Readers generalize adaptation to newly-encountered dialectal structures to other unfamiliar structures

Scott H. Fraundorf

University of Pittsburgh

T. Florian Jaeger

University of Rochester

Author Note

Scott H. Fraundorf, University of Pittsburgh; T. Florian Jaeger, University of Rochester.

This work was supported by National Science Foundation CAREER award (IIS-1150028) and National Institutes of Health grant (NIHCD R01 HD075797) to T. Florian Jaeger. We thank the members of the Human Language Processing Lab for their comments and suggestions, Andrew Watts for programming and data collection, and Alison Trude and Martha Fraundorf for assistance with stimulus construction. Study 2, in particular, benefited from and was inspired by conversations with Julie Boland and Michael Kaschak at, and after, the CUNY 2014 and CUNY 2015 Conferences on Human Sentence Processing.

Correspondence concerning this article should be addressed to Scott H. Fraundorf, Learning Research and Development Center and Department of Psychology, University of Pittsburgh, 3939 O'Hara Street, Pittsburgh, PA 15260. E-mail: scottfraundorf@gmail.com

Abstract

Growing evidence suggests that syntactic processing may be guided in part by expectations about the statistics of the input that comprehenders have encountered; however, these statistics and even the syntactic structures themselves vary from situation to situation. Some recent work suggests that readers can adapt to variability in the frequencies of known, but infrequent syntactic structures. But, the relation between adaptation to altered frequencies of familiar structures and learning to process unfamiliar, never-before-seen structures is under-explored. In two self-paced reading experiments, we investigated readers' adaptation to an unfamiliar structure used in some regional dialects of American English: the *needs*+past participle structure, such as using *The car needs washed* to mean *The car needs to be washed*. Study 1 used a novel Web-based recruitment method to target regions where participants were likely to be familiar (Ohio, western Pennsylvania) or unfamiliar (Colorado) with the *needs*+past participle structure. Participants unfamiliar with the structure initially read the structure more slowly, but over the course of the experiment came to read it more like the familiar participants. Study 2 further demonstrated that participants who have adapted to *needs*+past participle generalize this adaptation to a different, but related structure. These results suggest (a) that readers adapt to unfamiliar syntactic structures, (b) that, in doing so, they become more like existing users of those structures, and (c) that they can generalize this other structures that they may also be more likely to encounter. We discuss these results in the context of implicit learning accounts of exposure effects on syntactic processing.

Keywords: adaptation; sentence processing; dialect syntax; implicit learning; learning novel syntax

Understanding a sentence involves mapping the sentence's linguistic form onto the meaning the speaker intends to convey. But, the same meaning is not always expressed using the same structure: A speaker may say either *I gave him the book* or *I gave the book to him* to describe the same event. And, the meanings of newly encountered structures may be unclear (e.g., the use of *He be home* to mean *He tends to be home* in African American Vernacular English). A challenge for psycholinguistic theories, then, is to account for why we can often quickly and accurately process the syntax of even unfamiliar sentences.

An increasing accepted view of syntactic processing is that language users acquire the relevant frequencies of particular syntactic structures based on experience with the input statistics (Arai & Keller, 2013; DeLong, Urbach & Kutas, 2005; Dikker & Pylkkänen, 2013; Hale, 2001; Jurafsky, 1996; Levy, 2008; Kamide, Altmann, & Haywood, 2003; MacDonald, Pearlmutter, & Seidenberg, 1994; McRae, Spivey-Knowlton, & Tanenhaus, 1998; Staub & Clifton, 2006; MacDonald, 2013). For example, English speakers may understand both the double-object (*I gave him the book*) and prepositional dative (*I gave the book to him*), but know that the double-object is more common (Bresnan, 2007). Knowing which structures are relatively likely enables fast and accurate comprehension because it allows comprehenders to rapidly interpret the unfolding input and perhaps even to predict upcoming structures (Levy, 2008; MacDonald et al., 1994; MacDonald, 2013; Smith & Levy, 2013), consistent with proposals of a broad role for prediction in language comprehension (Dell & Chang, 2014; Federmeier, 2007; for review, see Kuperberg & Jaeger, 2016).

But, expectations about syntactic distributions are only useful to the extent that they match the statistics of the linguistic input (Fine et al., 2013), and these statistics vary across situations, such as different dialects, idiolects, and sociolects (e.g., Bresnan & Ford, 2010;

Finegan & Biber, 2001; Labov, 1969; Tagliamonte & Smith, 2005; Tagliamonte, Smith, & Lawrence, 2005; Weiner & Labov, 1983). This variability includes both differences in the relative frequency of structures (for instance, U.S. and New Zealand English differ in the frequency of the double-object structure Bresnan & Hay, 2008) and in the presence of entirely different structures (such as the use of, e.g., *The car needs washed* to mean *The car needs to be washed* in western Pennsylvanian English; Doyle, 2008; Murray, Frazer, & Smith, 1996; Tenny, 1998). In the face of such variability, previously acquired knowledge about syntactic distributions may be useless or even misleading.

One potential way that the language comprehension system may cope with such variability is by *adapting* to the current situation. Work on structural priming in comprehension suggests that syntactic processing is indeed sensitive to which of several *known* structures have been recently encountered, a phenomenon that has been attributed to implicit learning about the distribution of syntactic structures. This account of structural priming implies that comprehenders should also be able to learn a distribution that contains entirely new structures. In two studies, we test how comprehenders adapt to, and *generalize*, experience with unfamiliar syntactic structures. We begin by summarizing what the literature has revealed about how exposure to known structures facilitates their subsequent processing and whether similar effects might emerge after exposure to novel structures. We then elaborate on the critical role of generalization during exposure to novel structures, which has received little to no attention in previous work.

Recent Exposure Can Facilitate Processing of Known Structures

Beginning with Levelt and Kelter (1982) and Bock (1986), numerous experiments have revealed that experience reading or speaking a particular structure biases speakers to *produce* it

again (for review, Pickering & V. Ferreira, 2008). More recently, it has been shown that recent experience with a structure also speeds its subsequent *comprehension* (Arai & Mazuka, 2014; Fine et al., 2013; Thothathiri & Snedeker, 2008; Tooley & Bock, 2014; Tooley, Traxler, & Swaab, 2009; Traxler, 2008). This facilitation occurs with even just one exposure (Arai, van Gompel, & Scheepers, 2007; Branigan, Pickering, & McLean, 2005; Traxler, 2008) and accumulates with repeated exposure (Farmer, Fine, Yan, Cheimariou, & Jaeger, 2014; Fine, Qian, Jaeger, & Jacobs, 2010; Fine et al., 2013; see also Kamide, 2012).

