CrossMark

BRIEF REPORT

Influences of speech familiarity on immediate perception and final comprehension

Lynn K. Perry[1] · **Emily N. Mech**[2] · **Maryellen C. MacDonald**[3] · **Mark S. Seidenberg**[3]

Abstract Unfamiliar speech—spoken in a familiar language but with an accent different from the listener's—is known to increase comprehension difficulty. However, there is evidence of listeners' rapid adaptation to unfamiliar accents (although perhaps not to the level of familiar accents). This paradox might emerge from prior focus on isolated word perception and/or use of single comprehension measures. We investigated processing of fluent connected speech spoken either in a familiar or unfamiliar accent, using participants' ability to "shadow" the speech as an immediate measure as well as a comprehension test at passage end. Shadowing latencies and errors and comprehension errors increased for Unfamiliar relative to Familiar Speech conditions, especially for relatively informal rather than more academic content. Additionally, there was evidence of less adaptation to Unfamiliar than Familiar Speech. These results suggest that unfamiliar speech imposes costs, especially in the immediate timescale of perceiving speech.

Keywords Speech perception · Accented speech · Speech shadowing · Listening comprehension

As language users, we are confronted with highly variable input as we attempt to understand speakers of different ages, voice qualities, speaking rates, and accents. Investigations of

the impact of this variability on comprehension are informative about the nature of speech perception processes and also could shed light on whether it affects communication across speech communities. We focus on one type of variability, the degree to which a speaker's speech style—accent, speaking rhythm, and other variations—is familiar to the perceiver. There is a fascinating paradox in the literature about effects of speech familiarity on intelligibility. Some previous research has found that nonnative accents disrupt intelligibility of speech, as do regional accents that differ from the perceiver's (Adank, Evans, Stuart-Smith, & Scott, 2009; Floccia, Goslin, Girard, & Konopczynski, 2006; Major, Fitzmaurice, Bunta, & Balasubramanian, 2005). Other research emphasizes that both adults and children can rapidly adapt to unfamiliar speech given sufficient exposure, even over the course of a short experimental session (Clarke & Garrett, 2004; Maye, Aslin, & Tanenhaus, 2008; Sumner, 2011; Sumner & Samuel, 2009; White & Aslin, 2011). One reason for this seeming discrepancy is that adaptation often is found in tasks measuring recognition of isolated words (Sumner, 2011) or identification of the final word of short sentences (Clarke & Garrett, 2004). The nature of adaptation may be very different in continuous speech, however, where there is both more context but also more speech that must be recognized. Additionally, even within the context of continuous meaningful speech, comparing perception of continuous speech at different timescales—both immediate processing and end-state comprehension—could yield different effects.

Suggestive findings were reported by Sabatini (2000), who presented nonstandard English (strongly accented American English and Indian nonnative English) speech to native Italian listeners trained as professional English-language interpreters, measuring accuracy of comprehension, shadowing, which requires participants to repeat aloud the speech they hear as quickly and accurately as possible (Marslen-Wilson, 1973),

✉ Lynn K. Perry
lkperry@miami.edu

[1] Department of Psychology, University of Miami, 5665 Ponce De Leon Blvd, Coral Gables, FL 33146, USA

[2] University of California, Riverside, Riverside, CA, USA

[3] University of Wisconsin-Madison, Madison, WI, USA

and simultaneous interpretation (translating English into Italian out loud). Participants performed more poorly with unfamiliar than familiar speech on all tasks, but comprehension measures yielded better performance than the other two tasks. These differences might reflect variations in task difficulty but also the different timescales of shadowing/interpreting versus comprehension. The longer time scale of comprehension tasks may allow background knowledge and downstream speech information to clarify the interpretation of earlier input, whereas shadowing provides an index of the difficulties listeners encounter in early stages of processing unfamiliar speech. Shadowing tasks offer a continuous fine-grained measure of speech processing that could prove useful in studying perception of unfamiliar speech even by native speakers.

