"Does this make sense?": The effect of matching guise in regional accent on grammatical acceptability judgments

Nour Kayali[*]

Abstract. Syntax and sociophonetics are typically treated as wildly disjointed (possibly even incompatible) theoretical pursuits. This paper seeks to unite sociophonetic speech perception and syntax research by presenting participants with matching or mismatching social expectations during a structural grammaticality judgment task. Place is the unifying social association between guise and structure. Participants completed a between-subjects matched guise survey with place-based grammatical structures spoken in either a matching place-based, local accent or a nonlocal accent. Place-based structures are consistently rated more acceptable in the local accent than the nonlocal. These results suggest that judgment of grammaticality results from an interplay of sociocultural expectations with accent and sentence structure. Judgment of structural grammaticality is not independent of social expectation.

Keywords. syntax; speech perception; grammaticality judgment task; language and place.

1. Introduction. Structural grammaticality is traditionally understood as intuitive, the result of an internal sense of grammar and functionally independent of other aspects of language, such as semantics and sociolinguistics (Chomsky 1965). These theories are at the core of generativist models of grammar. In contrast, exemplar theories are non-generativist models that propose that experience shapes an individual's sense of structural grammaticality (Bod 2006) and that experience logs structure alongside other factors of language, such as semantic meaning or social associations (Hay & Bresnan 2006; Squires 2013). Exemplar theories root the nascent body of work on the syntax-phonetics interface, potentially because they offer – and necessitate – a neatly intralinguistically interwoven framework.

The present study seeks to unite sociophonetic speech perception and syntax research by considering the potential role of social expectations. Social expectations affect sociophonetic speech perception (McGowan & Babel 2020), though the extent to which they may touch sentence structure is under-researched. There is evidence of sentence structures indexing certain language varieties in similar ways as accent. Given this, accent and sentence structure may be matched for the same social association.

In this study, place is considered the unifying social association between the chosen syntactic structures and purported social information through a matched guise task. Place is established in sociolinguistic research as a social association that elicits both conscious and subconscious language ideologies (e.g., Carmichael 2016; McGowan & Babel 2020). This study focuses on Southern American English spoken in western Kentucky as the place-based variety – hereafter 'local' – for matched accent and target structures. The target grammatical structures considered are personal datives (i.e., "I got me a new car" rather than "I got myself a new car"; henceforth 'PD') and double modals (i.e., "I might could go with you" rather than "I might be able to go

with you"; henceforth 'DM'). In contrast to the local accent is a nonlocal accent largely unmarked for place; this acts as a mismatched guise with the targeted structures. A set of structurally unremarkable filler sentences (i.e., "I want a cookie"; 'fil' in subsequent figures) appear along with the place-based structures.

By considering auditory stimuli in matching or mismatching conditions, the syntax-phonetics interface is examined in engagement with social associations.

2. Theoretical and methodological considerations.

2.1. EXEMPLAR THEORY. Exemplar theory is a family of usage-based frameworks that exist across different linguistic subfields (Bod & Cochran 2007; Hay & Bresnan 2006). This study considers phonetic and syntactic exemplar theories and their potential unification as a theoretical framework.

Phonetic exemplar theory has been discussed and considered in psychology, speech production, and speech perception (Hintzman 1984; Johnson 1997). Lexical items are stored in an abstract, underlying form as they are experienced; this form includes phonetic detail as well as non-linguistic information, such as the speaker's identity and range of voice quality (Hawkins, 2003). In these models, there is inherent "matching" of linguistic and social-indexical information based on past language experiences. For example, if one is exposed to two variant pronunciations of "cat" – [kʰæt] and [kʰæʔ] – and one hears the former more often from femme-identifying people and the latter from masc-identifying people, those subcategories of "cat" will include both the phonetic details and the perceived gender of the common speaker of that variant (Docherty & Foulkes 2014). Phonological knowledge is highly individualized in this theory, since exemplars are built up from exposure and experience, which vary among even individuals in the same community (see Docherty & Foulkes 2014 for a detailed discussion).

More recent than its phonetic counterpart, syntactic exemplar theory posits that grammar is a product of stored "chunks" of previous language experiences (Bod & Cochran 2007). Chunks may vary in size, from words to whole sentences, and new expressions may be built by combining chunks analogically. As Bod (2006) highlights, exemplar-based syntax differs from the notions of Universal Grammar in generativist linguistics in that it does not propose preexisting rules or understandings of structure— rather, "a statistical ensemble of language experiences" produces knowledge of language. To this point, because exemplar theory suggests that grammar is the product of experience, different exposures naturally lead to different grammars (Hay & Bresnan 2006).

