the ability to recognize words produced by talkers with the native dialect. Figure 2 shows the correlations for each pair of speech varieties.

2. Cognitive-linguistic variables

The intelligibility scores for each participant were also analyzed with linear mixed effects regression models to determine whether the two cognitive assessments and/or the vocabulary assessment predicted speech perception performance.[3] The cognitive-linguistic scores that were entered into the model as fixed factors were as follows. Two z-scores from the IED test were entered: Intra-Dimensional Shift and Extra-Dimensional Shift. These two scores were employed as measures of cognitive flexibility. The flanker conflict cost score, which was calculated using the logarithm of the mean response time (MRT) for correct incongruent and neutral trials in the formula [(MRT incongruent-MRT neutral)/MRT neutral], was employed as the measure of inhibitory control of attention. The receptive vocabulary score was the raw score from the PPVT. Because these cognitive-linguistic measures were employed to account for the participant variable, participant data were not entered as random items. In addition to the cognitive-linguistic scores, speech type (i.e., nonnative, dialect, and dysarthria) was entered into the regression as a fixed factor. Speech type was contrast coded to compare Irish English to the other two varieties, and then to compare the dysarthric speech to the Spanish-accented English. Interactions between speech type and each of the cognitive factors above were also entered as fixed factors. Therefore, the full model included as fixed factors: the two contrast-coded measures of speech type, PPVT, Intra-dimensional Shift, Extra-dimensional Shift, flanker cost score, and all interactions. Significance was determined using model comparisons to determine whether each fixed factor contributed significantly to the model. The final model with the best fit included the fixed factors of speech type (i.e., dialect, dysarthria, nonnative) and PPVT score. The best fitting model is summarized in Table III. None of the interactions between cognitive variables and speech variety were significant predictors of intelligibility performance on the speech perception tasks (all p-values > 0.2). Neither of the IED scores nor the flanker test emerged as significant predictors of intelligibility performance on the speech perception task (all p-values > 0.17), suggesting that listeners' inhibitory control of attention and cognitive flexibility scores do not directly relate to performance on this word recognition task. PPVT scores, however, did emerge as a significant predictor of performance on the speech perception task ($\chi^2 = 11.003$, $p < 0.001$). This result suggests that individuals with higher receptive vocabulary scores were more successful at recognizing speech under adverse conditions. Participants' scores on this assessment ranged widely; standard scores ranged

TABLE III. Linear mixed effects model summary.

Predictor	Estimate	Standard error	t-value
(Intercept)	0.6799	0.0198	34.31
Dialect vs others	0.0584	0.0089	6.56
Dysarthria vs Nonnative	0.1509	0.0103	14.67
Receptive vocabulary	0.0009	0.0003	3.38

from 79 to 132 ($M = 108$; SD = 10.8) and percentiles ranged from 8 to 98 ($M = 67$; SD = 21.6). Speech type was also a significant predictor of model fit ($\chi^2 = 83.308$, $p < 0.0001$).

While the two measures of cognitive flexibility and the inhibitory control of attention measure did not directly predict performance on the speech recognition task, each of these variables was significantly correlated with the measure of receptive vocabulary (see Table IV). Therefore, we hypothesized that there might be an indirect relationship between these variables and intelligibility that is mediated by receptive vocabulary scores. To test this hypothesis, we performed a path analysis of our data using SmartPLS (Ringle *et al.*, 2015), which uses a partial-least squares approach to path analysis. Because our sample size is relatively small for this type of analysis, we used bootstrapping to construct confidence intervals for each of these estimates. Figure 3 shows the path model for the predictors of intelligibility with standardized coefficients. As predicted, there is a direct relationship between intelligibility and vocabulary size ($\beta = 0.076$, $p < 0.001$). Further, the relationship between intelligibility and each of the other three factors (cognitive and perceptual flexibility and inhibitory control of attention) is mediated by receptive vocabulary size (extra-dimensional shift: $\beta = 0.320$, $p < 0.001$; intra-dimensional shift: $\beta = 0.331$, $p < 0.001$; inhibitory control of attention: $\beta = 0.181$, $p < 0.001$). A multi-group analysis revealed no significant differences among the three speech varieties with regard to this type of analysis, suggesting that the cognitive and linguistic factors examined here may similarly influence intelligibility of these speech varieties.