Both types of facilitation have been attributed to *implicit learning* (Fine & Jaeger, 2013; Fine et al., 2013), building on implicit learning accounts of exposure effects in language production (Bock & Griffin, 2000; Chang, Dell, & Bock, 2006; V. Ferreira, Bock, Wilson, & Cohen, 2008; Kaschak, Kutta, & Jones, 2011; Jaeger & Snider, 2013; Reitter, Keller, & Moore, 2011). Alternative accounts attribute trial-to-trial facilitation, also referred to as *structural* or *syntactic priming*, to short-term boosts in the activation of recently processed structures (Pickering & Branigan, 1998; Traxler & Tooley, 2008); in these accounts, the cumulative effects observed in recent studies require another, different explanation (for further discussion, see Fine & Jaeger, 2016). One piece of support for the implicit-learning hypothesis comes from the observation that structural priming in comprehension seems to be sensitive to how unexpected the prime was: Less expected structures show greater priming in comprehension (Arai & Mazuka, 2014; Fine & Jaeger, 2013) and in production (Bernolet & Hartsuiker, 2010; Jaeger & Snider, 2013). Sensitivity to the unexpectedness (or prediction error) of a prime is expected under error-based and similar learning accounts (Chang et al., 2006; Dell & Chang, 2014; Jaeger & Snider, 2013). Specifically, there is some evidence that exposure to particular structures may lead comprehenders to implicitly adapt expectations about the distribution of syntactic structures

(Farmer et al., 2014; Fine et al., 2010, 2013). Evidence that the facilitatory effects are indeed due to changes in expectations comes from visual world eye-tracking studies, in which exposure to talkers with different syntactic preferences leads to talker-specific expectations about upcoming syntactic structures (Arai & Mazuka, 2013; Kamide, 2012). Finally, repeated exposure can produce facilitation that is detectable several days later when comprehenders are assessed in the same environment (Wells, Christiansen, Race, Acheson, & MacDonald, 2009), consistent with the idea that implicit learning allows comprehenders to adapt their expectations for that environment.

Learning to Process Unfamiliar Structures

Thus, it appears that comprehenders can sometimes rapidly adapt their expectations about the relative frequency of *known* structures based on recent exposure. But, as comprehenders navigate varied linguistic environments, they do not just encounter altered frequencies of *known* structures; they also sometimes encounter *new* structures. For instance, comprehenders may encounter a dialect that includes idiosyncratic structures not part of their own dialect, such as using *The car needs washed* to mean *The car needs to be washed* (western Pennsylvania) or *habitual be* (African American Vernacular English). Learning to comprehend these unfamiliar structures is important for mastering the form-to-meaning mappings of a newly-encountered dialect.

However, it is not clear *a priori* that mechanisms for priming of known structures also support priming newly-encountered structures. Under an implicit learning account, it is certainly plausible that comprehenders could also learn about, and adapt to, entirely new structures since they are also part of the distribution of syntactic structures that comprehenders are experiencing. And, the ability of adults to learn syntactic structures has also been demonstrated for entirely,

novel miniature languages in the artificial grammar learning paradigm; in some cases, learners can even track multiple "dialects" (Gebhart, Aslin, & Newport, 2009). However, the ability to learn genuine, attested dialectal structures that are tied to a comprehender's rich native-language knowledge has received little attention. It is possible that the implicit learning measures in syntactic processing might be limited to adjusting processing of known structures rather than creating new structural representations.

Some evidence does suggest that people can rapidly learn to comprehend unfamiliar structures. Kaschak and Glenberg (2004) tested learning of the western Pennsylvania Dialectal *Need* structure (further described below) by subjects previously unfamiliar with it. Participants experimentally assigned to receive exposure to this dialectal structure were later faster at reading it than participants encountering the structure for the very first time, suggesting they had learned something about processing it (see also Kaschak, 2006, and ongoing work first presented as Boland, de los Santos, Carranza, & Kaschak, 2015). However, these studies did not contain any comparisons to native users of the Dialectal *Need* structure. Thus, it is unclear if participants were truly learning to comprehend the structure the way a native user would or if they were instead acquiring some unusual task-specific strategy, such as becoming more accommodating of unusual or erroneous input (Kim & Mauner, 2006; cf., Traxler & Tooley, 2008).

How comprehenders adapt to novel syntactic input also speaks more generally to the mechanisms of structural priming. If comprehenders adapt to novel syntactic input in ways similar to the adaptation observed for familiar input, that could be readily accommodated by implicit learning accounts. But it is less clear that such a pattern could be easily explained by accounts based on short-term activation because it is uncertain if transient activation-based priming applies to structures not grammatical to the comprehender (for discussion, see

Hofmeister, Jaeger, Arnon, Sag, & Snider, 2013).

Generalizing Syntactic Learning

A second challenge for syntactic processing is to *generalize* from some limited exposure to build expectations for further input from the same talker (or writer). Comprehenders' exposure to any given talker or writer is comprised of only a finite, specific set of observations. Using this information to guide future processing requires a (possibly implicit) inference about how those observations can be generalized to a set of expectations about what the talker or writer will produce in the future—including, perhaps, differences that one has not specifically observed. Thus, for instance, in speech perception, generalization from a limited sample of a talker's speech to broader expectations about the talker is considered critical for effective processing (Kleinschmidt & Jaeger, 2015), and listeners have indeed demonstrated such an ability (e.g., Best & Shaw, & Clancy, 2013; Bradlow & Bent, 2008; Sidaras, Alexander, & Nygaard, 2009; Ying, Shaw, & Best, 2013; for a review of this evidence, see Weatherholtz & Jaeger, in press). Given that comprehenders also have only limited exposure to a talker or writer's *syntactic* preferences, syntactic processing could also benefit from the ability to generalize from exposure to one structure to other structures that are similar or related.

As we have reviewed above, a general account of syntactic processing is that comprehenders expect possible structures to differing degrees based on the statistics of the input (Hale, 2001; Jurafsky, 1996; Levy, 2008); as a result, structures are difficult to the extent they are unexpected (as evidenced in, e.g., Boston, Hale, Vasishth, & Kliegl, 2010; Demberg & Keller, 2008; Jurafsky, 1996; Levy, 2008; MacDonald, 1999). But, it is an open question how these expectations are structured with respect to each other. Researchers have typically tested the difficulty only of processing single structures considered in isolation.

It is plausible that syntactic knowledge may be further organized in terms of which features *covary* in the indexical structure of the world (for related ideas, see Kleinschmidt & Jaeger, 2015; Pajak, Fine, Kleinschmidt, & Jaeger, 2016). Syntactic and other linguistic frequencies do not vary arbitrarily, but often are conditioned as a function of environmental features such as dialect or register (for a quantification of this dependence, see Weatherholtz, Seifeldin, Kleinschmidt, Kurumada & Jaeger, 2016). If recent exposure indeed leads to *learning* about a *distribution* of syntactic structures—not merely one structure taken alone—it is conceivable that exposure to one structure could thus lead to inferences about syntactic structures beyond the specific one that comprehenders just experienced. Indeed, the set of structures to which learners generalize can reveal which structures are organized together in syntactic knowledge. (A similar approach has been taken in research on phonological learning, where the presence or absence of generalization to novel items has been used to infer the existence and nature of abstract categories, e.g., among many others, Finley, 2011; Finley & Badecker, 2009; Pajak & Levy, 2014.)