Although developed independent of each other, phonetic and syntactic exemplar theories share common features. Both theories find that linguistic knowledge results from storage of language experiences. Because it is built on experience, one's linguistic knowledge is highly individualized and continually growing and changing; there is also the potential to store additional detail as part of an exemplar, such as social information or semantic meaning.

The primary difference between the two is that phonetic exemplar models focus on classification while syntactic exemplar models focus on composition (Bod & Cochran 2007); by examining different aspects of language, however, there is a possibility of unifying the two, since they complement each other. Unification of these theories has been explored in language acquisition and use (see Bod & Cochran 2007 for more detail), as well as in research on the syntax-phonetics interface (Hay & Bresnan 2006; Squires 2013).

Studies rooted in exemplar theory have varied in methodology used in examining the interaction of syntax and phonetics (see Hay & Bresnan 2006 for a corpus study; see Squires 2013 for

an image-based forced-response task). The current study utilizes a previously underutilized methodology in speech perception: a syntactic acceptability survey.

2.2. SYNTACTIC ACCEPTABILITY WITH AUDITORY STIMULI. Historically, acceptability judgment tasks have relied on text sentences; this practice has recently been scrutinized for its exclusionary features, such as requiring a standardized writing system (Sedarous & Namboodiripad 2020). The bleaching of stimulus presentation to solely written word – excluding presentation of a signing space or speech – also removes social information that may color linguistic processing. Since syntactic acceptability judgments are a tool to better understand structure, ignoring how these structures more commonly appear to language users – as signing or speech – has created a potential gap in the understanding of how language works.

Combining an acceptability judgment task with matched guise, Remirez (2019) examined implications of exemplar theory presented by Sumner et al. (2014). Auditory stimuli for the judgment task varied by speaker accent, and stimulus sentences of variant-specific syntactic structures; acceptability and reaction time were measured. Results indicate both higher acceptability ratings and faster reaction time when hearing a socially congruent accent and structure.

Remirez's experiment focused on guises with established stereotypical social weighting in American culture (Southern Standard British English and African American English; see Sumner et al. 2014 for details). The current study expands on Remirez's findings by focusing on place-based guises familiar to the American South.

2.3. LANGUAGE AND PLACE. Studies of language variation and change have long considered place an influencing factor. Early studies of dialectology and development of linguistic atlas projects acknowledged place as a potential influence on language (Kurath 1972). It is a more recent development in sociolinguistics to consider *relationship to place* and a speaker's sense of identity as factors as well (e.g., Labov 1963; McAndrew 1998; Carmichael 2017; Reed 2018).

The effect of place on a speaker's language and identity is dynamic (Llamas 2007). In the case of Kentucky, Cramer's (2013) examination of language in Louisville, Kentucky highlights the fluidity and multidimensionality of identity construction through language. Participants from Louisville, Kentucky's largest city, style-shifted to present identities that are both Southern and non-Southern. This dynamic sense of identity appears in non-Louisvillian Kentuckians as well; Kentucky is part of the northern border of the typical American South. Due to this, Cramer et. al explained, "residents often feel conflicted about their regional affiliation. For some, Kentucky's Southernness is completely unchallenged; for others, the lure of calling oneself "Midwestern" to avoid being subjected to the stereotypes associated with the South is too tempting" (2018:453). Despite this, "Southern" was the second most common keyword elicited by Kentuckian participants in a mental mapping task (Cramer et. al 2018).

Other common words elicited in that task variably indexed identities that were Southern or Appalachian, a salient geography-based and negatively-stereotyped identity located partially in Kentucky. The Appalachian Mountains run through the eastern half of the state, creating a clear geographic split. Whatever other lines they drew, participants generally distinguished eastern Kentucky – part of Appalachia – as different from the rest of the state. Cramer et. al (2018) suggest that Kentuckians acknowledge negative stereotypes of Southernness and use Appalachia as a scapegoat of those negative stereotypes; in other words, Kentucky has an acceptable kind of Southern (associated largely with western and central Kentucky) and a negative kind of Southern (associated largely with eastern Kentucky).

The current study focused on a western Kentucky variety as the local guise. Western Kentucky is indisputably not Appalachian, and, in perceptual studies, appears in the middle of the

linguistic hierarchy of Kentucky (Cramer et. al 2018). It is also recognizable in Lexington, Kentucky, where the study took place. Previous research links both local structures, personal datives and double modals, to the American South, which western Kentucky typically maps to (see Webelhuth & Dannenberg 2006 and Conroy 2007 for examinations of personal datives; see Fennell & Butters 1996 and Hasty 2011 for examinations of double modals).

2.4. RESEARCH QUESTIONS AND HYPOTHESES. The current study aims to answer the following research questions:

RQ1: What effect does matching guise and structure have on acceptability judgments? When structure and accent match in social associations, is an utterance more acceptable? Conversely, when structure and accent are mismatched in social associations (or one of the two is unmarked), is an utterance less acceptable?