We compared shadowing and offline (end-of-passage) listening comprehension measures to assess perception of continuous speech at two timescales. We manipulated the familiarity of the speaker's accent and the formality of the passages recorded by the speakers. Passages were either informal narratives or more academic in character, written in a more formal style with more technical vocabulary. Academic passages are likely to be more difficult overall than informal narratives, but this effect may interact with speech familiarity. Because informal narratives allow for more variation in speaking style than with more constrained academic prose, we predicted an interaction between speech familiarity and passage type, such that effects of speech familiarity would be greater for informal narratives than for academic prose. Previous sociolinguistic research suggests that speakers adjust phonological, morphological, and lexical aspects of their speech depending on conversational partner and topic (Gumperz, 1958). It is reasonable to expect similar variation in speaking styles with passage formality, which may affect listeners' comprehension difficulty. The manipulation of both speech familiarity and formality of the passage, plus the measure of both shadowing and comprehension, allowed us to investigate how speech quality affects perception of continuous speech in a range of contexts and timescales.

Methods

Participants

A total of 59 (36 females) white, native English speakers who indicated that they were originally from and currently residing in the Midwest United States participated. Twenty-nine participants were assigned to the Familiar Speech condition (16 females), in which they heard passages spoken by a speaker from the same speech community as their own, and 30 to the Unfamiliar Speech condition (20 females). Participants were recruited from introductory psychology classes at the University of Wisconsin-Madison and received course credit for their participation.

Materials

Four text passages, two academic and two informal, were used as stimuli. The academic passages were drawn from reading comprehension portions of the Test of English as a Foreign Language (TOEFL) exam and addressed scientific topics. The informal passages were written transcriptions of two personal narratives from the radio program, "This American Life." Passages were 308-350 words in length.

Each passage was recorded by two female native English speakers, both graduate students in their mid-20s. The speaker in the Familiar Speech condition was a white woman from the upper Midwest, the same region as the participants, whereas the speaker in the Unfamiliar Speech condition was an African American woman from the southeastern United States. The speakers were chosen, because their accents differed markedly. The speakers differ in race but that was not a focus of the study. Speakers were informed that the recordings would be used for a study of speech perception and accents. They read through the passages before making the recordings and were instructed to speak as naturally as possible. As recorded, passage duration was well-matched across speakers (Familiar Speech: $M = 134$ seconds; Unfamiliar Speech: $M = 133$ seconds), $t(11) = 0.71$, $p = 0.496$.

Speech norming Fourteen white participants from the upper midwest who did not participate in the main experiment rated the recordings for similarity to their own speech on a scale from 1 (least similar) to 10 (most similar). Familiar Speech condition ($M = 8$) recordings were rated as more similar to the raters' own speech than were the Unfamiliar Speech condition recordings ($M = 5.7$), $t(13) = 4.00$, $p = 0.002$. To be sure that these ratings reflected speech familiarity and not speaker-specific or recording-specific idiosyncrasies, an additional 11 African American participants, primarily from the southern United States, rated speech familiarity of the recordings. Although these participants only rated the African American English speech recordings ($M=7$) as marginally more similar to their own speech than the white midwestern participants had rated these recordings ($M = 5.7$), t(23) = 1.83, $p = 0.081$, they judged the recordings from the white speaker as significantly less similar to their own speech ($M = 5$) than did white midwestern participants ($M = 8$), $t(18) = 3.78$, $p = 0.001$. These results establish that familiarity of the stimuli depended on the listener's own speech background, suggesting that differences in perceived familiarity are not due to other properties of the recorded stimuli.

Procedure

Speech familiarity (Familiar, Unfamiliar) was manipulated between participants. Task (shadowing, comprehension), and Passage Type (Academic, Informal) were manipulated within participants, with the order of conditions randomized for each participant. Passages were presented over headphones in a quiet lab room. Each participant heard all four passages once, with assignment of conditions counterbalanced across participants. Counterbalancing allowed us to examine adaptation over the course of the task (e.g., were participants who completed the shadowing task second faster and more accurate than participants who completed it first because the prior comprehension task allowed opportunity to adapt to the speech?). All experimental tasks were run using E-prime 2.0 software. The experiment took approximately 20 minutes to complete.

A primary goal was to examine whether familiarity affected speech processing even after exposure to a given speaker over several passages. To accomplish this goal while simultaneously avoiding fatigue effects, we manipulated speech familiarity between subjects to reduce the time spent on the tasks. Manipulating familiarity within subjects will be an important direction for future research.