IV. GENERAL DISCUSSION

A wide range of variability across listeners in their abilities to successfully perceive speech that deviated from native dialect norms was observed. This range of variability was robust for each of the three speech varieties, with intelligibility scores for each variety in experiment 2 reflecting at least a 30-percentage point range across listeners. This range of intelligibility scores is particularly striking given that all the listeners were young adults with normal hearing and little to no experience with any of the three varieties included. The inclusion of listeners with varying sensory abilities, age, and/or experience with the speech varieties would have likely led to an even greater performance range. Likewise, embedding the stimuli in noise would also be likely to increase the performance range across listeners.

TABLE IV. Correlations among factors (* indicates a significant correlation).

	Percent words correct	Receptive vocabulary	Flanker conflict	Intra-dimensional shift
Receptive vocabulary	0.395*			
Flanker conflict	−0.002	0.148*		
Intra-dimensional shift	0.020	−0.342*	0.014	
Extra-dimensional shift	−0.002	−0.102*	−0.251*	−0.797*

Although a wide range of variability in intelligibility performance has been reported in previous work with a single variety of speech degradation (Janse and Adank, 2012; Atagi and Bent, 2015; Borrie *et al.*, 2016), this study is, to our knowledge, the first of its kind to investigate whether the ability to perceive one unfamiliar speech variety is related to the ability to perceive another variety. That is, are listeners globally skilled at perceiving speech that deviates from native dialect norms or do listeners possess specific expertise with certain types of speech deviations? Here, we observed that the ability to perceive nonnative-accented speech was related to the ability to perceive both other varieties (i.e., the regional and disordered varieties); however, this relationship was not evident between perception of the regional and disordered varieties. These results suggest that listeners may not be globally skilled at perceiving talkers whose speech deviates from native dialect norms, but rather may possess an aptitude to map particular types of acoustic-phonetic deviation onto words in their lexicons. While there are a number of ways in which the similarities and differences across the three speech varieties employed here could be compared, in an attempt to explain the observed patterns, two are explored below.

Listeners may differ in their abilities to navigate specific types of deviations in the speech signal. Broadly, some listeners may be more adept at understanding speech with significant deviations at the suprasegmental level whereas others may be more skilled at understanding speech that differs from native dialect norms in segmental features. Given the pathological rhythmic speech deficits that characterize ataxic dysarthria, we can assume that this speech type exhibited the greatest deviations from native norms in terms of suprasegmental properties. In contrast, Irish English and standard American English have very similar rhythmic properties. Thus, in terms of suprasegmental features, the unfamiliar dialect likely exhibited the least amount of deviation from the native dialect norms of our participants. If we consider this characteristic along a continuum from greatest to least, Spanish-accented English, which frequently deviates from standard English in the temporal patterns of stressed and unstressed vowel realizations (Shah, 2003), likely falls somewhere in between. Therefore, listeners who showed aptitude at understanding speech that deviated from native dialect norms in suprasegmental characteristics might be at a greater advantage for understanding the dysarthric and nonnative varieties than the Irish variety. In contrast, for segmental features, specifically vowels, Irish talkers and nonnative talkers generally show greater deviations from native American English dialect norms than dysarthric talkers. Both Irish and Spanish-accented talkers show substantial patterns of vowel substitutions compared to standard