> H1: I predicted that participants would find stimuli with target structures, which are both associated with Southern American English, more acceptable when heard in the local guise, matching listeners' expectations' of the coherence of structure and accent. Conversely, participants would find the target structures less acceptable when heard in the nonlocal guise due to the lack of aligned associations.

RQ2: How are different place-based structures affected by matching and mismatching guises? In what ways are the effects of the matched guise on each structure similar? In what ways are the effects of the mismatched guise on each structure similar?

> H2: I predicted that personal datives in both guises would be found more acceptable than double modals, due to my personal perception that there is greater usage of personal datives (compared to double modals) in popular culture. I predicted that acceptability of both structures (double modals and personal datives) would increase in the local guise and decrease in the nonlocal guise.

3. Experiment.

3.1. PARTICIPANTS. A total of 81 University of Kentucky undergraduate students who have lived in Lexington at least six months participated in this experiment. This requirement worked as a proxy to establish participants had a baseline familiarity with Kentuckian varieties. Four of the 81 also participated in a follow-up interview about their relationship to Kentucky and thoughts on the acceptability task.

3.2. STIMULI. The sentence stimuli consisted of 67 sentences designed to test participants' grammaticality judgments under different accent conditions. 31 of these sentences are structurally unremarkable control sentences; the remaining 36 sentences are split between the two target structures: 16 double modal and 20 personal dative sentences. Sentences were either compiled from multiple resources or constructed for the experiment. Two native speakers of the local variety checked all sentences for naturalness. To ensure that social associations were primarily indexed by accent, I reviewed all sentences for words and topics that may be associated with stereotypical notions of Southernness, such as farming and firearms (see Preston 2018 for an indexical field of "Southerners").

All audio stimuli were produced by a single speaker. The speaker, a cis-gender white man natively from western Kentucky, is bidialectal in his home variety and a less-marked, nonlocal variety. The speaker reviewed sentences for naturalness prior to recording. The speaker then recorded all sentences in one guise before taking a break and recording in the other guise. Sentences were repeated three times. I selected from the three utterances the most natural-sounding version

then extracted them using Praat (Boersma & Weenink 2023). The chosen single utterance became the stimulus for that respective guise and sentence. From these 67 sentences, I created three lists, each a unique, pseudo-random ordering of all stimuli.

Each list was split into three sections. The first and last sets of 24 stimuli, each consisting half of fillers, appear as text sentences; the middle set of 19 stimuli, twelve of which contain target structures, appears as audio. No sentence is repeated as text and audio, so participants only interact with each unique stimulus once.

Block 1
- 24 stimuli
- Text

Block 2
- 19 stimuli
- Audio

Block 3
- 24 stimuli
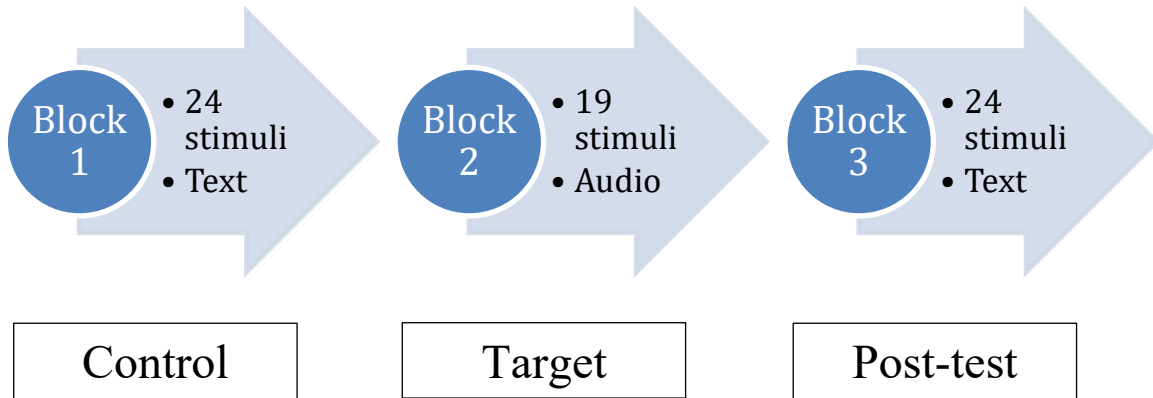- Text

Control   Target   Post-test

Figure 1. Survey flow

3.3. DESIGN. As shown in Figure 1, the survey consisted of three distinct blocks. All blocks presented a single stimulus at a time, required a response, and did not allow participants to return to previously answered questions. Within each block, Qualtrics survey software controlled random presentation of sentences to participants (Medeiros et. al. 2021:430). Block 1 presented 24 stimulus sentences as text; this block acted as a control, gathering participants' baseline intuitions with minimal social information called on outside their personal social presuppositions. Block 2 presented 19 stimulus sentences as audio. Participants could only hear the audio once, which was noted in the instructions, though they chose when to play each clip. Block 3 mirrored the first, presenting the last 24 stimulus sentences as text.