Continuing the example of an unfamiliar dialect, imagine a speaker of Standard American English who encounters for the first time the Western Pennsylvania English sentence *The car needs washed*. At this point, the comprehender is likely to realize that he or she is listening to an unfamiliar variant of English. The language processing system might then infer—if it is sensitive to the contingencies between syntactic frequencies—a higher probability of encountering other unfamiliar structures and a lower probability of encountering familiar ones. There is some evidence that this property holds for speech perception: Learning how a talker produces one consonant or vowel changes expectations about how that talker produces other sounds that tend to vary in the same way (Maye, Aslin, & Tanenhaus, 2008; Kraljic & Samuel,

2006; Weatherholtz, 2015; for review, see Weatherholtz & Jaeger, in press). However, the prediction remains untested within the domain of syntactic processing.

Present Work

In two studies, we tested how an unfamiliar syntactic structure is processed and generalized. We used the *need*+past participle structure (hereafter, Dialectal *Need*) of Ohio and western Pennsylvania (Doyle, 2008; Murray et al. 1996; Tenny, 1998), in which a sentence like (1) below has the same meaning as a conventional *need to be*+past participle sentence (hereafter, Conventional *Need*), such as (2). The Dialectal *Need* structure is particularly well-suited for assessing readers' implicit expectations because its online comprehension is likely to be strongly modulated by comprehenders' prior familiarity. For comprehenders unfamiliar with the structure, the string *The photocopier needs recycled*… is likely to receive a garden-path interpretation[1] in which *recycled* is assumed to modify an upcoming noun, as in (3). When the word *because* violates this expectation, as in (1), processing should be slowed, which provides a simple measure of the implicit expectations that comprehenders have for Dialectal *Need*.

(1) The photocopier needs recycled because it no longer works. **[Dialectal *Need*]**

(2) The photocopier needs to be recycled because it no longer works. **[Conventional *Need*]**

(3) The photocopier needs recycled paper to comply with our environmental policy. **[Conventional Modifier]**

In Study 1, we follow up previous investigations of how subjects previously unfamiliar with Dialectal *Need* change, over the course of an experiment, how they read this structure (Boland et al., 2015; Kaschak & Glenberg, 2004; Kaschak, 2006). Further, we examine whether

these changes lead readers previously unfamiliar with Dialectal *Need* to process it in a way that increasingly resembles native users. This prediction follows from an implicit learning account of exposure effects of syntactic processing: If existing users of this structure are more fluent in processing it simply because they have more exposure to it, then experimentally providing exposure with the structure to naïve, unfamiliar subjects should gradually lead to increasingly native-like processing. By contrast, if within-experiment adaptation to unfamiliar structures reflects a task-specific strategy for dealing with unfamiliar or "ungrammatical" structures, readers previously unfamiliar with Dialectal *Need* might adjust their reading of it, but not in a way that resembles how a native user would process the structure.

In Study 2, we test how readers *generalize* exposure to an unfamiliar syntactic structure, a question bears on the larger problem of how syntactic knowledge is organized. Specifically, we ask whether exposure to one unfamiliar structure can affect processing of other parts of a syntactic distribution, even those for which the comprehender does not yet have direct evidence of change. The ability to generalize from other structure to another is predicted by an implicit learning account in which comprehenders are learning about an entire *distribution* of structures.

Study 1

Study 1 had two goals. First, we seek to determine whether the rapid adaptation to an unfamiliar structure (Kaschak & Glenberg, 2004) can also be observed in a web-based crowdsourcing paradigm (which we also employ in Study 2). Second, we conduct the first test of whether such within-experiment adaptation actually leads readers to process the Dialectal *Need* structure more like who have already had long-term (potentially life-long) exposure to the dialect that contains this syntactic feature, as would be predicted under an implicit learning account. Alternately, we might find that readers unfamiliar with the Dialectal *Need* structure

adapt their reading behavior in a way that does not bring them in line with existing users of the construction. If so, that would suggest that the changes in reading and syntactic processing seen in experiments like this one instead reflect atypical, task-specific strategies (Kim & Mauner, 2006) or non-strategic changes in processes other than learning (as sometimes argued for comprehension priming more generally; Traxler & Tooley, 2008).

Method

Participants. We recruited participants ($N = 224$) using the Amazon Mechanical Turk crowd-sourcing website. All participants were native speakers of English.

Our hypotheses concerned the effects of participants' prior exposure (or lack thereof) to the Dialectal *Need* structure. To facilitate this comparison, we targeted participant recruitment (see Procedure) to areas in which prior work (Doyle, 2014; Murray et al., 1996) indicated that participants were likely to be familiar with the structure (Ohio and western Pennsylvania, $N = 114$) or unfamiliar with the structure (Colorado, $N = 110$). A small number of participants ($N = 5$) signed up outside of the targeted regions; we retained these participants because our procedure eventually evaluated familiarity or unfamiliarity with Dialectal *Need* at the individual level..

While we expected that participants in some geographic regions would be more likely to be more familiar with the Dialectal *Need* structure, we did not expect participants' current geographic residence would be perfectly predictive of their familiarity with the structure. For example, a participant currently residing in Colorado might have grown up in Ohio and learned the structure there. Thus, a debriefing questionnaire (described below; full questionnaire in Appendix B) queried whether individual participants were familiar with the Dialectal *Need* structure. Based on question 5 of this questionnaire, we sorted participants into *familiar* participants who reported prior knowledge of the structure and *unfamiliar* participants without

such knowledge.

The geographic distribution of participants (after performing an additional exclusion described in the Results section below) is depicted in Figure 1. As can be seen, the targeted participant recruitment was largely successful in identifying participants from the regions of interest. Moreover, as expected, most participants in the Ohio and western Pennsylvania region reported familiarity with the Dialectal *Need* structure; some participants in Colorado also reported familiarity with the structure, but this region still accounted for the majority of unfamiliar participants.

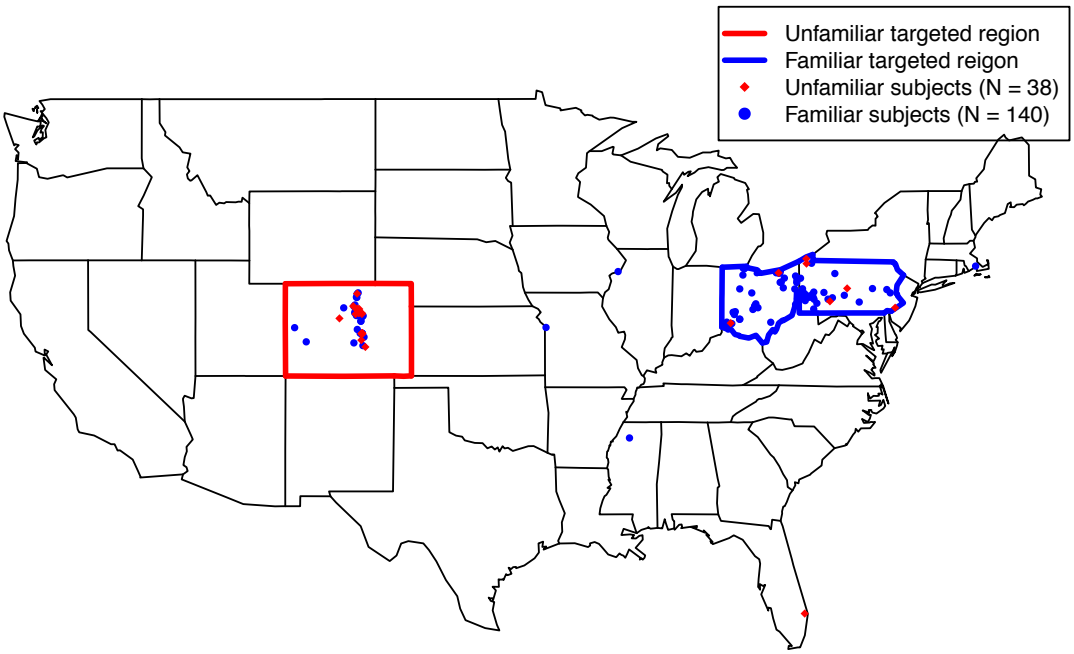

Figure 1. Map of target recruitment for Study 1. Outlines indicate regions targeted for recruitment of participants unfamiliar with *needs*+participle (red) and those already familiar with it (blue). Points indicate individual participants, sorted on the basis of a post-experiment questionnaire into those unfamiliar with *needs*+participle (red) and those familiar with it (blue).