Shadowing task Participants were instructed to repeat the passage as they were hearing it as quickly and accurately as possible, speaking into a microphone directly in front of them. One Academic passage and one Informal passage were presented. This procedure allowed us to examine how participants process continuous passages of speech in real time. Shadowing is both a listening comprehension task and a speech production task. The important finding from classic shadowing studies was that production was sensitive to linguistic properties of the materials, indicating that they were comprehended as heard. For example, shadowers override anomalous words that have been embedded in the materials based on what has been comprehended; latencies are longer for semantically coherent passages compared with incoherent ones (Marslen-Wilson, 1973, 1975, 1985; Marslen-Wilson & Welsh, 1978). Shadowing allows the use of more naturalistic continuous speech stimuli than in studies that measured the processing of the final word. For these reasons, shadowing offers useful compromise between methodological control and approximation of real-world perceptual experiences.

Listening comprehension task Participants were instructed to listen to each passage and were told their comprehension would be tested afterwards. After hearing each passage, they answered six written true/false questions, presented individually onscreen, by pressing T and F keys on the keyboard. One Academic passage and one Informal passage were presented.

Coding

Shadowing task Two trained research assistants, blind to the experimental hypotheses, coded each participant's speech shadowing for errors and latency. Coders were not blind to speaker condition, but conditions were labeled with speakers' names rather than "familiar" or "unfamiliar" so as not to highlight the hypotheses.

Accuracy Any errors or deviations from the original transcript were coded as omissions, constructive errors, or delivery errors (Marslen-Wilson, 1973, 1985). Omissions were whole words that participants omitted in shadowing. Constructive errors included any added words or changes to words that resulted in a different word or a nonword. Delivery errors included slurred hesitations, stuttering, and unintelligible responses.

Latency Every tenth word of participants' shadowing was coded for latency relative to its occurrence in the recorded passage. Latency was measured using Praat software and was defined as the delay from word onset in the passage to onset of the participant's production.

Listening comprehension Participants' comprehension accuracy was defined as the number of true/false questions answered correctly for each passage.

Analyses

All analyses were conducted using mixed effects regression models. Latencies were analyzed using linear regression, and accuracy analyses were conducted using logistic regression. To determine the best-fit model, we used chi-square tests comparing models with and without the factor of interest. For interactions, we report coefficients and confidence intervals from the full model, and the chi-square test of model fit from the comparison to a model with the interaction removed. For main effects, we report coefficients and confidence intervals from the full model, and the chi-square test of model fit from the comparison to a model with the predictor main effect removed. To determine appropriate random effects, we began with completely specified random effects structures, including random slopes for all variables in a given model. Using model comparison, we systematically removed uninformative random effects (Jaeger, 2009). All final models included random intercepts for subjects and items.

Results

We first report performance for each task separately and then examine the relationships between them.

Shadowing task

The principal analyses concerned the speed and accuracy of shadowing responses as a function of speech condition (Familiar vs. Unfamiliar) and passage type (Academic vs. Informal). Because the analysis of shadowing accuracy focuses on the occurrence of errors, we report the proportion of errors rather than proportion correct both in the text and in related tables and figures.

Shadowing accuracy As shown in Table 1, participants were highly accurate overall but made a larger proportion of errors in the Unfamiliar Speech condition than in the Familiar Speech condition. Although participants in both conditions were generally more likely to make errors for Academic passages than Informal passages, this difference was larger for those in the Familiar Speech condition than in the Unfamiliar Speech condition (Table 1).

Model comparisons revealed a main effect of speech condition, $b = 0.34$, 95% CI [0.02, 0.67]; $X^2(1) = 11.46$, $p < 0.001$, such that participants in the Familiar Speech condition were more accurate than participants in the Unfamiliar Speech condition, and a main effect of passage type, $b = -0.91$, 95% CI [−1.14, −0.68]; $X^2(1) = 8.49$, $p = 0.004$, such that participants were less accurate at shadowing Academic than Informal speech. The interaction between passage type (Academic vs. Informal) and speech condition (Familiar vs. Unfamiliar) also was significant, $b = 0.59$, 95% CI [0.43, 0.75]; $X^2(1) = 51.54$, $p < 0.00001$. This interaction was driven primarily by participants' performance on Informal passages, with better performance in Familiar than Unfamiliar speech in the Informal passages, $b = 0.90$, 95% CI [0.56, 1.24]; $X^2(1) = 23.35$, $p <$