3.4. PROCEDURE. Participants were randomly assigned to (1) one of the three lists and (2) one of the two guises. Qualtrics controlled group assignment and balanced distribution. Guise (local or non-local) only affected Block 2. The local and non-local groups differed by the accent heard in audio stimuli in Block 2. Randomization was handled by Qualtrics within blocks independent of guise.

Participants were randomly presented with one of the three lists and asked to rate each stimulus along a scale of 0 to 100. Only the end points of the scale were labelled in order to give participants flexibility in the remainder of the scale (Jamieson 2020:8). The rating 0 is labeled as "makes no sense" and 100 as "makes sense". These labels attempted to help participants view stimuli more descriptively and overcome socially enforced notions of grammaticality and

associations of non-standardness with "incorrect" language. Participants' comments on how this scale affected judgments are considered further in the discussion.

Prior to Block 1, there were two practice blocks. The first practice was for sound settings. Instructions told participants to adjust sound settings so that a norming audio, which could be re-played, could be heard comfortably. They were also informed that following this block, all audio recordings would only play once. The practice audio thanked participants for their time and reiterated the written instructions. It was recorded by a different speaker than the stimuli. In contrast with the stimulus speaker, this speaker presents as female and non-Kentuckian. The second practice block introduced participants to the rating scale. They were given one filler sentence and a wholly ungrammatical sentence. Instructions stated to rate each sentence 0 if it made no sense and 100 if it makes perfect sense. Participants were told this same scale would be used throughout the survey. Following Block 3, participants completed a demographic survey and self-selected for optional interviews on the contents of the survey. Demographic information collected included age group, preferred pronouns, where they identify as being from, and whether they identified as having a place-based accent (and if so, what accent).

4. Results and Analysis.

4.1. AVERAGE RATINGS BY LOCAL/NONLOCAL GROUP AND BLOCK. For analysis of average ratings, responses are first organized by guise group, which dictated which accent– local or nonlocal– they heard in the audio block. As described in Section 3.4, guise group had no effect on either text block (Blocks 1 and 3) or the content of sentences heard.

	filler		personal dative		double modal	
	local	nonlocal	local	nonlocal	local	nonlocal
Block 1	97.9	95.7	73.5	70.7	48.1	43.6
Block 2	91.4	94.2	79.3	56.3	65.2	46.1
Block 3	98.1	97.2	77.9	70.8	53.0	48.1

Table 1. Average ratings by modality and stimulus type, local and nonlocal

Table 1 presents averages of ratings, organized by modality (pre-text, audio, post-text), stimulus type (fil, pd, dm), and guise heard (local, nonlocal). In the local guise, fil and PD show little change across all three blocks; the largest gap in the fillers is less than eight points and in PD, less than six. DM, however, show more dramatic variation in average ratings by block. The Block 1 average at 48.1 is almost 20 points lower than the audio average of 65.2. The Block 3 average is 53.0– within five points of the pre-text rating but barely within 13 points of the audio.

Average ratings for the nonlocal group differ variably from the local group. The Block 1 row notably does not vary strongly between local and nonlocal ratings; this suggests both groups had, on average, similar baseline intuitions when beginning the survey. Also like the local group, the average ratings of the filler vary little across the three blocks. There is a strong decline – about 14 points – between the PD in Block 1 and Block 2 in the nonlocal group, from 70.7 to 56.3. Despite this, Block 3 ratings are nearly identical to Block 1, with an average of 70.8. Ratings of DM vary slightly across the three blocks, rising by two to two-and-a-half points between each block.

