Title

Vocabulary knowledge and social contact in L2-accented speech perception: Evidence from low-exposure listeners

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 48(0)

Authors

Zhou, Xinming
Myers, Emily
Heffner, Christopher

Publication Date

2026

Vocabulary knowledge and social contact in L2-accented speech perception: Evidence from low-exposure listeners

Xinming Zhou (xinming.zhou@utexas.edu)[1], Christopher C. Heffner (ccheffne@buffalo.edu)[2] & Emily B. Myers (emily.myers@uconn.edu)[3]

[1]Department of Speech, Language, and Hearing Sciences, The University of Texas at Austin
[2]Department of Communicative Disorders and Sciences, University at Buffalo
[3]Department of Speech, Language, and Hearing Sciences, University of Connecticut

Abstract

Laboratory training improves L2-accented speech recognition, but it is unclear whether everyday exposure produces similar benefits. Usage-based theories propose that lifetime exposure to diverse accents facilitates recognition, whereas resource-dependent models argue that recognition depends primarily on linguistic processing resources. To evaluate these competing hypotheses, we tested 78 young monolingual English adults on their word recognition of Italian-accented English in noise. A multidimensional questionnaire was developed to measure the frequency, duration, amount, and diversity of participants' real-world accent exposure. Receptive vocabulary strongly predicted recognition accuracy ($p < .001$), which remained significant after controlling for social exposure ($p = .003$) in a sample with low exposure to L2-accented speech. Composite measures of social interaction did not independently predict performance. Our findings suggest that receptive vocabulary is a primary predictor of accented speech recognition in noise among monolingual listeners with limited exposure to L2-accented speech, whereas direct effects of real-life interactions were not observed.

Keywords: L2-accented speech perception; word-recognition; speech-in-noise; receptive vocabulary; individual differences

Introduction

English has emerged as a global language for cross-cultural communication (Crystal, 2003; Rao, 2019). The prevalence of English as a second language means that a substantial proportion of English users communicate in English with an L1-influenced accent. L2-accented speech differs from native speech in both segmental and suprasegmental properties (Allen et al., 2003; Nazzi et al., 2003), which may increase the perceptual load for listeners unfamiliar with these accents. Consequently, such listeners often struggle to accurately process both the phonetic (Cho & Ladefoged, 1999) and semantic (Munro & Derwing, 1995) content of speech. Nevertheless, listeners often adapt rapidly, with measurable improvements sometimes occurring within minutes (Clarke & Garrett, 2004). The central theoretical question is whether this rapid adjustment is primarily driven by listeners' prior experiences or by their linguistic processing resources.

The mechanisms underlying flexibility in L2-accented speech perception have been examined from two primary theoretical perspectives. The first perspective, the usage-based account (e.g., Exemplar Theory; Johnson, 1997), posits that speech perception improves as listeners accumulate episodic traces of varying accents. This exposure benefit is documented in studies showing that training with specific accents and exposure to multiple talkers improve recognition (Bradlow & Bent, 2008; Maye et al., 2008). These findings suggest that listeners benefit from learning the overarching acoustic-phonetic properties of an accent. Beyond laboratory settings, this exposure benefit extends to real-world contexts. For example, residing in linguistically diverse environments is associated with improved L2-accented speech recognition (Atagi & Bent, 2014; Porretta et al., 2016). According to Kutlu et al. (2023), speech adaptation occurs throughout the lifetime, and greater linguistic diversity, such as a larger social network, may facilitate more flexible speech categorization. Similarly, Laturnus (2018) reported that individuals with increased social interaction with L2-accented speakers, particularly among family and friends, performed better on accent recognition tasks.

An alternative perspective focuses on resource-dependent models, such as the Ease of Language Understanding (ELU) model (Rönnberg et al., 2013). This model posits that processing distorted speech imposes significant demands on cognitive resources. The ELU model highlights the importance of working memory and lexical access speed. Critically, the quality and breadth of the mental lexicon, as indexed by receptive vocabulary, may also provide essential constraints for speech repair (Davis et al., 2005). Indeed, receptive vocabulary has been shown to predict speech recognition success in adverse listening conditions for both younger (Banks et al., 2015) and older adults (Janse & Adank, 2012). Nevertheless, findings remain inconsistent in the literature. For example, Laturnus (2018) found that exposure, but not vocabulary, predicted recognition, whereas other studies (e.g., McLaughlin et al., 2018) indicate that vocabulary is the dominant factor. Notably, few studies have examined the effects of vocabulary and lifetime L2-accented experiences within the same study, leaving their relative contributions unresolved.