Materials. Participants read 24 paragraphs constructed to resemble e-mails at a fictive small business. Ten of the paragraphs were *critical* paragraphs containing the Dialectal Need structure, ten constituted *filler* paragraphs, and four were practice paragraphs presented at the

beginning of the experiment. An example paragraph is presented in (4) below, and the full list of stimuli is provided in Appendix A.

(4) I'm certainly excited to have a new award to display. The display case needs polished before we add the trophy. Let's do that this afternoon.

Each critical paragraph began with one or two context sentences to provide context for the Dialectal *Need* structure. The next-to-last sentence was the critical sentence that contained the Dialectal *Need* structure. Finally, each paragraph concluded with a final, wrap-up sentence. The filler paragraphs followed the same format except that none of the sentences in a filler paragraph contained the Dialectal *Need* structure. Thus, the Dialectal *Need* structure appeared in 15% of the total set of sentences. The participle used in the Dialectal *Need* structure was different in each of the critical items. The words *need* and *needs* did not appear anywhere in any of the other items (whether as part of the Dialectal *Need* or otherwise).

Reading times increase at the beginning and end of a line of text (Rayner, 1998), which might overwhelm the effects of interest. We thus constructed the critical items so that at least two words intervened between the start of a line and the critical regions (the word *needs*, the participle, and the two disambiguating words afterwards) and between the critical regions and the end of a line.

Two *yes* or *no* comprehension questions were constructed for each paragraph (both critical and filler). We presented more than one comprehension question per paragraph because the paragraphs were longer than the single-sentence items used in many self-paced reading experiments. For instance, example paragraph (4) above was tested using the comprehension

questions presented below (5). In example (5), the correct answer to the first question is *yes* and the correct answer to the second question is also *yes*, but across all paragraphs, the four possible patterns of correct answers to the two comprehension questions (*yes—yes*, *yes—no*, *no—yes*, and *no—no*) were equally frequent.

(5) Q1. Will the display case be polished this afternoon?

Q2. Is there a new award?

Because the effects of interest involved learning over the course of the experiment, it was important to deconfound the effects of particular serial positions from the items occupying those serial positions. Thus, we constructed four different possible stimulus orders: two different semi-random orderings of critical and filler items (with the constraint that there were never more than two critical or two filler items in a row), and the reverse of those two orderings. Across all stimulus orders, whether a given serial position was occupied by a critical paragraph or a filler paragraph was constant.

The stimulus orders were crossed with whether the participant was recruited from a region likely to be familiar or unfamiliar with the structure (although, as noted above, this did not guarantee an individual participant's familiarity or unfamiliarity), resulting in a 2 x 4 between-participants design: recruitment location (Colorado versus Ohio and western Pennsylvania) x stimulus ordering[2].

Finally, we constructed a structured debriefing questionnaire to probe (a) participants' short-term awareness of the structures presented during the experiment and (b) their prior, long-term experience with the Dialectal *Need* structure. One potential concern with asking

participants to self-report their experience with dialectal variants is that participants might be motivated by demand characteristics to report seeing the structure (either in the experiment or over the long term) even if they were actually unfamiliar with it. To catch participants who simply affirmed all dialectal variants, we also included a "catch" question (question 4) that queried whether participants had "seen" over the course of the experiment a different dialectal variant (modal stacking such as *I might could pass this test*; Di Paolo, 1989) that actually never appeared in the experiment. The complete debriefing questionnaire for Study 1 appears in Appendix B.

Procedure. The initial phase of the experiment determined participant's geographic region, which was used to identify participants who lived in regions where they were likely to be familiar or likely to be unfamiliar with the structure of interest. Participants were instructed to sign up for the experiment on Amazon Mechanical Turk only if they resided in one of the targeted regions, which were described both verbally and by a highlighted section of a map of the United States. As an initial step in the experiment, participants then had to indicate in which of the targeted regions they resided via a drop-down box on a Web form. Finally, with consent from participants, we used the geolocation information in participants' Web browsers to identify their exact coordinates of latitude and longitude; we then presented what we had identified as the participants' current city so that participants could confirm its accuracy. If participants refused the geolocation, if the geolocation information could not be obtained from the Web browser, or if participants reported that the automatically identified location was incorrect, participants were asked to self-report their city of residence instead.

The self-paced reading task then began with a cover story instructing participants that they would be reading e-mail messages. The instructions then described the self-paced reading

task to the participants. After reading the instruction, participants first completed two single-sentence practice items, one of which further reminded participants of the location of the fictive company, followed by four practice paragraphs designed to acclimate participants to the self-paced reading task. The twenty paragraphs in the experimental list (10 critical and 10 filler) then followed.

Each paragraph was presented using the self-paced moving window paradigm (Just, Carpenter, & Woolley, 1982) so that reading times could be collected for individual words (via a Flash applet written by Harry Tily, Nuance Technology, and adjusted for our purpose by Andrew Watts, Human Language Processing Lab, University of Rochester). The paragraph was initially displayed on the screen with only the first word visible; the other words were masked by lines. Participants pressed the space bar to advance; after each press, the next word was displayed and the previous word replaced by a line. After reading each paragraph, participants answered two comprehension questions, presented one at a time, by pressing one of two keys to indicate *yes* or *no*. The purpose of the comprehension questions was to encourage participants to read each item for comprehension. If a comprehension question was answered incorrectly, a "Wrong answer—please try your best to answer correctly" message appeared on the screen for 2000 ms so that participants could not complete the experiment more quickly by answering the questions spuriously.

Once participants had finished the self-paced reading task, participants completed the debriefing questionnaire via a web form. The questions were presented one at a time in order from least revealing about the structure of interest to most revealing and appeared after participants had completed the entire self-paced reading task.

Results

Data processing and exclusion. As noted above, to enhance the validity of the self-report debriefing questionnaire, we included a "catch" question in the debriefing questionnaire to identify participants who simply affirmed all dialectical variants. Participants who falsely affirmed seeing the *I might could* (modal-stacking) structure in the experiment, which actually never appeared, were excluded, leaving 140 participants in the *familiar* group and 38 in the unfamiliar group. (The differing numbers of participants across groups reflects the fact that participants were categorized on their self-reported prior, out-of-laboratory familiarity with Dialectal *Needs*, which could not be experimentally manipulated.)