Using the *lme4* package in R (Bates et al. 2015; RStudio Team 2020), I performed a linear mixed effects analysis of the relationship between rating, voice, and stimulus type. I entered stimulus type, voice, and modality type (without interaction term) as fixed effects into the model. I had intercepts for participant and items, as well as by-subject and by-item random for the effect of stimulus type. Visual inspection of residual plots did not show any obvious deviations from homoscedasticity or normality. P-values were found by likelihood ratio tests of the full model with the effect in question (rating by voice and stimulus type) against the model without stimulus type as a fixed effect. Voice and stimulus type affected rating ($\chi^2(1) = 97.64$, $p < 0.001$).
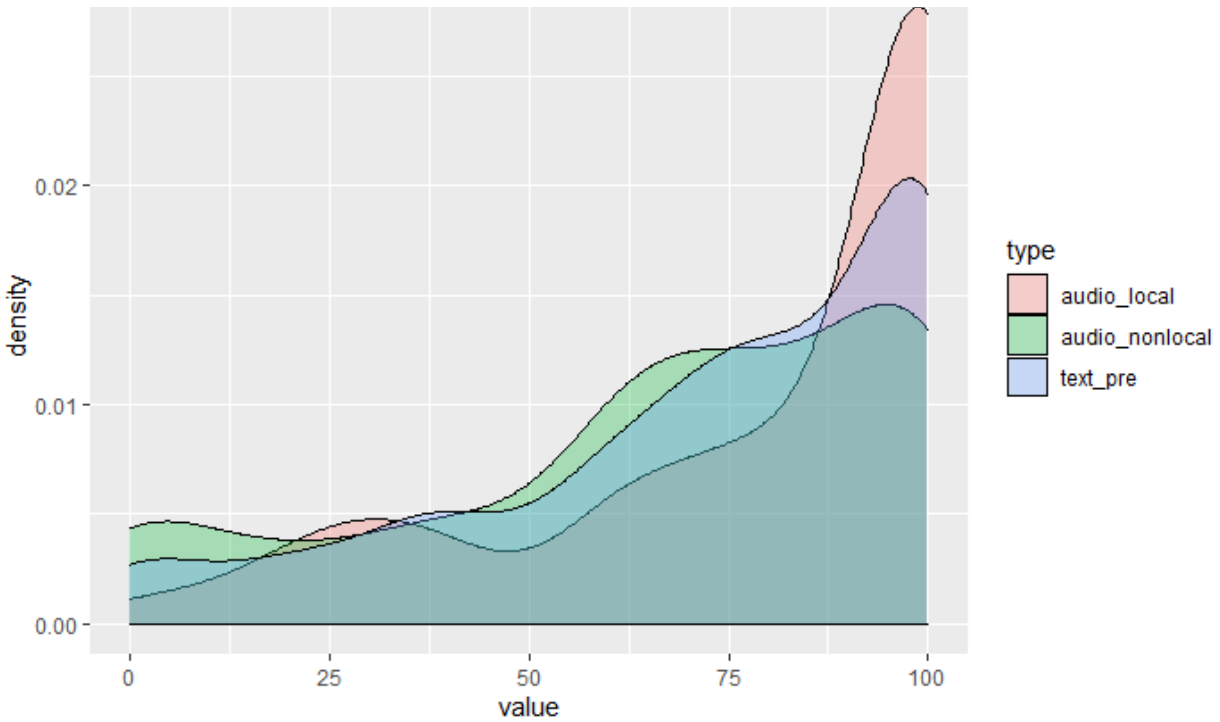

Figure 2. Distribution of ratings of personal datives (PD)

4.2. DISTRIBUTION OF RATINGS OF PERSONAL DATIVES. Ratings of both targets (PD and DM) varied heavily by modality. Figure 2 presents the distribution of ratings for PD across three stimulus types: the pretext block (Block 1), local audio, and nonlocal audio (the varieties of Block 2). All three types peak, with the highest distribution of ratings, at approximately 100; however, the height of each peak differs significantly. The local audio has the greatest density, just beyond 0.025. Following it is the pretext, with a density around 0.02. The nonlocal audio, although still peaking at 100, is more modest at 0.015 – especially compared to its second densest area (0.01) at a rating of 60. This variation suggests that PD match less with a nonlocal accent than when read. This is corroborated by the almost identical average ratings of PD in both text blocks by participants in the nonlocal group (seen in Table 1).
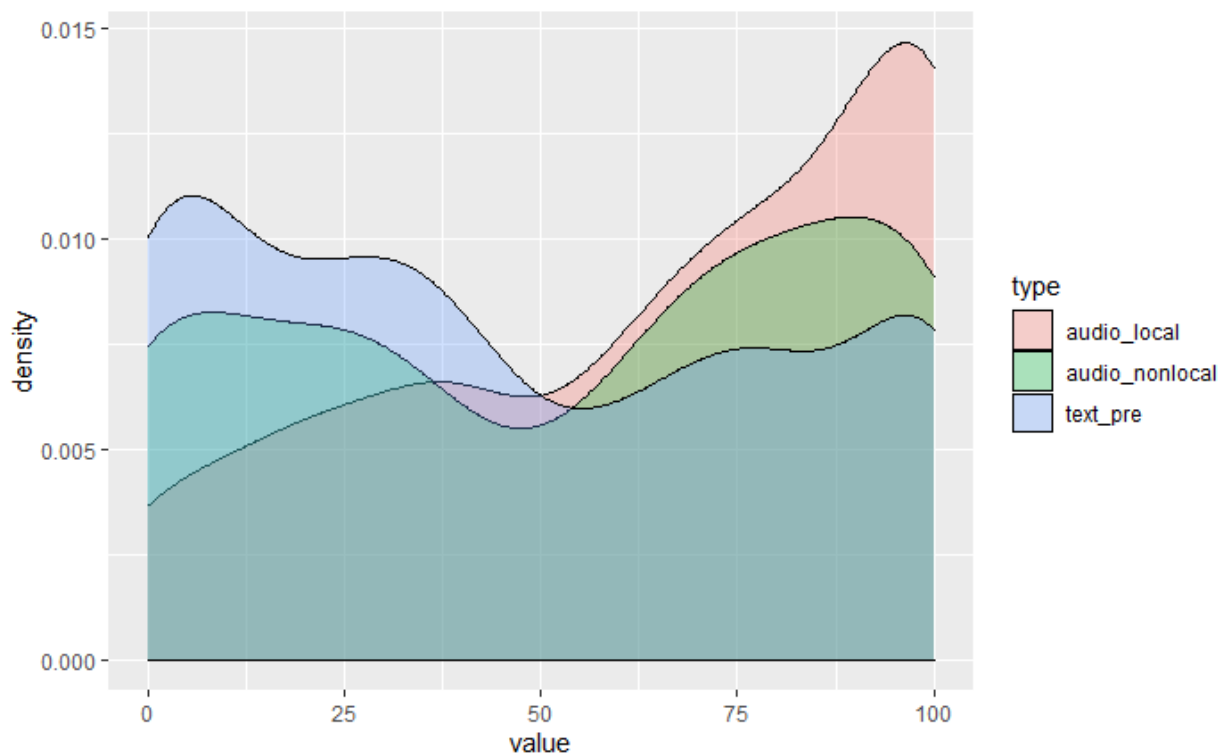
Figure 3. Distribution of ratings of double modals (DM)