To address these mixed results, the present study aims to clarify the relationship between receptive vocabulary,

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

real-life L2-accented experiences, and L2-accented speech recognition. We extended the work of Laturnus (2018) by employing a multidimensional questionnaire that quantified L2-accented interactions on a continuous scale, capturing speaker type, number of speakers, and lifetime exposure (including international travel; Roberson, 2018). This approach goes beyond the binary high/low exposure classifications used in prior work. By preserving continuous variation in exposure, it allows us to test whether usage-based benefits require a critical threshold of exposure, a possibility that binary classifications cannot address. We embedded unpredictable sentence materials in varying levels of noise to tax listeners' cognitive resources and directly test the predictions of the ELU model.

We hypothesized that: (1) If usage-based mechanisms operate even at low levels of real-world exposure, real-life exposure metrics should independently predict recognition accuracy. (2) If resource-dependent mechanisms are the primary drivers of adaptation in low-exposure populations, receptive vocabulary should remain the primary predictor after controlling for social exposure.

Methods

Participants One hundred and seven participants were recruited from the University of Connecticut Psychological Sciences Participant Pool and Prolific (online). Twenty-nine participants were excluded for incomplete data, multilingualism, or language disorders, resulting in a final sample of 78 participants (M_{age} = 21.69, SD = 4.99; 62 female). The sample size was chosen to align with previous research that examined individual differences in L2-accented speech perception (Heffner & Myers, 2021). All participants self-identified as monolingual first-language speakers of American English with normal hearing and vision. Although pure-tone audiometry was not conducted, extreme signal-to-noise ratio (SNR) thresholds on the accent word recognition task were monitored as a screener for significant, undetected hearing loss and none was observed. This study received approval from the University of Connecticut Institutional Review Board (X21-0137). Participants received either class credit or monetary compensation for their participation. Task performance did not differ significantly between cohorts; however, testing modality was included as a covariate to account for any environment-specific variance. The study comprised four tasks designed to assess how receptive vocabulary and prior exposure to L2-accented speech influence individuals' ability to generalize to a novel accent.

Materials

Accent Experience Questionnaire To move beyond the binary exposure metrics used in previous research (e.g., Laturnus, 2018), a multidimensional questionnaire was developed to quantify experience on continuous scales (see Appendix). This measure captures the intensity and quality of interactions, maximizing the detection of individual differences in the usage-based experience. We assessed L2-accented interaction across three key domains to reflect the cumulative nature of exposure and to evaluate social depth: (1) Timeline: weekly interaction frequency at four life stages (childhood, adolescence, adulthood, current; question c), (2) Social Context: frequency of interaction with specific interlocutors (friends, family, coworkers, teachers; question d), and (3) Diversity: the number of distinct L2-accented speakers encountered weekly (question e; Laturnus, 2018).

Scoring. We converted responses into numerical values using the grading scale provided in the Appendix. Residence data (question a) were excluded from the analysis due to insufficient geographic variability in the sample. We derived a total of seven scores from the questionnaire across four categories:

1. *Travel experience.* Participants reported the country, duration, and frequency of interaction for each international trip. For each trip, we calculated an integrated score by multiplying the duration by the interaction frequency. These values were then summed across all reported trips to determine a total exposure score for each participant.

2. *L2-accented interactions throughout the lifetime.* We assessed interaction frequency at four life stages: before age 10, ages 10 to 18, 18 to present, and currently, yielding four separate scores.

3. *L2-accented interaction scores by speaker type.* We assessed the interaction frequency with four categories of speakers: friends, family members, coworkers, and teachers. These four scores were summed to generate a composite variable representing overall interaction frequency across social domains.

4. *L2-accented interaction scores by number of speakers.* Participants reported the estimated number of L2-accented speakers they interact with regularly, capturing social diversity.

Receptive Vocabulary Task Vocabulary size is a strong predictor of L2-accented speech recognition (e.g., Bent et al., 2016). To index linguistic proficiency, we assessed each participant's receptive vocabulary using a 100 multiple-choice task developed by Drown and colleagues (2023). In each trial, participants saw a target word inside an example sentence and selected the most appropriate definition from four options. Accuracy was calculated as the percentage of correct responses across all 100 trials. Data from one participant were excluded from further analysis due to uniformly rapid response times of less than 500 ms across all trials.