0.001, but only a marginal effect of speech condition on Academic passages, $b = 0.34$, 95% CI [−0.03, 0.71]; $X^2(1) = 3.18$, $p = 0.07$. Additional follow-up comparisons revealed that both the participants in the Familiar Speech condition, $b = -0.34$, 95% CI [−0.46, −0.22]; $X^2(1) = 6.22$, $p = 0.02$, and those in the Unfamiliar Speech condition, $b = -0.92$, 95% CI [−1.21, −0.62]; $X^2(1) = 8.39$, $p = 0.004$, demonstrated a significant effect of passage type, with more errors on Academic than Informal passages. As shown in Table 1, error types were similar across speech conditions and passage types. There was a main effect of error type, $X^2(1) = 588.38$, $p < 0.00001$, with more constructive errors than delivery errors or omissions. There also was an interaction between Speech condition and error type, $X^2(1) = 17.99$, $p = 0.0001$, such that constructive errors occurred more often in the Unfamiliar Speech condition than the Familiar Speech condition.

Finally, we examined whether shadowing accuracy changed over time, which would suggest that participants were able to adapt to the speech, improving performance. Model comparisons revealed that the interaction between speech condition and block was not significant, $X^2(1) = 2.13$, $p = 0.14$, nor was the three-way interaction between speech condition, block, and passage type, $X^2(1) = 0.40$, $p = 0.53$, suggesting that adaptation to the speaker did not affect shadowing accuracy (Table 1).

Shadowing latency In addition to making more errors, participants in the Unfamiliar Speech condition also shadowed more slowly compared with those in the Familiar Speech condition (Table 2). Model comparisons revealed only a marginally significant interaction between speech condition and passage type, $b = 0.05$, 95% CI [−0.005, 0.10]; $X^2(1) = 3.27$, $p =$

Table 1 Mean proportion (SD) of shadowing errors made in each Speech Condition, for each passage type

			Familiar speech Overall 0.06 (0.24)	Unfamiliar speech Overall 0.09 (0.29)
Passage type	Academic	Overall	0.08 (0.28)	0.10 (0.31)
		Constructive	0.04 (0.20)	0.05 (0.22)
		Omission	0.03 (0.17)	0.04 (0.20)
		Delivery	0.01 (0.12)	0.01 (0.11)
		Block 1	0.10 (0.31)	0.11 (0.31)
		Block 2	0.07 (0.25)	0.10 (0.31)
	Informal	Overall	0.04 (0.19)	0.08 (0.27)
		Constructive	0.02 (0.14)	0.04 (0.19)
		Omission	0.01 (0.10)	0.03 (0.17)
		Delivery	0.01 (0.08)	0.01 (0.11)
		Block 1	0.05 (0.21)	0.08 (0.26)
		Block 2	0.03 (0.17)	0.08 (0.27)

Both the overall proportion of errors, the distribution of specific types of errors, and a comparison of error rate for those doing the shadowing task first versus second are reported

Table 2 Mean shadowing latencies in ms. (SD) in each Speech Condition, for each passage type

			Familiar speech Overall 1032 (480)	Unfamiliar speech Overall 1108 (513)
Passage type	Academic	Overall	1069 (509)	1118 (452)
		Block 1	1199 (578)	1192 (571)
		Block 2	968 (421)	1019 (418)
	Informal	Overall	1001 (452)	1098 (509)
		Block 1	1129 (539)	1159 (540)
		Block 2	898 (333)	1026 (460)

Both the overall latency and a comparison of latencies for those doing the shadowing task first versus second are reported

0.07. Planned follow-up comparison revealed that this marginal interaction was driven by participants in the Familiar Speech condition being significantly faster to shadow Informal than Academic passages, $b = -0.07$, 95% CI [−0.14, −0.01]; $X^2(1) = 4.34$, $p = 0.037$. There were no other significant effects.