4.3. DISTRIBUTION OF RATINGS OF DOUBLE MODALS. The distribution of ratings for DM across the same three stimulus types are shown in Figure 3. Both local and nonlocal audio had the highest distribution of ratings at approximately 100, with the local at a density slightly higher than 0.015 and the nonlocal 0.01. Both waves also noticeably dip at a rating of 50 – the midway point of the scale. This may suggest that hearing DM may be less likely to elicit a neutral judgment. The pretext also drops in density around a rating of 50 – however, unlike both audio blocks, its peak is near the 0 rating. At approximately 0.0125, the most popular rating for DM in the pretext finds DM unacceptable. These points suggest that DM are less acceptable when written than heard, regardless of guise.

5. Discussion. I posited two hypotheses: overall, stimuli of target structures in the local guise would be rated more highly than stimuli in the nonlocal guise (H1) and, in the target structures, PD would be generally more acceptable than DM, though both would have higher acceptability ratings in the local, matched guise (H2) (see Section 2.4 for the fully-stated hypotheses). The data supports H1 and H2. Neither hypothesis, however, considered degree of difference in ratings relative to the control (Block 1) or the post-test (Block 3).

5.1. QUANTITATIVE RESULTS. This study examined two structures of Southern American English, personal datives and double modals, compared across two modalities (text and audio) and, within audio, two guises (a socially congruent local and socially incongruent nonlocal). The control of both structures found differing baseline acceptability between the two, though trends between them differ greatly depending on modality and guise.

In Block 1, both local and nonlocal groups rated PD comparably acceptable. This is in drastic contrast to Block 2: the local audio is rated on average 79.3– about six points higher than Block 1– but the nonlocal audio is rated on average 56.3– 14 points lower than Block 1 and 23

points lower than the local audio. The two groups came back together in the post-text block, though the nonlocal group's rating, at an average of 70.8, is still numerically lower than the local group's at 77.9. The local group's trend line peaks in the audio block, though the difference between each point is undramatic; the nonlocal group's trend line valleys in the audio block instead– rather dramatically, since the dip is over twenty points lower.

This data suggests the following:
1. PD are comparably acceptable in written language and a matched accent.
2. When presented as spoken language, PD as a structure are considerably less acceptable when mismatching the accent of the speaker.
3. Exposure to a mismatched spoken PD has negligible effects on acceptability ratings of read PD.

Figure 3 suggests that DM are less acceptable when written than heard, regardless of guise. This does not mean that all auditory DM are equally acceptable nor that text DM are all lower in acceptability. In Block 1, the local group rated DM on average a bit higher (48.1) than the nonlocal (43.6), though they are within 4.5 points of each other. Block 2 saw a steep change in ratings however: the local audio was rated at an average of 65.2, almost twenty points more acceptable than the nonlocal audio (46.1). The difference in average ratings diminishes in the post-text; the nonlocal group, at 48.1, is still lower in their ratings than the local (53.0). Both groups rate these DM, however, almost five points higher on average than Block 1. The local group's ratings of DM spike at the audio block whereas the nonlocal group steadily ascends with each block.