Accented Word Recognition Task To assess individual differences in L2-accented speech recognition, we determined the signal-to-noise ratio (SNR) threshold for each participant using an objective L2-accented speech-in-noise task developed by Heffner and Myers (2021). The target stimuli comprised English sentences produced by a native Italian speaker. A single speaker was selected to reduce the confound of inter-speaker variability and limit familiarity effects. Consistent with Heffner and Myers (2021), the speaker demonstrated naturalistic L2-accented phonetic patterns that did not diminish recognition performance under quiet conditions. To isolate bottom-up acoustic-phonetic processing from top-down semantic restoration, stimuli comprised grammatically correct but semantically anomalous sentences (e.g., "The glass was living to be cool at both flints"). To simulate adverse listening conditions, speech-shaped noise was generated from the long-term average spectrum of the target sentences to provide consistent energetic masking. The task used an adaptive staircase procedure. Following 15 L1-speech practice trials, participants completed 90 experimental trials starting at +8 dB SNR. After each sentence, participants performed a keyword identification task by selecting three keywords from forced-choice pairs, consisting of the target word and phonetically similar foils to ensure objective scoring and virtual piloting without requirement of a live coder. A trial was considered accurate only if all three target words were correctly identified, reducing the chance level to 12.5%. Following a 2-down/1-up rule, the noise level increased (SNR decreased) after two consecutive correct trials and decreased after any incorrect trial. Step sizes were reduced from 6 dB to 2 dB after two reversals.

Scoring. The outcome measure was the average SNR across all reversals, estimating the participant's recognition threshold. Because lower SNRs indicate better noise tolerance, scores were multiplied by -1 to align positively with other performance measures for analysis.

Accent Familiarity Survey To mitigate prior exposure confounds, participants rated Italian accent familiarity on a 5-point Likert scale.

Procedure All participants provided informed consent before the experiment. First, we collected demographic information, including age, gender, language background, highest parental education completed, musical training, and history of language disorders. Participants then completed the following tasks in a fixed order: (1) Accent Experience Questionnaire, (2) Receptive Vocabulary Task, (3) Accented Word Recognition Task, and (4) Accent Familiarity Survey (Table 1). This sequence was implemented to prevent subjective experience reports from influencing performance on objective speech recognition tasks. The Accent Experience Questionnaire focused on retrospective social contact rather than specific familiarity with Italian accents, minimizing potential priming or bias for the subsequent recognition task. Furthermore, the questionnaire was administered first before the cognitively demanding speech-in-noise task. The entire session lasted approximately 40 minutes. Upon completion, participants were redirected to their respective platforms to receive course credit or monetary payment.

Order	Task
1	Accent Experience Questionnaire
2	Receptive Vocabulary Task
3	Accented Word Recognition Task
4	Accent Familiarity Survey

Table 1: Tasks administered in order.

Statistical Analysis Strategy The dependent measure was the average SNR across all reversals for each participant, estimating their threshold for accurate perception of accented speech in noise. Item-level variation could not be modeled due to the single calculated threshold output.

Results

Preliminary Analyses A Welch's t-test on the accent word recognition performance revealed no significant differences ($t(55.96) = 0.02, p = .987$) between participants tested in-person and those recruited online. Furthermore, preliminary screening showed that neither parental highest education level ($p = .34$) nor music training ($p = .45$) significantly predicted recognition performance.

Descriptive Statistics *Accent Experience* Overall, participants reported limited engagement with L2-accented speakers. The Social Context Score (summed across friends, family, and coworkers) averaged 1.68 ($SD = 1.95$; range 0–9). As summarized in Figure 1, interactions were generally infrequent, skewed towards the lower end of the scale. The Diversity Score (number of L2-accented speakers regularly interacted with) also remained low ($M = 0.74, SD = 0.65$), with few participants interacting with a wide variety of speakers. This restricted range of diversity scores in our sample limits our ability to detect usage-based benefits that may only emerge in high-immersion contexts.

Receptive Vocabulary Participants demonstrated a mean accuracy of 0.75 ($SD = 0.075$, range 0.54–0.88).