We analyzed reading times from the remaining 178 participants for the 20 paragraphs of interest (excluding the practice paragraphs). Following standard procedure, we excluded implausibly long (greater than 2000 ms) reading times, affecting 0.4% of the data, short reading times (less than 100 ms), affecting another 0.6% of the data, and reading times from paragraphs for which participants did not correctly answer at least one of the comprehension questions[3], affecting another 0.7% of the data.

Following the self-paced reading literature, reading times were then corrected for word length and for participants' baseline reading speed (F. Ferreira & Clifton, 1986). Note that, in this experiment, we expected systematic differences between subjects in their overall reading time because some subjects were familiar with the critical structure and others were not. Thus, reading times on the critical items would not be informative about participants' baseline reading speed. Instead, we performed the correction by first performing a mixed-effects regression exclusively on the reading times on the filler items. This regression included as predictors the word length, a random intercept capturing each participant's baseline reading speed, and a random slope capturing how word length affected reading time for each participant. The

materials in this experiment, unlike those in most self-paced reading experiments, spanned several lines. Since line breaks are known to affect reading times at the beginning and end of lines (e.g., Kuperman, Dambacher, Nuthmann, & Kliegl, 2010), we also included as additional predictors (in both the fixed effects as random slopes) whether a word began a line and whether a word ended a line. (None of the *critical* words ever appeared in those positions, but screen position could influence reading time on the other words used to perform the residual reading time regression.) Indeed, the inclusion of both of these screen position parameters was justified by significance tests of the fixed effects; $t = 15.0$, $p < .001$, for line beginnings, and $t = 17.9$, $p < .001$ for line ends.

As a final predictor in this residual reading time regression, we included the log serial position of the item within the experiment (following Fine and Jaeger, 2016). Reading times have been observed to generally decrease over the course of self-paced experiments as participants become accustomed to the self-paced reading task (e.g., Fine et al, 2010, 2013); this general speed-up is separate from, and can be empirically dissociated from, adaptation to any particular syntactic structures (Fine et al., 2010). Indeed, we observed such a relation in the present data: Figure 2 depicts the mean per-word reading times on the filler paragraphs as a function of the serial position of the paragraph within the experiment; the decrease in reading time over the course of the experiment appears to be well-fit by a logarithmic function. To separate this general speed-up from any effects of within-experiment adaptation to the Dialectal *Need* structure in particular, we included log serial position as an additional predictor in the residual reading time regression to regress out the general adaptation to the self-paced reading task. (The inclusion of this predictor was justified; $t = -8.77$, $p < .001$). Thus, any remaining effects of serial position on reading of Dialectal *Need* represent changes in reading time over and

above this general adaptation effect. The results of all experiments reported below remain unchanged regardless of this decision[4].

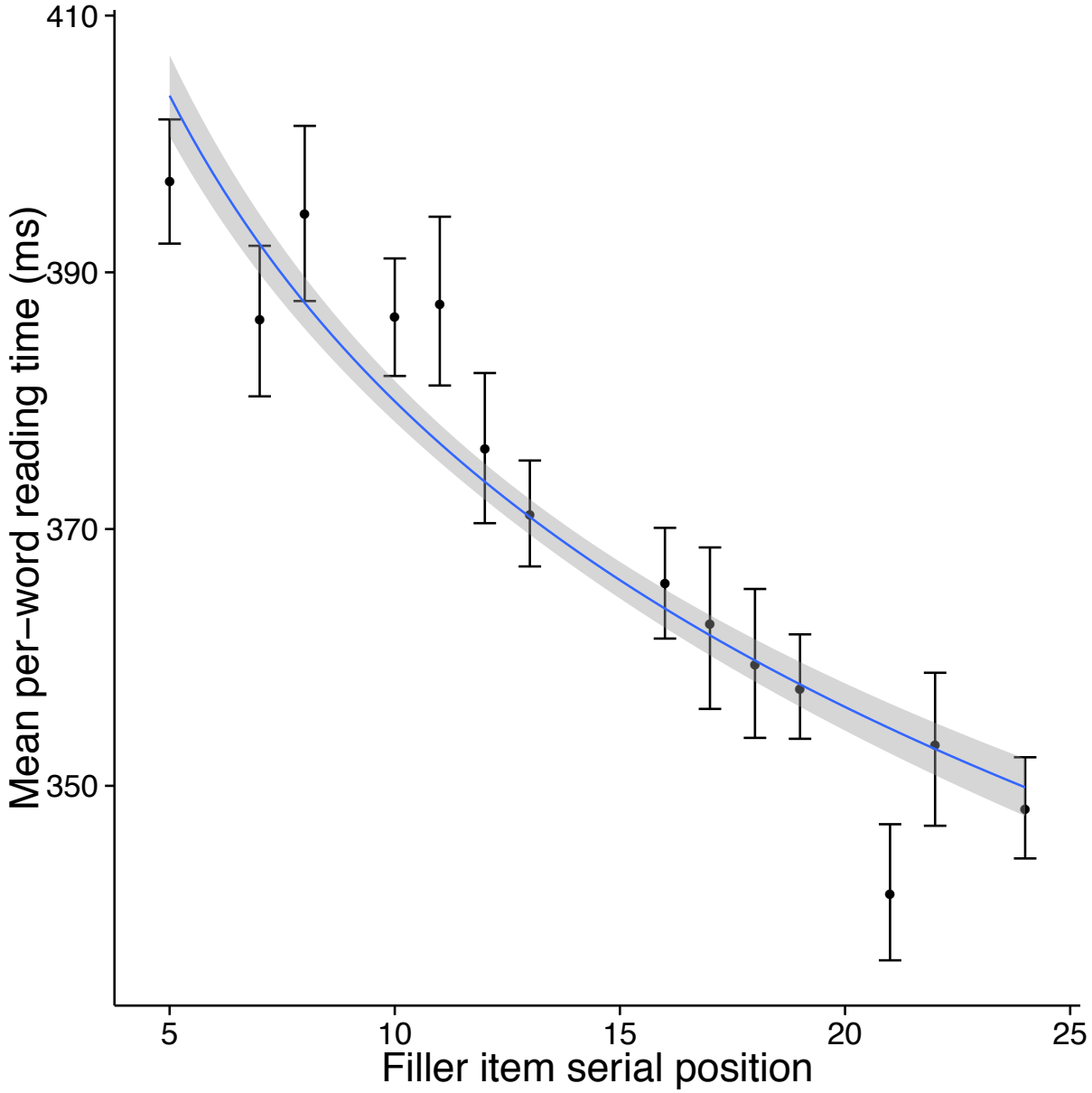

Figure 2. Mean per-word reading time in filler items in Study 1 as a function of serial position within the experiment. (Note that not all serial positions are represented because of our experimental design: Our list design held the serial position of items and fillers constant, and as a consequence some serial positions never contained a filler item.)

Having performed the residual reading time regression, we eliminated extreme reading times with standardized residuals greater than 3.0, affecting 1.9% of the data, and then re-ran the regression. Finally, this regression was then used to predict reading times on the critical items; residuals from these predictions were then used as a measure of reading time that corrected for length, screen position, and task practice.