This data suggests the following:
1. DM are significantly more associated with spoken language than written.
2. When presented as spoken language, DM as a structure are more acceptable when matching the accent of the speaker.
3. Exposure to spoken DM may marginally increase acceptability of written DM.

5.2. QUALITATIVE RESULTS. Immediately following the acceptability task, participants had the opportunity to comment freely on the task, stimuli, or general survey.

Some comments point to stimuli having a reasonable degree of naturalness. For example:

(1) a. Some audio I feel I had less issues understanding what was being said and more of an issue not being sure if I was actually hearing the whole audio. They made sense, but I couldn't tell if I heard it right.
b. A lot of the sentences I have heard people say before. Even if it didn't make sense.

(1a)'s feeling of sensing a disconnect due to a perceived lack of context suggests that the utterances could pass as having greater context, therefore being relatively natural. (1b) further confirms a sense of naturalness in the stimuli, though rather than discussing the auditory experience, they point to the content of the stimuli. The comment on the legibility of the audio is unconcerning. The difference between experiencing uncertainty of "actually hearing" in the acceptability task and in real life is that the task did not allow audio to be replayed; in a live interaction with another person, one may ask for repetition or clarification. It is worth noting that unlike a conversation, all stimuli were single utterances— with no distinct social or contextual relationship drawn to anything else.

One participant noted the effect of the rating scale.

(2) A lot of the sentences I rated high even though I knew they were not grammatically correct because they still made sense.

This response highlights the popular definition of grammaticality and its ties to prescriptive ideologies as well as the significance (and consequences) of how an acceptability scale is labelled. (2) distinguishes "grammatically correct" from "[makes] sense" in a way that suggests that the target features – features of Southern American English, a stigmatized variety (Preston, 1996, 2018; Cramer, 2013) – would be at an immediate disadvantage if participants had been asked to rate 'grammaticality' or 'correctness'.

Studies in perceptual dialectology suggest a view of 'grammatical correctness' that corroborates the decision not to elicit conscious notions of grammar, since those ideas would likely reflect language ideologies – not linguistic intuitions (Cramer et al. 2018). Language ideologies affect speech perception (Lindemann 2002; McGowan 2016). Although ideology and identity play a significant role in this study (see Section 3.4), understanding the role of social associations on the syntax-phonetics interface requires that the language ideologies activated in the judgment task be only those that function on a subconscious level – those that may be stored as a grouping in an exemplar. The effect of subconscious language ideologies unites phonetics and syntax.

6. Conclusions and future work. Judgment of structural grammaticality is not independent of social expectation. The main goal of this study was to investigate the relationship between sociophonetic speech perception and syntax. These results suggest that judgment of grammaticality results from an interplay of sociocultural expectations with accent and sentence structure. The implications of this deviate from theories of syntax separating structure from other aspects of language. Further research may consider the bearing of sociolinguistic elements on syntax as well as how interwoven different aspects of language, such as syntax and phonetics, need to be to function.

References

Bates, Douglas, Martin Mächler, Ben Bolker, & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. https://doi.org/gcrnkw.

Bod, Rens. 2006. Exemplar-based syntax: How to get productivity from examples. *The Linguistic Review* 23(3). 291–320. https://doi.org/10.1515/TLR.2006.012.

Bod, Rens & David Cochran. 2007. Introduction to exemplar-based models of language acquisition and use. In (eds. n/a), *Proceedings of the ESSLLI Workshop "Exemplar-Based Models of Language Acquisition and Use"*. Publisher n/a.

Boersma, Paul & David Weenink. 2023. Praat: doing phonetics by computer. Computer program. Version 6.3.09. https://www.praat.org/.

Campbell-Kibler, Kathryn. 2016. Towards a cognitively realistic model of meaningful sociolinguistic variation. In Anna Babel (ed.), *Awareness and control in sociolinguistic research*. 123–151. Cambridge: Cambridge University Press. https://doi.org/j46d.

Carmichael, Katie. 2016. Place-linked expectations and listener awareness of regional accents. In Anna Babel (ed.), *Awareness and control in sociolinguistic research*. 152–176. Cambridge: Cambridge University Press. https://doi.org/j46f.

Carmichael, Katie. 2017. Displacement and local linguistic practices: R-lessness in post-Katrina Greater New Orleans. *Journal of Sociolinguistics* 21(5). 696–719. https://doi.org/gcmbhc.

Conroy, Anastasia M. 2007. The personal dative in Appalachian English as a reflexive pronoun. *University of Maryland Working Papers in Linguistics* 16. 63–88.