As shown in Fig. 1, shadowing latencies were shorter if the task was performed after the comprehension task (second block) compared with when the task was performed first, and this difference was larger for those in the Familiar Speech condition than those in the Unfamiliar Speech condition (see Table 2 for means). Given the extant adaptation literature, we investigated the effect of task block on shadowing latencies. The statistical analysis yielded a main effect of block, $b = -0.24$, 95% CI [−0.46, −0.02]; $X^2(1) = 6.43$, $p = 0.011$. Planned follow-up comparisons revealed a significant effect of block for participants in the Familiar Speech condition, $b = -0.24$, 95% CI [−0.46, −0.03]; $X^2(1) = 4.86$, $p = 0.028$, but not in the Unfamiliar Speech condition, $X^2(1) = 2.19$, $p = 0.14$. However, there was no overall interaction between speech

condition and block, $X^2(1) = 0.30$, $p = 0.584$. Thus, for participants shadowing familiar speech, exposure to the speaker's voice during the comprehension task facilitated shadowing; participants shadowing unfamiliar speech did not show this benefit.

Next, we compared shadowing errors with shadowing latency to assess whether there were speed/accuracy tradeoffs. As shown in Fig. 2, participants who shadowed more slowly also tended to make more shadowing errors, such that latency was a significant predictor of shadowing errors, $b = 0.10$, 95% CI [0.07, 0.14]; $X^2(1) = 39.76$, $p < 0.001$. However, there was no interaction between latency and speech condition in predicting the number of speech errors, $X^2(1) = 2.04$, $p = 0.16$. Figure 2 also shows that this relationship was not significantly different for participants in the two speech conditions, demonstrating that there was not a speed/accuracy tradeoff in shadowing. Rather, speech latency and errors are mutually consistent measures of shadowing ability for both familiar and unfamiliar speech.
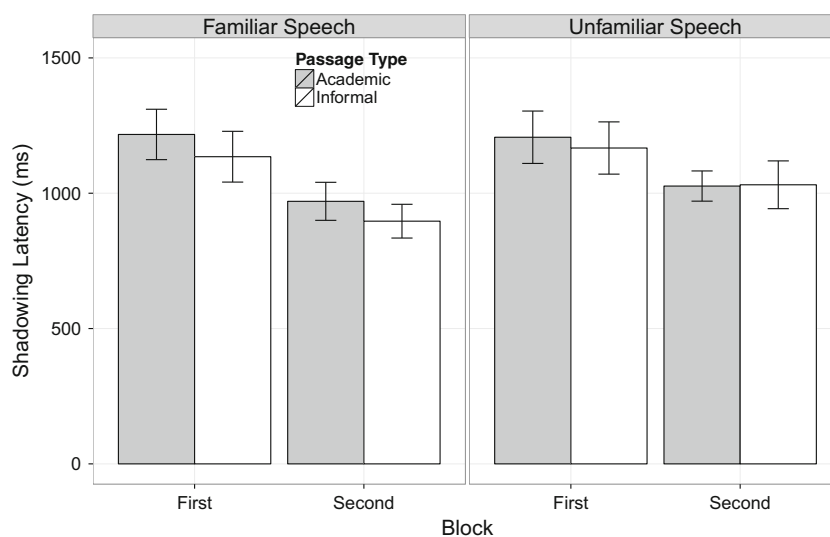

Fig. 1 Latency of participants' shadowing of speech of each passage type depending on whether they did the shadowing task first or second. Within each block, there were two passages, meaning that, e.g., those who did the shadowing task second had heard two passages of the same speaker in the listening comprehension task first. Error bars depict standard error of mean
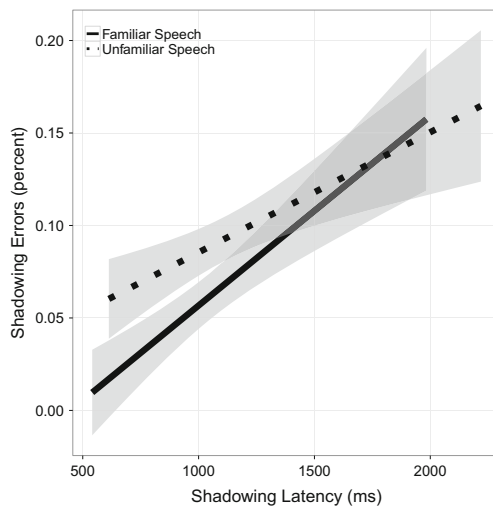
Fig. 2 Relationship between shadowing errors and shadowing latency for participants in each speech condition. Error bands depict standard error of mean

Listening comprehension task

Table 3 depicts the results of the listening comprehension task as proportion of questions out of six that were answered correctly. Overall accuracy was very high. Model comparisons indicated only marginal effects of both speech condition, $b = -0.49$, 95% CI [−1.04, 0.06]; $X^2(1) = 3.02$, $p = 0.082$, and passage type, $b = -0.24$, 95% CI [−0.05, 0.96]; $X^2(1) = 3.01$, $p = 0.083$. The interaction between speech condition and passage type was not significant, $X^2(1) = 0.84$, $p = 0.36$.