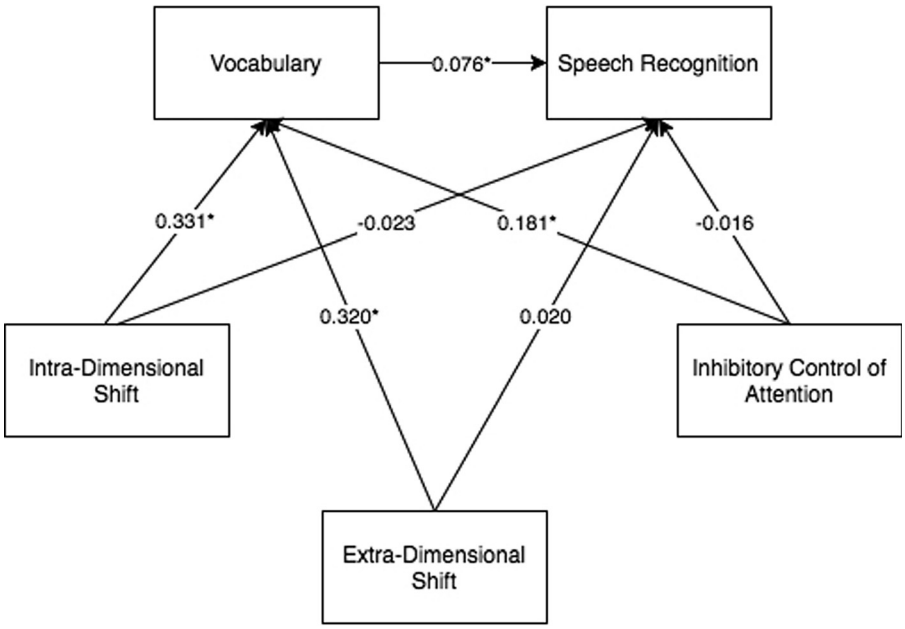

FIG. 3. Path model for the direct and mediated effects of cognitive and linguistic factors on percent words correct scores for three speech varieties. Numbers on each pathway represent coefficients. Pathways between the three variables on the bottom and percent words correct were not significant and were mediated by vocabulary.

American English (Flege and Bohn, 1989; Flege *et al.*, 1997; Hickey, 2004; Hughes *et al.*, 2012). For talkers with ataxic dysarthria, there is a lack of precision with making the correct articulatory contacts; however, vowel category substitutions are very rarely observed (Duffy, 2005). Thus, a listener who was adept at recovering from vowel substitutions would excel at understanding the Irish talkers and also be advantaged for the nonnative talkers, but this skill would afford very little advantage for understanding dysarthric speech. Considering deviations from native norms on a continuum for both suprasegmental and segmental attributes, the nonnative talkers pattern with the Irish talkers on one of these dimensions (vowels) but with talkers with dysarthria on the other (suprasegmentals). Conversely, ataxic dysarthria and Irish English present with less overlap on these particular dimensions. These patterns may explain why intelligibility performance on nonnative speech was correlated with intelligibility performance on both Irish English and dysarthric speech, and why correlations with intelligibility performance were not observed between Irish English and dysarthric speech.

These individual differences in overcoming segmental or suprasegmental deviations may reflect a sensitivity and/or preference for a particular strategy for processing speech in adverse listening conditions. For example, listeners with a preferential reliance on stress cues to inform word segmentation decisions—a critical process in recognizing ambiguous speech—may be at a disadvantage when navigating speech that differs from native norms in the suprasegmental domain. In contrast, word segmentation and recognition may be less affected by suprasegmental deviations for listeners who have a preferential reliance on segmental cues. Identifying words with segmental deviations can require listeners both to cope with sounds that are distorted versions of the target speech sound as well as with those that map to a different phoneme category. Listeners who expand the boundaries of their phoneme categories under appropriate circumstances may be better able to handle phoneme distortions. Recovering from phoneme substitutions may require listeners to suppress mappings that do not result in real words to uncover the intended lexical item. Given the present data, it is not possible to differentiate between a listener's sensitivity to perceiving unfamiliar speech and the various strategies a listener may or may not use to overcome deviations in various linguistic domains; however, this topic should be addressed in future studies.