Accented Word Recognition Performance on the accented word recognition in noise task is summarized in Figure 2. The mean reversed SNR threshold was -2.87 ($SD = 3.14$, range -9.54 to 5.61), suggesting participants could accurately identify keywords when the background noise was 2.87 dB louder than the speech signal. In reverse order, higher values reflect better recognition (tolerance for higher noise levels). Data screening confirmed that all 78 participants performed within 3 standard deviations (SD) of the mean. Therefore, no data points were identified as outliers for exclusion. Furthermore, the data followed a normal distribution (skewness = 0.19), suitable

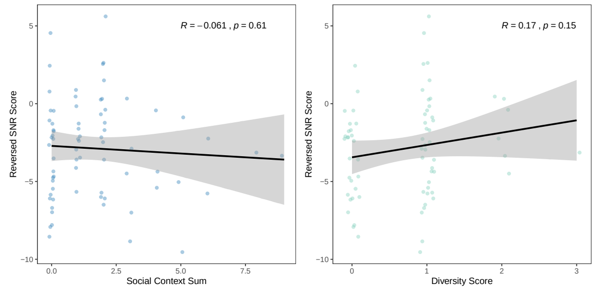

Figure 1: Predictors of individual variability: social context vs. diversity score.

for parametric analysis. This level of performance is consistent with findings in young adults with normal hearing listening to accented speech in challenging acoustic environments (Heffner & Myers, 2021).

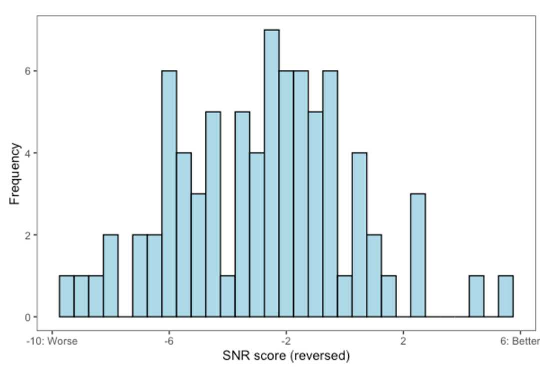

Figure 2: Histogram showing distribution of SNR reversals during the accent comprehension task. The x-axis represents reversed SNR scores.

Accent Familiarity Responses to the post-task survey confirmed that the Italian accent was a valid control for prior familiarity. The majority of the participants (54%) reported being "somewhat familiar" with the accent, while only 3% reported being "very familiar," minimizing potential ceiling effects from prior exposure.

Correlational Analysis To examine the interrelationships among predictors prior to regression modeling, we computed a Pearson correlation matrix (Figure 3). As expected, *Vocabulary Knowledge* showed a moderate positive correlation with accented word recognition performance (Reversed SNR, Figure 4). In contrast, experience-related variables (Social Context Sum and Diversity Score) did not exhibit significant correlations with performance. Furthermore, the correlations between predictor variables were generally low to moderate ($r < .50$), confirming that multicollinearity was not a concern for the subsequent regression analyses.

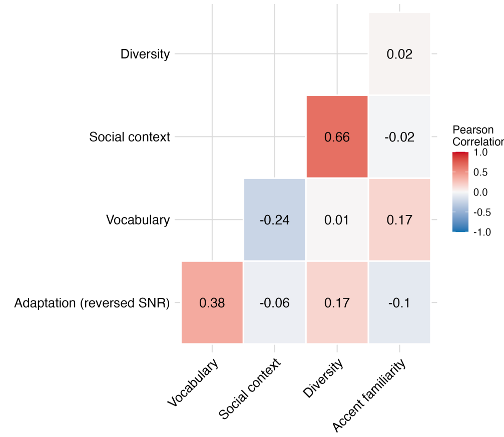

Figure 3: Pearson correlation matrix of all continuous variables included in the regression models. Color intensity represents the strength of the correlation (r).

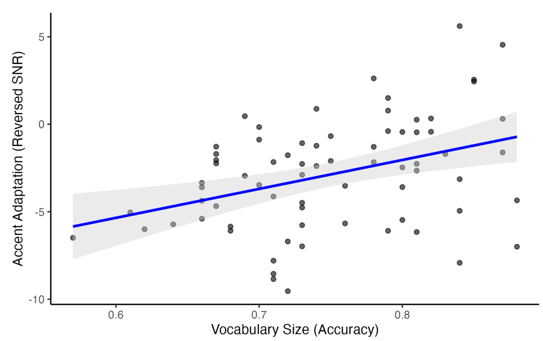

Figure 4: The relationship between vocabulary knowledge and L2-accented speech recognition performance. The shaded area represents the 95% confidence interval.

Regression Analysis A hierarchical linear regression (Table 2) on accented word recognition performance (Reversed SNR) was conducted to evaluate competing hypotheses and address potential confounds.