We also examined whether listening comprehension changed over the course of the experiment. Accuracy did not differ as a function of whether participants completed the listening comprehension task before or after shadowing (Table 3). Model comparison revealed that there was not a significant interaction between speech condition and block, $X^2(1) = 1.62$, $p = 0.20$, nor was there a main effect of block on listening comprehension, $X^2(1) = 0.03$, $p = 0.86$. Thus, participants' listening comprehension did not improve with prior speech exposure from the shadowing task.

Between task comparison

Finally, we examined the relationships between listening comprehension and shadowing. As shown in Fig. 3, participants with better listening comprehension tended to make fewer speech errors in shadowing; however, listening comprehension was only a marginally significant predictor of shadowing errors, $b = -0.08$, 95% CI [−0.16, 0.003]; $X^2(1) = 3.57$, $p = 0.06$, and listening comprehension and speech condition did not interact, $X^2(1) = 0.40$, $p = 0.53$.

Importantly, there was a main effect of speech condition on shadowing errors, even when both shadowing latency and listening comprehension were included as covariates, $b = 0.02$, 95% CI [0.007, 0.04]; $X^2(1) = 7.39$, $p = 0.008$, suggesting that immediate error assessments (shadowing errors) provide a sensitive measure of effects of speech familiarity.

Discussion

The main finding from this study is that Unfamiliar Speech was more difficult to shadow than Familiar speech, as indicated by longer latencies to produce words and small increases in errors. The impact of Familiarity on speech shadowing was larger for more informal passages than academic passages. These results suggest that when processing natural, connected speech in an immediate timescale (as captured by shadowing), adults have greater difficulty with relatively unfamiliar speech. Over the longer time scale of listening comprehension—at least in the relatively easy listening conditions of our task—speech familiarity effects were not evident.

We also found an interaction between passage type and speech familiarity. It may seem somewhat counterintuitive that informal passages should have exaggerated the effects of familiarity on shadowing. Intuitively, an unfamiliar accent should have a larger effect on more difficult material. However, a speaker's accent and speaking style is not fixed but varies depending on conversational partner and conversational topic (Gumperz, 1958). In our stimulus passages, the

Table 3 Mean listening comprehension accuracy in percent correct (SD) in each Speech Condition, for each passage type

			Familiar speech Overall 0.92 (0.27)	Unfamiliar speech Overall 0.88 (0.34)
Passage type	Academic	Overall	0.90 (0.31)	0.87 (0.34)
		Block 1	0.92 (0.28)	0.85 (0.36)
		Block 2	0.87 (0.34)	0.88 (0.32)
	Informal	Overall	0.95 (0.22)	0.89 (0.31)
		Block 1	0.96 (0.20)	0.88 (0.32)
		Block 2	0.94 (0.24)	0.90 (0.30)

Both the overall accuracy and a comparison of accuracy for those doing the listening comprehension task first versus second are reported
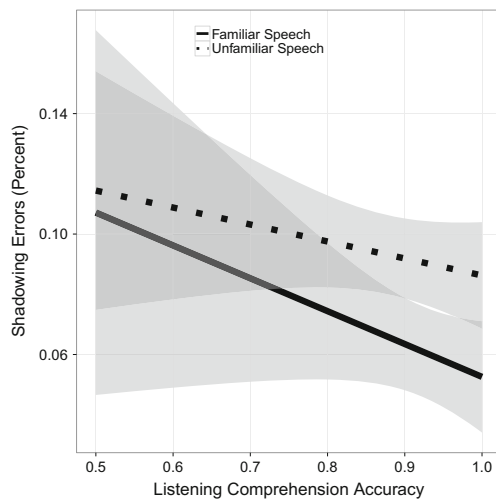
Fig. 3 Relationship between shadowing errors and listening comprehension for participants in each speech condition. Error bands depict standard error of mean

academic prose likely promoted a more formal speaking style, reducing the differences between the speech of the two speakers compared with the differences between them in the informal narrative passages. Quantifying such differences and relating them to comprehension difficulty is an obvious step for future research.