Cramer, Jennifer. 2013. Styles, stereotypes, and the South: Constructing identities at the linguistic border. *American Speech* 88(2). 144–167. https://doi.org/10.1215/00031283-2346753.

Cramer, Jennifer, Susan Tamasi & Paulina Bounds. 2018. Southernness and our linguistic planets of belief: The view from Kentucky. *American Speech* 93(3-4). 445–470. https://doi.org/j46g.

Docherty, Gerard J. & Paul Foulkes. 2013. An evaluation of usage-based approaches to the modelling of sociophonetic variability. *Lingua* 142. 42–56. https://doi.org/gf7ntr.

Fennell, Barbara A. & Ronald R. Butters. 1996. Historical and contemporary distribution of double modals in English. In Edgar Schneider (ed.), *Focus on the USA*. 265–288. Amsterdam: John Benjamins. https://doi.org/10.1075/veaw.g16.14fen.

Hasty, Daniel J. 2011. I might not would say that: A sociolinguistic study of double modal acceptance. *University of Pennsylvania Working Papers in Linguistics* 17(2). 11. https://repository.upenn.edu/pwpl/vol17/iss2/11.

Hay, Jennifer, & Joan Bresnan. 2006. Spoken syntax: The phonetics of giving a hand in New Zealand English. *The Linguistic Review* 23(3). 321–349. https://doi.org/10.1515/TLR.2006.013.

Kurath, Hans. 1972. *Studies in area linguistics*. Bloomington: Indiana University Press. https://doi.org/10.2979/StudiesinAreaLinguis.

Labov, William. 1963. The social motivation of a sound change. *Word* 19(3). 273–309. https://doi.org/10.1080/00437956.1963.11659799.

Lindemann, Stephanie. 2002. Listening with an attitude: A model of native-speaker comprehension of non-native speakers in the United States. *Language in Society* 31(3). 419–441. https://doi.org/10.1017/S0047404502020286.

McAndrew, Francis. 1998. The measurement of 'rootedness' and the prediction of attachment to home-towns in college students. *Journal of Environmental Psychology* 18(4). 409–417. https://doi.org/10.1006/jevp.1998.0078.

McGowan, Kevin B. 2016. Sounding Chinese and Listening Chinese: Awareness and Knowledge in the Laboratory. In Anna Babel (ed.), *Awareness and Control in Sociolinguistics Research*, 25–61. Cambridge: Cambridge University Press. https://doi.org/j46j.

McGowan, Kevin B. & Anna Babel. 2020. Perceiving isn't believing: Divergence in levels of sociolinguistic awareness. *Language in Society* 49(2). 231–256. https://doi.org/ggz8j4.

Preston, Dennis. 1996. Where the worst English is spoken. In Edgar Schneider (ed.), *Focus on the USA*. 297–360. Amsterdam: John Benjamins. https://doi.org/10.1075/veaw.g16.16pre.

Preston, Dennis. 2018. Changing research on the changing perceptions of Southern US English. *American Speech* 93(3-4). 471–496. https://doi.org/10.1215/00031283-7271283.

Reed, Paul E. 2018. The importance of Appalachian identity: A case study in rootedness. *American Speech* 93(3-4). 409–424. https://doi.org/10.1215/00031283-7271250.

Reed, Paul E. 2020. Place and language: Links between speech, region, and connection to place. *Wiley Interdisciplinary Reviews: Cognitive Science* 11(3). e1524. https://doi.org/j46k.

RStudio Team. 2020. RStudio: Integrated Development for R. RStudio. http://www.rstudio.com/.

Remirez, Emily. 2019. Phonetic cues influence judgment of syntax. *The Journal of the Acoustical Society of America*. 146(4). 3055–3055. https://doi.org/10.1121/1.5137596.

Sedarous, Yourdanis & Savithry Namboodiripad. 2020. Using audio stimuli in acceptability judgment experiments. *Language and Linguistics Compass* 14(8). e12377. https://doi.org/10.1111/lnc3.12377

Squires, Lauren. 2013. It don't go both ways: Limited bidirectionality in sociolinguistic perception. *Journal of Sociolinguistics* 17(2). 200–237. https://doi.org/10.1111/josl.12025.

Sumner, Meghan, Seung Kyung Kim, Ed King & Kevin B. McGowan. 2014. The socially weighted encoding of spoken words: A dual-route approach to speech perception. *Frontiers in Psychology* 4. 01015. https://doi.org/10.3389/fpsyg.2013.01015

Webelhuth, Gert & Clare J. Dannenberg. 2006. Southern American English personal datives: The theoretical significance of dialectal variation. *American Speech* 81(1). 31–55. https://doi.org/dvv27r.