The ability to overcome other dimensions of deviation may, of course, also be at play. One other possibility is the extent of intraspeaker variability and the nature of the degradation source. On this dimension, it is plausible that, given the presence of irregular articulatory breakdowns, talkers with ataxic dysarthria produce more acoustic-phonetic variability than both Spanish-accented and Irish talkers. It is also conceivable that Spanish-accented talkers are more variable than Irish talkers—a hypothesis backed by reports of greater variability in some measures for nonnative talkers as compared to native talkers (Wade *et al.*, 2007; Hanulikova and Weber, 2012; Baese-Berk and Morrill, 2015). Therefore, on this scale, dysarthric speech and the Irish English may, again, be the least similar speech varieties. Thus, listeners who are skilled at navigating intra-talker and/or inter-talker signal variation are likely to show improved perceptual performance in word recognition for both the dysarthric and Spanish-accented speech varieties, whereas such a skill may be less pertinent to successful word recognition with the more predictable signal that characterizes the Irish speech. Because the talkers in the current study were producing speech naturally, the extent of deviation from native dialect norms was not explicitly controlled for and the talkers differed from typical standard American English in multiple ways, along multiple dimensions. In future work, artificial accents could be created that introduce deviations, to specific degrees, along particular dimensions, while adhering to native dialect standards on all other dimensions. This type of approach has been taken primarily with vowel substitutions (e.g., Adank and Janse, 2010; Banks *et al.*, 2015). However,

constructing multiple artificial accents that deviate along different segmental and suprasegmental dimensions would afford a controlled experimental design to investigate our speculation that listeners may possess specific skills for navigating certain types of acoustic-phonetic deviation. Including other types of challenging listening conditions (e.g., environmental degradation) is also warranted. Furthermore, it is important to note that only two talkers were used to represent each speech variety in the present investigation. The inclusion of two talkers with relatively homogeneous exhibition of the cardinal features of each variety was done to minimize variability arising from multiple talkers. The inclusion of many talkers with different constellations of deviations from the cardinal speech features of the variety may interact with intelligibility outcomes. The generalization of these results should be assessed in future investigations with different talkers or a larger number of talkers.

In addition to investigating the relationship among intelligibility scores on the three speech varieties, three cognitive-linguistic skills and their relationship to the word identification were assessed. Specifically, participants' receptive vocabulary, cognitive flexibility, and inhibitory control of attention were tested. The receptive vocabulary assessment was the only significant direct predictor of performance on the word recognition task. Previous work has demonstrated a link between vocabulary size and perception of speech under a variety of adverse conditions (Adank and Janse, 2010; McAuliffe *et al*., 2013; Tamati *et al*., 2013; Banks *et al*., 2015). The relationship between vocabulary size and word identification accuracy has been shown for two talker-related adverse conditions: dysarthric speech (McAuliffe *et al*., 2013) and constructed unfamiliar accents (Janse and Adank, 2012; Banks *et al*., 2015). Although these studies have demonstrated a relationship between vocabulary knowledge and word identification accuracy, there are methodological differences across the studies that may influence the interpretation of this relationship. Specifically, Janse and Adank (2012) and Banks *et al*. (2015) both used meaningful sentences with speeded sentence verification task in quiet and sentence repetition task in noise, respectively. With meaningful sentences, listeners with greater vocabulary knowledge may be better at using their linguistic and contextual knowledge to predict upcoming elements in a sentence. This ability may assist with perception of unfamiliar accents (Banks *et al*., 2015) as well as adaptation in these conditions (Janse and Adank, 2012). With a sentence verification task, listeners, who are skilled at predicting upcoming words in a sentence, would be quicker to reject sentences that are false. When the production of sentences deviates from native language norms, these prediction abilities may help listeners narrow the range of possible interpretations of lexical items that do not precisely map to previously stored exemplars. In contrast, in both McAuliffe *et al*. (2013) and the current study, semantically anomalous phrases were used. In fact, the two studies used the same anomalous phrases, although in McAuliffe *et al*., participants repeated what they heard rather than typing their responses as in this study. Both studies showed a relationship between word identification and receptive vocabulary. Thus, the current study extends McAuliffe *et al*. by demonstrating the link between receptive vocabulary knowledge and the ability to accurately perceive multiple varieties of speech that deviate from native norms in different ways. Here, vocabulary scores were predictive of word recognition both when assessing the relationship between vocabulary scores and all varieties combined as well as between vocabulary scores and each individual variety (i.e., nonnative, dialect, or disordered). For the anomalous phrases employed in this study, the ability to better predict possible upcoming words would not benefit performance. The reason, therefore, that listeners with larger receptive vocabularies were more adept at perception of anomalous phrases may be related to experience with a greater number of exemplars for each lexical item. Listeners with more exemplars in their mental lexicon may be more proficient at matching incoming signals that deviate from their native dialect norms, even if these incoming signals do not match precisely to any of their previous experiences. Listeners with larger vocabularies may be more flexible in their interpretation of incoming words. Alternatively, it may be that speech perception skill in adverse conditions facilitates the building of the lexicon. That is, listeners who are better able to perceive speech in adverse conditions may be better at learning words from a variety of talkers in a wide range of environments.