For each critical item, we identified four sentence regions relevant to the hypotheses of interest. Following Kaschak and Glenberg (2004), these regions were: (a) the word *needs*, (b) the participle, and (c) the disambiguating word after the participle, and (d) a spillover region made up of the word after the disambiguation. For instance, in critical sentence (6) below (excepted from the paragraph-long example stimulus 4 above), region (a) is the word *needs*, the participle region (b) is the word *polished*, the disambiguation (c) is the word *before*, and the spillover region (d) is the word *we*.

(6) The display case needs polished before we add the trophy.

Because our hypotheses concerned learning about the Dialectal *Need* structure over the course of the experiment, we also coded the trial number within the sequence of critical trials. We use the trial number among the critical items only because the filler items never contained the Dialectal *Need* structure and thus contributed no information about it.

Analysis approach. Reading times were submitted to a mixed effects regression (Baayen, Davidson, & Bates, 2008). The regression included three fixed effects (critical trial number, sentence region, and prior familiarity) and their full factorial interaction. Sentence

region was treatment-coded, contrasting each of the other regions to the baseline at the word *needs*, which precedes the ambiguous participle. Note that these contrasts (like other regression contrasts) test differences between particular regions rather than an overall omnibus effect of Region; this property is desirable given that our hypotheses predicted difference in a specific region (i.e., the disambiguation region). The other predictor variables were mean-centered; doing so produces estimates of main effects averaging over the other variables (e.g., the effect of familiarity at the *average* critical trial number), analogous to those obtained in an AN(C)OVA.

In this and in all other regressions, participants and items (paragraphs) were included as random effects. We included a near-maximal random effects structure with the omission only of correlations between the random effects parameters and of random slopes not justified by the design (i.e., there were no random by-subject slopes for between-subject effects). All regressions were performed in the R Project for Statistical Computing using the *lmer()* function of the *lme4* package (Bates, Maechler, Bolker, & Walker, 2015). In all analyses reported in this paper, fixed effect correlations were within acceptable ranges for balanced designs like the present one (absolute *r*s < .7).

Reading time results. Reading times are displayed in Figure 3 as a function of critical item order, sentence region, and prior familiarity with the Dialectal *Need* structure. (To facilitate comparison to prior studies and interpretation, we visualize the data with the same three bins of critical trial numbers used by Kaschak and Glenberg, 2004: critical trials 1-3, critical trials 4-7, and critical trials 8-10. However, the regression incorporated the full range of individual critical trial numbers in order to incorporate all of the information available about changes over the course of the experiment, following Fine et al., 2013; Fine & Jaeger, 2016)

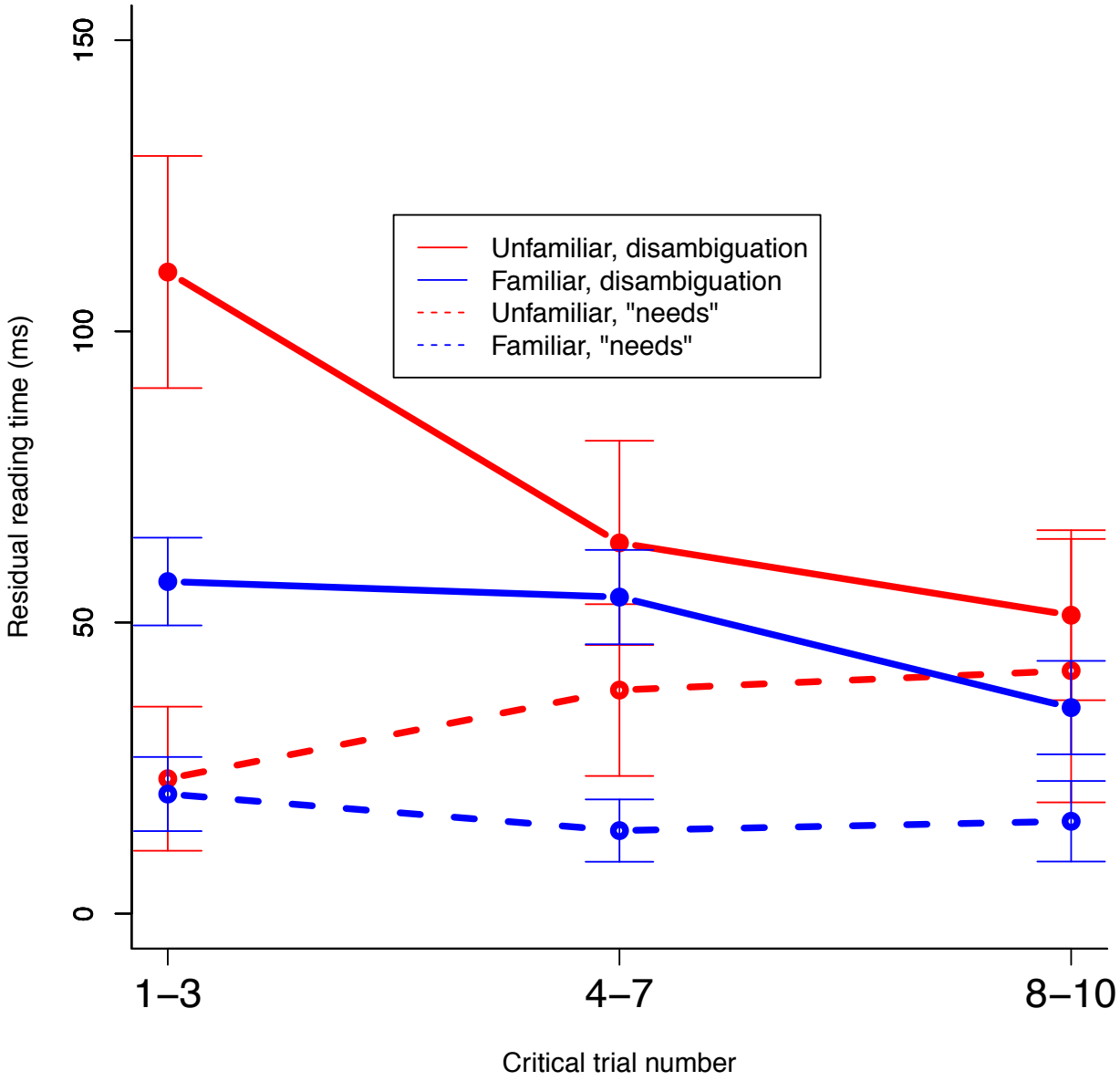

Figure 3. Residual reading times (corrected for word length, screen position, and task practice) in Study 1 as a function of participant group, sentence region, and critical trial number. Error bars represent ±1 standard error.

Table 1 summarizes the results. For regressions with thousands of observations, such as

the present one, absolute *t* values greater than 2 indicate significance at the α = .05 level (Baayen, 2008, p. 270).

Table 1
Fixed Effect Estimates for Linear Mixed Effects Regression of Residual Reading Times in Study 1.