Step 1: Control Variables and Vocabulary. In the first step, we entered the control variables testing modality (dummy coded: 0 = In-person, 1 = Online) and *Italian Accent Familiarity*, along with the primary predictor of interest, *Vocabulary Knowledge*. The base model significantly predicted recognition performance ($R^2 = .18, F(3, 68) = 4.81, p = .004$). Crucially, testing modality was not a significant predictor ($\beta = -0.41, p = .591$), confirming testing environment (online vs. in-person) did not account for significant variance in performance. Consistent with resource-dependent frameworks, *Vocabulary Knowledge* emerged as a robust predictor ($B = 18.39, p < .001$), suggesting that individuals with larger vocabulary sizes were able to tolerate significantly higher levels of noise.

Step 2: L2-accented Experience. In the second step, we introduced the experience-related variables: *Social Context Sum* and *Diversity Score*. The addition of these usage-based measures did not significantly improve model fit ($\Delta R^2 = .04, F(2,66) = 1.57, p = .215$). A 10-fold cross-validation confirmed the generalizability of the final model ($R^2 = .244$), indicating that the results are not attributable to overfitting. In the final model, neither the frequency of social interaction nor the diversity of speakers accounted for the unique variance in L2-accented speech recognition. However, *Vocabulary Knowledge* remained a significant predictor ($B = 16.36, p = .003$) after controlling for exposure metrics.

Table 2: Hierarchical regression analysis predicting L2-accented speech recognition (Reversed SNR).

Predictor	*B*	*SE B*	*t*	*p*
Step 1				
Testing Modality[a]	-0.41	0.75	-0.54	.591
Italian Familiarity	-0.66	0.50	-1.32	.193
Vocabulary	**18.39**	5.01	**3.67**	**< .001**
Step 2				
Testing Modality[a]	-0.19	0.75	-0.26	.798
Italian Familiarity	-0.70	0.50	-1.42	.162
Vocabulary	**16.36**	5.29	**3.09**	**.003**
Social Context	-0.24	0.25	-0.97	.336
Diversity Score	1.23	0.70	1.75	.085

Note. Step 1: $R^2 = .18$ ($p < .01$). Step 2: $\Delta R^2 = .04$ ($p = .215$). [a]Dummy coded: 0 = In-person, 1 = Online.

Non-linear Regression & Mediation Analysis We used a quadratic term (Social Contact) to test for non-linear threshold effects of social contact. Despite the non-linear model, the fit did not improve significantly, $F(1,66) = 1.573$, $p = .225$. An exploratory mediation analysis was conducted to further explore the relationship between exposure and linguistic resources. While social contact was not directly associated with adaptation ($p = .736$), a significant indirect effect emerged via vocabulary knowledge ($p = .038$).

Exploratory Interaction Analysis To test if vocabulary benefits are conditional on exposure levels, we added interaction terms between vocabulary knowledge and our composite experience measures (Social Context Sum and Diversity Score) into the regression model. Results indicated that these interaction terms did not reach statistical significance ($p > .05$). This finding suggests that the advantage provided by a larger receptive vocabulary is robust and independent across the limited range of social exposure observed in our current sample.

Discussion

Our results indicate that receptive vocabulary, but not real-world exposure, was the primary predictor of L2-accented speech recognition in noise among these low-exposure monolingual listeners. Although accent familiarity and the number of regular L2-accented interactions were included in the regression model, neither emerged as a significant predictor at any step ($p > .05$).

Why does vocabulary predict L2-accented speech recognition? We observed that a larger vocabulary was positively associated with greater tolerance for background noise during L2-accented word recognition. Within the low-exposure sample, this result aligns with resource-dependent frameworks and previous studies indicating that vocabulary knowledge predicts L2-accented speech recognition (Banks et al., 2015; Bent et al., 2016). According to the Ease of Language Understanding (ELU) model, processing of L2-accented speech in noise triggers an explicit retrieval process when the incoming acoustic signal does not match stored phonological representations. This mismatch-driven retrieval process depends critically on the listener's lexical resources. This vocabulary advantage likely operates through several mechanisms. At a structural level, robust lexical networks strengthen word-level lexical constraints, enabling more efficient top-down restoration when phonetic signals are ambiguous due to L2-accented input. Complementing this, greater familiarity with individual words provides more detailed lexical representations, which sharpen the ability to recognize accented variants (Banks et al., 2015; Davis et al., 2005). Beyond specific lexical retrieval, these benefits might also reflect a broader cognitive flexibility (Marchman & Fernald, 2008), which supports adaptation to linguistic variations. In addition to cognitive resources, factors such as parental education and musical experience did not reach significance in our sample. Vocabulary size may instead serve as a proxy for broader language engagement that facilitates more efficient phonological processing. In contrast to the clear role of vocabulary, our experience-based measures did not independently predict recognition in the present study.