An alternative account for this vocabulary size effect is that the vocabulary size measure used here (PPVT) reflects not only receptive vocabulary size, but is also influenced by a number of other factors. The path analysis suggests that the relationship between speech recognition and the measures of cognitive flexibility and inhibitory control of attention tested here is, in fact, mediated by vocabulary size. One may then ask how factors like cognitive flexibility and inhibitory control of attention may impact speech recognition. The mapping of novel pronunciations to lexical items may require a type of linguistic, perceptual flexibility that influenced some of the individual differences seen here. However, this flexibility does not appear to be broad cognitive ability. Participants who showed greater flexibility in the intra-extra-dimensional shift set visual task did not demonstrate better performance in the speech perception task. Cognitive flexibility in Adank and Janse (2010) was related to better performance on a constructed accent. However, their task—the Trail Making task—may have tapped more into language abilities, as letters and numbers are employed. The intra-extra-dimensional shift set task uses visually presented shapes, which do not straightforwardly map onto words. Thus, verbal processing may be limited in the task. It is possible that linguistic flexibility is a separate ability from visual cognitive flexibility. Inhibitory control of attention was also not directly related to performance in the word recognition task in the current study. Yet previous work has demonstrated a connection between inhibitory control of attention and adaptation to a constructed accent using the Stroop task (Banks *et al*., 2015) and the Flanker task (Janse and Adank, 2012). Inhibitory control of attention, however, may be more directly related to perceptual adaptation to unfamiliar speech, rather than overall speech recognition.

The relationship of adaptation to the measured cognitive-linguistic variables could not be assessed because listeners in experiment 2 did not show adaptation during the course of the experimental session. This absence of perceptual adaptation when listeners were presented with multiple, intermixed speech varieties (experiment 2), but presence of perceptual adaptation for nonnative and dysarthric varieties when listeners were presented with a single variety during the course of the experiment (experiment 1) mirrors previous findings. Specifically, in Bradlow and Bent (2008), listeners showed adaptation across the course of a single session with one foreign-accented talker, but did not show adaptation under the same conditions when multiple nonnative talkers from several language backgrounds were employed. Therefore, the adaptation process appears to be restricted when listeners do not have consistent input from one speech variety that allows them to tune in to the specific deviations present in that speech variety. Furthermore, when the listeners were presented with a single speech variety, they only heard two talkers whereas in the experiment in which they heard all three speech varieties, there were six talkers. These six talkers not only varied in the ways in which their speech deviated from native language norms, but also represented a variety of ages. Under the conditions in which there were only two talkers, listeners may have been able to more easily learn the talker specific patterns of deviation from typical, native norms and had more exposure to the specific talkers. In contrast, in the condition with multiple speech varieties, the greater number of talkers, in addition to the multiple speech varieties, may have inhibited perceptual adaptation because listeners would be unable to track how specific talkers deviate from the norm and use that knowledge to assist in the interpretation of novel phrases later in the experiment. However, experimental conditions in which listeners are presented with meaningful sentences multiple times across two days have resulted in adaptation across nonnative talkers with the same accent as well as across different foreign accents (Baese-Berk *et al.*, 2013). Under different experimental conditions, it is possible that listeners could show adaptation in multi-variety conditions and even demonstrate generalization of learning across the speech varieties used in this study, although the varieties used here likely differ more from one another than different foreign accents. This explanation seems particularly likely given that participants did adapt to the Spanish-accented English and the dysarthric speech used in this experiment when they were the only varieties presented during the experiment. The multiple speech varieties, lack of feedback, lack of repetition, greater number of talkers, talkers representing multiple ages, and anomalous stimuli may have inhibited listeners' abilities to perceptually adapt to the unfamiliar speech during the course of the experiment. If learning generalizes across these varieties in future studies, the strategy at play may be one of general expansion rather than the learning of specific acoustic-phonetic patterns. More work will be necessary to determine whether there are experimental conditions that would promote learning across widely different speech varieties, and whether adaptation to any or all of these varieties can be predicted by cognitive-linguistic variables.