Fixed effect	$\hat{\beta}$	*SE*	*t*	*p*
Intercept (baseline residual reading time on *needs*)	24.81	7.90	3.14	< .01
Region: Participle (vs. *needs*)	1.72	6.05	0.29	.77
Region: Disambiguation (vs. *needs*)	40.56	6.78	5.98	< .001
Region: Spillover (vs. *needs*)	15.65	7.29	2.15	< .05
Critical trial number	1.95	1.54	1.27	.20
Prior familiarity	-15.97	10.96	-1.46	.14
Participle region (vs, *needs*) x critical trial number	-3.41	2.21	-1.54	.12
Disambiguation region (vs. *needs*) x critical trial number	-7.41	2.04	-3.63	< .001
Spillover region (vs. *needs*) x critical trial number	-7.02	1.96	-3.58	< .001
Participle region (vs. *needs*) x familiarity	16.10	12.85	1.25	.21
Disambiguation region (vs. *needs*) x familiarity	-12.08	13.60	-0.89	.37
Spillover region (vs. *needs*) x familiarity	-20.33	12.59	-1.62	.11
Critical trial number x familiarity	-4.18	3.08	-1.36	.18
Participle region (vs. *needs*) x critical trial number x familiarity	3.41	4.42	0.77	.44
Disambiguation region (vs. *needs*) x critical trial number x familiarity	11.78	4.08	2.89	< .01
Spillover region (vs. *needs*) x critical trial number x familiarity	11.09	3.92	2.83	< .01

Note. SE = standard error.

There were some baseline differences across sentence regions in the speed at which they were read (perhaps stemming from syntactic class or linear sentence position), reflected in main effects of sentence region and in interactions between trial number and sentence region.

Of primary interest, however, were the effects of participants' prior familiarity with Dialectal *Need*. A reliable three-way Disambiguation Region x Familiarity x Critical Trial

Number interaction revealed that the difference between the familiar and unfamiliar participants at the disambiguation of Dialectal *Need* relative to the baseline at the word *needs* (i.e., the two-way Disambiguation Region x Familiarity interaction) diminished over the course of the experiment, $t = 2.89$, $p < .01$. Similarly, the difference between the familiar and unfamiliar participants in the *spillover* region of Dialectal *Need*, relative to the words *needs*, also declined over the course of the experiment (Spillover Region x Familiarity x Critical Trial Number interaction: $t = 2.83$, $p < .01$). Note that these effects cannot be attributed to a general speed-up as a function of experience with the self-paced reading task because that effect was already regressed out of the reading times as part of the process of calculating residual reading times. Nor can this effect be attributed to differences between the groups in their overall reading speed across regions; there was also no significant main effect of Familiarity on overall reading speed, $t = -1.46$, $p = .14$.

In addition, the speed up in reading times for the unfamiliar group was specific to the disambiguation region: There was no such interaction at the participle ($t = 0.77$, $p = .44$) or at the unambiguous word *needs* ($t = -1.36$, $p = .18$). In fact, the effect at *needs* was numerically in the opposite direction; participants unfamiliar with Dialectal *Need* came to read this region numerically slower over the course of the experiment. We are reluctant to over-interpret that pattern given that it did not reach conventional levels of significance ($p = .18$) and was not replicated in Study 2. Further, the (numerical) increase in reading time at *needs* is of smaller magnitude than the decrease in reading time at the disambiguation, so participants previously unfamiliar with Dialectal *Need* were coming to read the structure faster as a whole.

To summarize, differences in within-experiment adaptation occurred specifically at the region that disambiguated the Dialectal *Need* structure and specifically for the participants

previously unfamiliar with it. Over the course of the experiment, the unfamiliar participants came to read the Dialectal *Need* structure at speeds more like those of the familiar group.

Discussion

Using the Web, we replicated prior findings (Kaschak & Glenberg, 2004) of rapid adaptation to unfamiliar syntactic structures: Participants previously unfamiliar with Dialectal *Need* were initially slower to read the sentence region that disambiguated this structure, but this effect diminished after just a few exposures to the structure. This result demonstrates that adaptation to unfamiliar syntactic structures can be observed even in a Web-based task. It also validates the targeted geographic recruitment procedure and questionnaire as capturing dialect differences relevant to online language processing.

Going beyond Kaschak and Glenberg (2004), a comparison to participants already familiar with Dialectal *Need* indicated the unfamiliar participants quickly came to read the structure increasingly like the familiar group. Indeed, by the end of the experiment, the differences in reading time between the groups had essentially vanished; the unfamiliar participants were reading the structure almost exactly like the familiar participants. Although it may seem surprising how rapidly subjects approached native-like processing here, this learning task was comparatively easy because unfamiliar subjects had to learn only a single novel structure to understand the dialect. Further, similarly fast learning has been observed after exposure to familiar structures (e.g., Farmer et al., 2014; Fine et al., 2013).

Because the decrease in reading times in the disambiguation brought the subjects unfamiliar with Dialectal *Need* increasingly in line with how a genuine existing user of the structure would read it, it seems unlikely that these changes in reading behavior reflect an atypical task-specific strategy (Kim & Mauner, 2006), such as relaxing one's usual constraints

on grammaticality and becoming more accepting of unusual or unusual input. Indeed, in the General Discussion below, we report an additional post-hoc analysis that suggests participants' explicit judgments of whether the experiment contained unusual grammar had no relation to whether and how they adapted to Dialectal *Need* over the course of the experiments. Further, the effects of long-term, out-of-laboratory experience with Dialectal *Need*—accrued over days, months, or years rather than from one trial to the next—suggest that exposure effects on syntactic comprehension are not purely a matter of short-term, transient priming or activation. While such interactions of long-term and recent experience can be accommodated within an activation-based account, they require reference to implicit learning (Reitter et al., 2011) and are therefore inconsistent with the idea that activation is purely transient or short-lived.

Rather, the results of Study 1 seem most consistent with an account in which exposure effects on syntactic processing reflect a process of continuous, implicit learning about the distribution of syntactic structures to be expected in a particular situation. Implicit, continuous learning would cause readers encountering the Dialectal *Need* structure for the first time to gradually process more and more like the readers who have already been exposed to (and, thus, learning about) the structure over the long term, and to do so regardless of whether or not participants perceived the grammar as unusual.

Study 2A

The pattern of results in Study 1 can be viewed as adaptation insofar as the changes in participants' reading behavior led them towards convergence with the input. But when readers adapt to Dialectal *Need* in this way, what exactly is changing? One possibility is that readers adapt to only the Dialectal *Need* structure itself. However, a language comprehension system sensitive to contingencies between syntactic frequencies in the environment should respond to

the fact that unfamiliar syntactic structures tend to co-occur (e.g., in a newly-encountered dialect) and, on the basis of encountering the unfamiliar Dialectal *Need* structure, also expect other unfamiliar structures.

In Study 2A, we investigated whether exposure to Dialectal *Need* facilitates a different unfamiliar structure; namely, a dropped *be* before a participle in the future tense, as in (7). This *be*-drop structure is not an attested part of any dialect, but has been introduced by Boland, de los Santos, Carranza, and Kaschak (2015) to examine adaptation to unfamiliar structures (and its use here thus facilitates comparison across studies).