Why did real-world exposure fail to predict performance? The absence of a significant exposure effect is inconsistent with a strong version of the usage-based account, which would predict that real-world contact, even at modest levels, should facilitate L2-accented speech recognition. A key divergence between the present findings and previous work (e.g., Laturnus, 2018) concerns the predictive power of social exposure. This discrepancy may partly reflect methodological differences. Laturnus (2018) compared pre-defined high- and low-exposure groups based on questionnaire responses, and the exposure effect was strongest among participants with substantial contact with L2-accented speakers, particularly family and friends. Our sample, by contrast, was characterized by uniformly low levels of self-reported L2-accent exposure, with few participants reporting regular contact. In such a restricted range, the statistical power to detect usage-based benefits is inherently limited. The present sample, consisting of university students and online workers, reported relatively low-to-moderate diversity scores ($M = 0.75$), indicating a predominantly monolingual environment where L2-accented interactions are occasional rather than immersive. Consequently, the exposure effect may be undetectable in this specific low-exposure population. According to usage-based mechanisms, such as Exemplar Theory, adaptation relies on the accumulation of episodic traces; however, these mechanisms may only facilitate recognition once a sufficiently high frequency or intensity of exposure has been achieved. Below this threshold, as observed in our sample, the cognitive system may lack robust episodic traces and may instead rely on internal linguistic processing resources, such as vocabulary, to compensate for signal distortion through top-down restoration (Davis et al., 2005). We speculate that vocabulary serves as a primary compensatory resource for low-exposure monolinguals, while environmental exposure may provide an additive benefit only when interaction is sufficient. Future research across more diverse populations is needed to capture the full picture of how linguistic capacity and real-world exposure interact.

Limitations and Future Directions Several limitations should be acknowledged. A primary limitation concerns the measurement of exposure. As in Laturnus (2018), this study relied on retrospective self-reports to quantify social contact. Although our multidimensional scale provides greater nuance than binary classifications, subjective recall remains susceptible to memory biases and fails to capture the listener's attentional engagement. Future research should consider real-world audio logging to objectively measure exposure. Regarding linguistic resources, while receptive vocabulary emerged as a robust proxy, our design did not explicitly assess other capacities central to the ELU model, such as working memory. Employing a broader cognitive battery would help isolate the unique contributions of lexical quality versus general processing capacity. Additionally, individual differences in vocabulary knowledge may reflect broader socioeconomic factors that are not fully captured in our design. Moreover, the cross-sectional nature of our mediation analysis limits causal conclusions. It remains plausible that individuals with larger vocabularies naturally navigate more diverse social networks, rather than social contact building lexical resources. Longitudinal work is needed to untangle these possibilities.

Task design and stimuli present additional limitations. Although our forced-choice paradigm utilized phonetically similar foils to restrict vocabulary-based guessing, it is fundamentally less cognitively demanding than the open-set transcription tasks used in previous research (e.g., Banks et al., 2015; McLaughlin et al., 2018). Future free-recall studies are needed to verify whether this robust vocabulary advantage persists under higher processing demands that require explicit lexical retrieval. Moreover, while background noise reflects real-world acoustic adversity, our use of semantically anomalous sentences measures phonetic processing capacity rather than overall communicative success in natural, context-rich interactions. Similarly, because stimuli featured only a single Italian speaker to control variability, generalizing these findings to other L2 accents requires caution.

Conclusion This study examined the influence of receptive vocabulary and real-life interactions on L2-accented word recognition. In this sample of low-exposure monolinguals, receptive vocabulary was the strongest predictor of L2-accented word recognition in noise. This suggests that vocabulary knowledge serves as the primary compensatory mechanism for degraded speech. Yet in the present study, neither social interaction frequency nor speaker diversity independently predicted performance. Usage-based accounts (Derwing et al., 2008) propose that diverse exposure contributes to accent recognition. However, this effect was not observed in our low-exposure sample, possibly due to the limited range in reported social diversity. Listeners likely rely primarily on internal cognitive resources to compensate for signal distortion when high-immersion environments are absent. That said, as Kutlu et al. (2023) noted, speech adaptation is a lifelong process, and a diverse social network may facilitate changes that promote more flexible speech categorization over time, provided sufficient experience is accumulated. Future research that samples listeners across a wider range of immersion environments will be essential to determine whether exposure-based benefits emerge beyond a critical threshold, and to clarify how internal linguistic resources and real-world experience jointly shape L2-accented speech recognition.