V. CONCLUSION

The present study investigated listeners' perception of three speech varieties—Spanish-accented English, Irish English, and ataxic dysarthria. Listeners' intelligibility scores for the Spanish-accented speech were related to their scores for both the Irish English and dysarthric speech; however, intelligibility of Irish English and dysarthric speech did not correlate. These results suggest that listeners may possess skills at navigating specific types of deviations from native dialect norms rather than a more global ability to recover from all acoustic-phonetic deviations, regardless of etiology. Furthermore, receptive vocabulary was the only cognitive-linguistic assessment that significantly predicted intelligibility scores; however, a path analysis revealed indirect effects of cognitive and perceptual flexibility and inhibitory control of attention. The directionality of and mechanism for the relationship between vocabulary and speech perception abilities should be explored further.

ACKNOWLEDGMENTS

We would like to thank our research assistants Lily Huston, Quinten Konyn, Erica Leishman, and Paul Olejarczuk from the University of Oregon and Haley Huffman from Indiana University for their assistance with data collection and analysis as well as Charles Brandt for writing experimental software. This work was partially supported by an Undergraduate Research Grant from the Indiana University Hutton Honors College.

[1]Here we describe the speech varieties as "unfamiliar." However, we acknowledge that our listeners may have encountered talkers representing these varieties previously. The essential characteristic of these varieties is that they are substantially less familiar and commonly encountered compared to the listeners' native dialects.

[2]See supplemental material at http://dx.doi.org/10.1121/1.4966677 for all of the anomalous phrases used in the practice and experimental trials in experiments 1 and 2.

[3]We performed analyses on the experiment 2 data to examine whether any of the cognitive-linguistic skills were significant predictors of performance on any single variety of speech in the experiment. Results were quite similar to those reported here, with only receptive vocabulary scores emerging as significant predictors of performance.

Adank, P., Evans, B. G., Stuart-Smith, J., and Scott, S. K. (**2009**). "Comprehension of familiar and unfamiliar native accents under adverse listening conditions," J. Exp. Psychol.: Human Percept. Perform. **35**, 520–529.

Adank, P., and Janse, E. (**2010**). "Comprehension of a novel accent by young and older listeners," Psychol. Aging **25**, 736–740.

Akeroyd, M. (**2008**). "Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults," Int. J. Audiol. **47**, S53–S71.

Atagi, E., and Bent, T. (**2015**). "Relationship between listeners' nonnative speech recognition and categorization abilities," J. Acoust. Soc. Am. **137**, EL44–EL50.

Baese-Berk, M. M., Bradlow, A. R., and Wright, B. A. (**2013**). "Accent-independent adaptation to foreign accented speech," J. Acoust. Soc. Am. **133**, EL174–EL180.