Whereas effects of speech familiarity on shadowing were apparent, there were few effects on comprehension accuracy. It is difficult to design truly challenging short passages and questions to assess comprehension, and thus it is not surprising that comprehension performance was very good. However, the small differences that we found between performances in the two speech conditions are notable. We would expect these small differences to be magnified under more difficult conditions, such as a noisy speech context, a preschool classroom, or crowded restaurant (see Van Engen & Peelle, 2014 for a discussion of accented speech and listening effort).

These results speak to a seeming paradox within the speech perception literature. Participants showed difficulty with the immediate processing of unfamiliar accents. Conversely, they managed well enough to exhibit good comprehension, at least with materials of the complexity used in this study. These findings suggest that the paradox is due in part to the type and timescale of processing demands. Our tasks used meaningful, continuous speech, and we found little evidence of adaptation to unfamiliar speech relative to familiar speech, especially in the immediate timescale of shadowing. Many of the previous studies finding rapid adaptation utilized single word (or single word at the end of a single sentence) measures of comprehension. Our closer analysis comparing processing at two timescales suggests that despite the fact that our participants in the Unfamiliar Speech condition were quite accurate in both the listening comprehension and shadowing tasks,

they still demonstrated a fair amount of difficulty. Of course adaptation to accents does exist, but our work suggests it varies with task demands.

The low levels of adaptation that we found are consistent with findings regarding another form of adaptation, syntactic alignment, namely the degree to which a listener subsequently uses the same sentence structures as a speaker. For example, Weatherholtz et al. (2014) found that participants' degree of syntactic alignment to recorded speech varied with the perceived "standardness" of the speaker's accent and familiarity of their speaking style. In our own task, speech in the Academic passages is arguably more standard and constrained than the Informal passages regardless of speech familiarity. This increased standardness could explain why participants' shadowing of Academic speech was relatively unaffected by speech familiarity. Future research will be needed to assess how speech familiarity changes as a function of speech content. In the following sections, we discuss additional areas for further investigation.

Speaker-specific differences

In the current study, we recorded one speaker per condition. A potential concern is that differences in participants' accuracy for the two speaker conditions could have actually been driven by idiosyncrasies of the particular speakers. However, the results of our norming study suggest this is unlikely. The norming study revealed that people consistently rated the recording of the speaker most like them as sounding more similar to the way they talk. If the participants' shadowing and listening comprehension were driven by the Unfamiliar Speech condition speaker being generally difficult to understand, regardless of the listeners' regional accent and ethnic group, then the norming results would have reflected this difficulty. Instead, we found that in the norming study, African American participants from the South rated the African American Southern speaker (i.e., the Unfamiliar Speech condition) as more similar to themselves than the Caucasian speaker from the Midwest (i.e., the Familiar Speech condition) and vice versa. A fruitful direction for future research will be to use additional speakers, and perhaps manipulate speaker familiarity along a continuum rather than dichotomous conditions. An additional direction for future research will be to use multiple speakers from each of several accents to better understand how people adapt to individual unfamiliar accents rather than unfamiliar speech in a more general way. Nevertheless, the results of the current study are a necessary first step in demonstrating that speaker familiarity has important consequences in speech processing over multiple timescales.

Applications to social and educational settings

The consequences of speaker familiarity are particularly relevant in a classroom context. Our results suggest that it is more difficult to process and comprehend unfamiliar speech, suggesting that having a teacher with unfamiliar manner of speaking could interfere with students' ability to learn material presented orally in class. Speech differences also could contribute to challenges in social and academic contexts, as peers and instructors may alter the quality or duration of their interactions with students with unfamiliar speaking styles, so as to reduce their own processing costs. Given that speech familiarity can vary as a function of a speaker's regional accent, race, and ethnicity (among other factors), and that there is a well-documented achievement gap between ethnic and racial majority and minority group students, speech differences between teachers and students could either add to this achievement gap or perhaps partially explain it. However, the negative consequences both inside and outside of the classroom may be mitigated by the ability of both interlocutors to accommodate their speech to the other or to the "standard". It will clearly be necessary to conduct future research on the role of speech familiarity in classroom learning and interactions.