Funding

This work was supported by a Summer 2021 Undergraduate Research Award from the Institute for the Brain and Cognitive Sciences at the University of Connecticut.

Acknowledgments

The authors used AI-assisted tools for language and writing support. All content was reviewed and finalized by the authors. In addition to the assistance of undergraduate research assistants in the Language and Brain Lab at the University of Connecticut with data collection, the authors also thank LAB current members and alumni for providing methodological and statistical analysis feedback and advice.

Author contributions

Xinming Zhou Conceptualization, Investigation, Methodology, Visualization, Writing – Original Draft Preparation, Review and Editing. **Christopher C. Heffner** Software, Writing – Review and Editing. **Emily B. Myers** Supervision, Writing – Review and Editing.

References

Allen, J. S., Miller, J. L., & DeSteno, D. (2003). Individual talker differences in voice-onset time. *The Journal of the Acoustical Society of America*, *113*(1), 544–552. https://doi.org/10.1121/1.1528172

Atagi, E., & Bent, T. (2014). Auditory free classification of native and nonnative speech by nonnative listeners. *Applied Psycholinguistics*, *37*(2), 241–263. https://doi.org/10.1017/s014271641400054x

Banks, B., Gowen, E., Munro, K. J., & Adank, P. (2015). Cognitive predictors of perceptual adaptation to accented speech. *The Journal of the Acoustical Society of America*, *137*(4), 2015–2024. https://doi.org/10.1121/1.4916265

Bent, T., Baese-Berk, M. M., Borrie, S. A., & McKee, M. (2016). Individual differences in the perception of regional, nonnative, and disordered speech varieties. *Journal of the Acoustical Society of America*, *140*(5), 3775–3786. https://doi.org/10.1121/1.4966677

Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition*, *106*(2), 707–729. https://doi.org/10.1016/j.cognition.2007.04.005

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics*, *27*(2), 207–229. https://doi.org/10.1006/jpho.1999.0094

Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *The Journal of the Acoustical Society of America*, *116*(6), 3647–3658. https://doi.org/10.1121/1.1815131

Crystal, D. (2003). *English as a Global Language*. Cambridge University Press.

Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A., Taylor, K., & McGettigan, C. (2005). Lexical information drives perceptual learning of distorted speech: Evidence from the comprehension of noise-vocoded sentences. *Journal of Experimental Psychology: General*, *134*(2), 222–241. https://doi.org/10.1037/0096-3445.134.2.222

Derwing, T. M., Munro, M. J., & Thomson, R. I. (2008). A longitudinal study of ESL learners' fluency and comprehensibility development. *Applied Linguistics*, *29*(3), 359–380. https://doi.org/10.1093/applin/amm041

Drown, L., Giovannone, N., Pisoni, D. & Theodore, R. (2023). Validation of two measures for assessing English vocabulary knowledge on web-based testing platforms: brief assessments. *Linguistics Vanguard*. https://doi.org/10.1515/lingvan-2022-0116

Heffner, C. C., & Myers, E. B. (2021). Individual differences in phonetic plasticity across native and nonnative contexts. *Journal of Speech, Language, and Hearing Research*, 1–14. https://doi.org/10.1044/2021_jslhr-21-00004

Janse, E., & Adank, P. (2012). Predicting foreign-accent adaptation in older adults. *Quarterly Journal of Experimental Psychology*, *65*(8), 1563–1585. https://doi.org/10.1080/17470218.2012.658822

Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In *Talker variability in speech processing* (pp. 145–165). Academic Press.

Kutlu, E., Chiu, S., Sorensen, E., Oleson, J., & McMurray, B. (2023). Network diversity leads to more flexibility in speech categorization. *In Proceedings of the Annual Meetings of the Cognitive Science Society*.

Laturnus, R. (2018). *Perceptual Adaptation to non-native speech: The effects of bias, exposure, and input variation* [Doctoral dissertation]. New York University.