Baese-Berk, M. M., and Morrill, T. H. (**2015**). "Speaking rate consistency in native and non-native speakers of English," J. Acoust. Soc. Am. **138**, EL223–EL228.

Banks, B., Gowen, E., Munro, K. J., and Adank, P. (**2015**). "Cognitive predictors of perceptual adaptation to accented speech," J. Acoust. Soc. Am. **137**, 2015–2024.

Benichov, J., Cox, L. C., Tun, P. A., and Wingfield, A. (**2012**). "Word recognition within a linguistic context: Effects of age, hearing acuity, verbal ability, and cognitive function," Ear Hear. **33**, 262–268.
Borrie, S. A. (**2015**). "Visual speech information: A help or hindrance in perceptual processing of dysarthric speech," J. Acoust. Soc. Am. **137**, 1473–1480.
Borrie, S. A., Lansford, K. L., and Barrett, T. S. (**2016**). "Rhythm perception and its role in recognition and learning of dysrhythmic speech," J. Speech, Lang., Hear. Res., in press.
Borrie, S. A., McAuliffe, M. J., and Liss, J. M. (**2012a**). "Perceptual learning of dysarthric speech: A review of experimental studies," J. Speech Lang. Hear. Res. **55**, 290–305.
Borrie, S. A., McAuliffe, M. J., Liss, J. M., Kirk, C., O'Beirne, G. A., and Anderson, T. (**2012b**). "Familiarisation conditions and the mechanisms that underlie improved recognition of dysarthric speech," Lang. Cogn. Process. **27**, 1039–1055.
Bradlow, A. R., and Bent, T. (**2008**). "Perceptual adaptation to non-native speech," Cognition **106**, 707–729.
Cambridge Cognition (**2014**). *CANTAB* (Cambridge Cognition Ltd., Cambridge).
Carlisle, R. S. (**1991**). "The influence of environment on vowel epenthesis in Spanish English interphonology," Appl. Linguist. **12**, 76–95.
Darley, F. L., Aronson, A. E., and Brown, J. R. (**1969a**). "Clusters of deviant speech dimensions in dysarthrias," J. Speech Hear. Res. **12**, 462–496.
Darley, F. L., Aronson, A. E., and Brown, J. R. (**1969b**). "Differential diagnostic patterns of dysarthria," J. Speech Hear. Res. **12**, 246–269.
Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A., Taylor, K., and McGettigan, C. (**2005**). "Lexical information drives: Perceptual learning of distorted speech: Evidence from the comprehension of noise-vocoded sentences," J. Exp. Psychol.-Gen. **134**, 222–241.
Dorn, A., O'Reilly, M., and Chasaide, A. N. (**2012**). "A preliminary analysis of speech rhythm in three varieties of Irish (Gaelic)," in *6th International Conference on Speech Prosody* (Shanghai, China).
Duffy, J. R. (**2005**). *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management* (Elsevier Mosby, St. Louis, MO).
Dunn, L. M., and Dunn, D. M. (**2007**). *Peabody Picture Vocabulary Test* (American Guidance Service, Circle Pines, MN).
Flege, J. E. (**1991**). "Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language," J. Acoust. Soc. Am. **89**, 395–411.
Flege, J. E., and Bohn, O.-S. (**1989**). "An instrumental study of vowel reduction and stress placement in Spanish-accented English," Stud. Second Lang. Acquisit. **11**, 35–62.
Flege, J. E., Bohn, O. S., and Jang, S. (**1997**). "Effects of experience on non-native speakers' production and perception of English vowels," J. Phon. **25**, 437–470.
Gilbert, J. L., Tamati, T. N., and Pisoni, D. B. (**2013**). "Development, reliability, and validity of PRESTO: A new high-variability sentence recognition test," J. Am. Acad. Audiol. **24**, 26–36.