Role for social influences on speech perception

Future research should also consider how social information modulates participants' processing of familiar and unfamiliar speech. Previous research has noted a role for social influence on speech perception (Babel, 2010; Casasanto, 2008; Kinzler, Dupoux, & Spelke, 2007; Weatherholtz et al., 2014). It will be interesting to explore further how social attitudes towards people from different racial and regional groups influence individuals' abilities to perceive and comprehend their speech. Given that familiarity, race, and region were in the current study, it is possible that participant attitudes may have been contributing to processing difficulties.

Conclusions

Our study is the first to our knowledge to compare listeners' ability to closely shadow and comprehend speech of speakers who both were native speakers of the listeners' language but who varied in regional accent and perceived speech familiarity. By comparing language processing in these two contexts, we are taking an important first step in understanding how listeners process speech that differs from their own. Together our results demonstrate that even when two speakers speak the same language, difference in speaking style can create difficulties in processing at multiple timescales.

References

Adank, P., Evans, B. G., Stuart-Smith, J., & Scott, S. K. (2009). Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *Journal of Experimental Psychology: Human Perception and Performance, 35*(2), 520–529. doi:10.1037/a0013552

Babel, M. (2010). Dialect divergence and convergence in New Zealand English. *Language in Society, 39*(4), 437–456. doi:10.1017/S0047404510000400

Casasanto, L. S. (2008). Does social information influence sentence processing? *Proceedings of the 30th Annual Conference of the Cognitive Science Society, BC Love, K McRae, VM Sloutsky*, 799–804.

Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *Acoustical Society of America, 116*. Retrieved from http://www.bcs.rochester.edu/courses/crsinf/561/archives/s06/clark-garrett_2004.pdf

Floccia, C., Goslin, J., Girard, F., & Konopczynski, G. (2006). Does a regional accent perturb speech processing? *Journal of Experimental Psychology: Human Perception and Performance, 32*(5), 1276–1293. doi:10.1037/0096-1523.32.5.1276

Gumperz, J. J. (1958). Dialect differences and social stratification in a North Indian Village 1. *American Anthropologist, 60*(4), 668–682. doi:10.1525/aa.1958.60.4.02a00050

Jaeger, T. F. (2009). Random effect: Should I stay or should I go? Retrieved February 8, 2013, from http://hlplab.wordpress.com/2009/05/14/random-effect-structure/

Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences, 104*(30), 12577–12580. doi:10.1073/pnas.0705345104

Major, R. C., Fitzmaurice, S. M., Bunta, F., & Balasubramanian, C. (2005). Testing the effects of regional, ethnic, and international dialects of english on listening comprehension. *Language Learning, 55*(1), 37–69. doi:10.1111/j.0023-8333.2005.00289.x

Marslen-Wilson, W. D. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature, 244*(5417), 522–523. doi:10.1038/244522a0

Marslen-Wilson, W. D. (1975). Sentence perception as an interactive parallel process. *Science, 189*(4198), 226–228. doi:10.1126/science.189.4198.226

Marslen-Wilson, W. D. (1985). Speech shadowing and speech comprehension. *Speech Communication, 4*(1), 55–73. doi:10.1016/0167-6393(85)90036-6

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology, 10*(1), 29–63. doi:10.1016/0010-0285(78)90018-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. doi:10.1080/03640210802035357

Sabatini, E. (2000). Listening comprehension, shadowing and simultaneous interpretation of two "non-standard" English speeches. *Interpreting, 5*(1), 25–48. doi:10.1075/intp.5.1.03sab

Sumner, M. (2011). The role of variation in the perception of accented speech. *Cognition, 119*(1), 131–136. doi:10.1016/j.cognition.2010.10.018

Sumner, M., & Samuel, A. G. (2009). The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language, 60*(4), 487–501. doi:10.1016/j.jml.2009.01.001

Van Engen, K. J., & Peelle, J. E. (2014). Listening effort and accented speech. *Frontiers in Human Neuroscience, 8*. doi:10.3389/fnhum.2014.00577

Weatherholtz, K., Campbell-Kibler, K., & Jaeger, T. F. (2014). Socially-mediated syntactic alignment. *Language Variation and Change, 26*(3), 387–420. doi:10.1017/S0954394514000155

White, K. S., & Aslin, R. N. (2011). Adaptation to novel accents by toddlers. *Developmental Science, 14*(2), 372–384. doi:10.1111/j.1467-7687.2010.00986.x