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, *11*(3), F9–F16. https://doi.org/10.1111/j.1467-7687.2008.00671.x

Maye, J., Aslin, R. N., & Tanenhaus, M. K. (2008). The WECKUD wetch of the WAST: Lexical adaptation to a novel accent. *Cognitive Science*, *32*(3), 543–562. https://doi.org/10.1080/03640210802035357

McLaughlin, D. J., Baese-Berk, M. M., Bent, T., Borrie, S. A., & Van Engen, K. J. (2018). Coping with adversity: Individual differences in the perception of noisy and accented speech. *Attention, Perception, & Psychophysics*, *80*(6), 1559–1570. https://doi.org/10.3758/s13414-018-1537-4

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *45*(1), 73–97. https://doi.org/10.1111/j.1467-1770.1995.tb00963.x

Nazzi, T., Paterson, S., & Karmiloff-Smith, A. (2003). Early word segmentation by infants and toddlers with williams syndrome. *Infancy*, *4*(2), 251–271. https://doi.org/10.1207/s15327078in0402_06

Porretta, V., Tucker, B. V., & Järvikivi, J. (2016). The influence of gradient foreign accentedness and listener experience on word recognition. *Journal of Phonetics*, *58*, 1–21. https://doi.org/10.1016/j.wocn.2016.05.006

Rao, P. S. (2019). The role of English as a global language. *Research Journal of English*, *4*(1), 65-79.

Roberson, D. N. (2018). Learning while traveling: The school of travel. *Journal of Hospitality, Leisure, Sport & Tourism Education*, *22*, 14–18. https://doi.org/10.1016/j.jhlste.2017.11.001

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., Dahlström, O., Signoret, C., Stenfelt, S., Pichora-Fuller, M. K., & Rudner, M. (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Frontiers in systems neuroscience*, *7*, 31. https://doi.org/10.3389/fnsys.2013.00031

Xu, J., & Qiu, X. (2020). Study abroad experiences count: Motivational profile of EFL listeners and its impact on top-down and bottom-up processing. *Applied Linguistics Review*. https://doi.org/10.1515/applirev-2020-0037

Appendix

Accent Experience Questionnaire

a. Where are some places you lived in for a year or longer (at most 4 longest experiences)?

b. Have you ever travelled outside of U.S.? If so, where, how long and how often do you speak English with nonnative speakers (up to 4 longest experiences)?

c. Estimate your weekly experience encountering people with a foreign accent at different points in life {before age of 10, age 10 to 18, 18 to current age, currently in a week}.

d. Currently, in a typical week, how often do you speak with any of the following people who have a foreign accent {Friends, Family member, Co-worker, Teachers} (Choose from Never, 1–2 times a week, 3–4 times a week, 5–6 times a week, and Everyday).

e. Thinking about the people you talk to in a typical week, how many of your friends, family speak with a foreign accent (people you know by name and interact with regularly) (Choose from 0, 1–5, 6–10, and over 10)?

Note. Question (a) and part of question (b) regarding countries of travel excluded from data analysis due to lack of reliable data about geographical variability.

Post-task Survey

a. How familiar are you with Italian accents? (Likert scale of 4 choices ranging from Not familiar at all to Very familiar).

Grading Scales

Accent Experience Questionnaire Grading Scale
Travel Abroad Score = duration × frequency speaking English with nonnative speakers.

Duration of Stay and Frequency (Questions b)

Items/Answer	< 1 mo	1–6 mo	6–12 mo	≥ 1 yr
Points	1	2	3	4

Items/Answer	Never	1–2/wk	3–4/wk	5–6/wk	Daily
Points	0	1	2	3	4

Frequency at Different Stages (Question c) & Social Circle (Question d)

Category	Never (0)	1–2/wk (1)	3–4/wk (2)	5–6/wk (3)	Daily (4)
Before 10 years old	0	1	2	3	4
10 to 18 years old	0	1	2	3	4
18 to now	0	1	2	3	4
Currently in a week	0	1	2	3	4
Friend(s)	0	1	2	3	4
Family member(s)	0	1	2	3	4
Co-worker(s)	0	1	2	3	4
Teacher(s)	0	1	2	3	4

Social Network Size (Question e)

Category	0 (0 pts)	1–5 (1 pt)	6–10 (2 pts)	> 10 (3 pts)
Individuals count	0	1	2	3

Italian Accent Familiarity (Post-task survey)

Accent Type	Not (1)	Slightly (2)	Moderately (3)	Very (4)
Italian Accent	1	2	3	4