Grabe, E., and Low, E. L. (**2002**). "Durational variability in speech and the rhythm class hypothesis," in *Papers in Laboratory Phonology*, edited by N. Warner and C. Gussenhoven (Mouton de Gruyter, Berlin), pp. 515–546.
Hanulikova, A., and Weber, A. (**2012**). "Sink positive: Linguistic experience with the substitutions influences nonnative word recognition," Attn., Percept., Psychophys. **74**, 613–629.
Hickey, R. (**2004**). *A Sound Atlas of Irish English* (Mouton de Gruyter, Berlin, Germany), pp. 43–50.
Hughes, A., Trudgill, P., and Watt, D. (**2012**). *English Accents and Dialects: An Introduction to Social and Regional Varieties of English in the British Isles* (Routledge, London), pp. 141–142.
Imai, S., Walley, A. C., and Flege, J. E. (**2005**). "Lexical frequency and neighborhood density effects on the recognition of native and Spanish-accented words by native English and Spanish listeners," J. Acoust. Soc. Am. **117**, 896–907.
Janse, E., and Adank, P. (**2012**). "Predicting foreign-accent adaptation in older adults," Q. J. Exp. Psychol. **65**, 1563–1585.
Kent, R. D., Weismer, G., Kent, J. F., and Rosenbek, J. C. (**1989**). "Toward phonetic intelligibility testing in dysarthria," J. Speech Hear. Disorders **54**, 482–499.
Kidd, G. R., Watson, C. S., and Gygi, B. (**2007**). "Individual differences in auditory abilities," J. Acoust. Soc. Am. **122**, 418–435.
Liss, J. M., Spitzer, S., Caviness, J. N., Adler, C., and Edwards, B. (**1998**). "Syllabic strength and lexical boundary decisions in the perception of hypokinetic dysarthric speech," J. Acoust. Soc. Am. **104**, 2457–2466.
Mattys, S. L., Davis, M. H., Bradlow, A. R., and Scott, S. K. (**2012**). "Speech recognition in adverse conditions: A review," Lang. Cogn. Process. **27**, 953–978.
McAuliffe, M. J., Gibson, E. M. R., Kerr, S. E., Anderson, T., and LaShell, P. J. (**2013**). "Vocabulary influences older and younger listeners' processing of dysarthric speech," J. Acoust. Soc. Am. **134**, 1358–1368.
Mueller, S. (**2014**). "PEBL: The psychology experiment building language (Version 0.14) [Computer experiment programming language]," Retrieved May 2014 from http://pebl.sourceforge.net (Last viewed February 2015).
Ringle, C. M., Wende, S., and Becker, J.-M. (**2015**). "SmartPLS 3" (SmartPLS, Bönningstedt).
Sereno, J., Lammers, L., and Jongman, A. (**2016**). "The relative contribution of segments and intonation to the perception of foreign-accented speech," Appl. Psycholinguist. **37**, 303–322.
Shah, A. P. (**2003**). "Acoustic indicators of Spanish-accented English," Canadian Acoust. **31**, 42–43.
Sidaras, S. K., Alexander, J. E. D., and Nygaard, L. C. (**2009**). "Perceptual learning of systematic variation in Spanish-accented speech," J. Acoust. Soc. Am. **125**, 3306–3316.
Sumby, W. H., and Pollack, I. (**1954**). "Visual contribution to speech intelligibility in noise," J. Acoust. Soc. Am. **26**, 212–215.
Tamati, T. N., Gilbert, J. L., and Pisoni, D. B. (**2013**). "Some factors underlying individual differences in speech recognition on PRESTO: A first report," J. Am. Acad. Audiol. **24**, 616–634.
Wade, T., Jongman, A., and Sereno, J. (**2007**). "Effects of acoustic variability in the perceptual learning of non-native-accented speech sounds," Phonetica **64**, 122–144.
Watson, C. S., Qiu, W. W., Chamberlain, M. M., and Li, X. F. (**1996**). "Auditory and visual speech perception: Confirmation of a modality-independent source of individual differences in speech recognition," J. Acoust. Soc. Am. **100**, 1153–1162.