(7) The copier will recycled because it no longer works. **[*be*-drop]**

The *be*-drop structure can be interpreted as structurally similar to Dialectal *Need*. Specifically, Dialectal *Need* might be interpreting as omitting the two function words *to be*[5], and the *be*-drop structure might also be interpreted as omitting *be*. Thus, readers might generalize expectations from one structure to the other. Further, this (hypothetical) property shared by both structures—omission of function words—also exists in other structures of *Standard* American English. For example, complementizer *that* in both complement clauses and certain types of relative clauses is frequently omitted in Standard American English (e.g., Bolinger, 1972; Elsness, 1984; Jaeger, 2006, 2010; Roland, Elman, & V. Ferreira, 2006; Wasow, Jaeger, & Orr, 2011). Similarly, the function word *to*, which is one of the words that might be considered omitted in the Dialectal *Need* structure, can be omitted in several other linguistic contexts of Standard American English (e.g., after the verb *help*; Rohdenburg, 1996). Standard American English also provides examples in which multiple function words are omitted; for example, in

subject-extracted passive relative clauses such as *A president (who is) admired by the people has no need to …* (for references, see Jaeger, 2011). These types of omissions are not limited to conversational speech, but also frequently occur in writing (e.g., Elsness, 1984; Jaeger, 2011; Rohdenburg, 1996; Roland et al., 2006). Therefore, it is plausible that comprehenders could attend to the apparent omission of function words and use the presence of one structure that can be interpreted as an omission of a function word (i.e., Dialectal *Need*) to infer that other such omissions (as in *be*-drop) will also occur.

Two further properties of English are expected to bias readers who are exposed to Dialectal *Need* for the first time towards expecting further novel structures omissions. First, there are other varieties of American English that contain structures that differ from the corresponding structure in Standard American English in that they omit function words. For examples, some varieties of American English omit complementizer *that* in environments where Standard American English does not allow omission (Tottie & Rey, 1997); the same alternation is observed in several varieties of British Isle English (e.g., Tagliamonte et al., 2005). Perhaps more importantly, copula drop of inflected forms of *be* is observed in certain linguistic contexts in African American Vernacular (Bender, 2000; Labov, 1969; Wolfram, 1974), which most American participants will have at least passive exposure to (at least through the media). That is, not only is function word omission frequently observed in Standard American English, many participants will likely also have implicit knowledge that some varieties of English allow omission in linguistics contexts where it is ungrammatical in Standard American English.

Finally, for alternations like *that*- or *to*-omission discussed above, the rate of omission tends to be correlated across structures. For example, omission rates across different alternations decrease with increasing formality (Finegan & Biber, 2001). All of these factors together are

expected to make *be*-drop relatively likely after exposure to the previously unfamiliar Dialectal *Need*.

Thus, if—as we hypothesize—comprehenders use their exposure to part of a syntactic distribution to infer what other structures might be present or frequent in that distribution, adaptation to Dialectal *Need* should generalize, in whole or in part, to the similar *be*-drop structure. In Study 2, we tested this prediction by comparing processing of this *be*-drop structure by participants exposed to that specific structure versus by participants exposed to a different, but similar unfamiliar structure (Dialectal *Need*).

Specially, we used an exposure-test paradigm similar to that used by Kaschak and Glenberg (2004). In Study 2, *all* participants were previously unfamiliar with the target structures. Rather, the variable of interest was one that was experimentally manipulated: the type of experience that was provided to participants within the experiment. In an initial *exposure* phase, participants saw either the Dialectal *Need* structure (*needs-exposure* condition) or the *be*-drop structure in (4) above (*be-drop-exposure* condition), intermixed with fillers in both cases. After the exposure phase, all participants transitioned—invisibly to the participant—to a subsequent *test* phase in which all participants saw the *be*-drop structure.

We expected that participants in the *be-drop-exposure* condition would gradually adapt to their exposure to the unfamiliar *be-drop* structure, paralleling adaptation to Dialectal *Need*. The critical condition was *need-exposure*, in which participants initially encounter the writer using an unfamiliar structure (Dialectal *Need*) that can be interpreted, like *be*-drop, as involving omitted function morphemes. If participants generalize this experience to expect other, similar dialectal syntactic differences, they should read the disambiguation of *be-drop* as quickly as *be-drop-exposure* subjects. This result would indicate that knowledge and expectations about one

syntactic structure are not independent of knowledge about other structures likely to be used by the writer. Figure 4 summarizes the design and predictions for this study and for Study 2B.

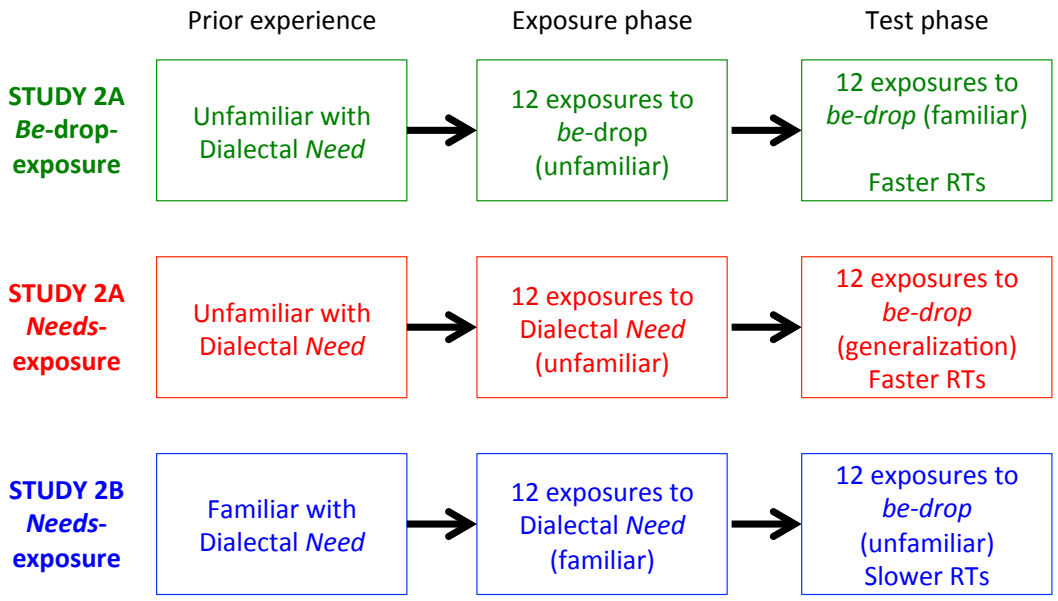

Figure 4. Design and predictions for Study 2A and Study 2B.

In Study 2, we also further developed our targeted geographic recruitment procedure. Although the procedure in Study 1 was generally successful at locating participants without prior exposure to the Dialectal *Need* structure, we did encounter a number of participants who reported possible experience to this structure even in the region we targeted for unfamiliarity with the structure (Colorado), resulting in an imbalance in the number of previously-familiar participants

versus previously-unfamiliar participants. To refine our procedure, and to facilitate the recruitment of the large number of people unfamiliar with Dialectal *Need* that we needed for Study 1, we revised our recruitment procedure in two ways. First, because Colorado may not have been ideal for recruiting participants without exposure to Dialectal *Need* (e.g., because of the high rates of in-migration to this area, many participants may know people from areas with this dialect or have grown up in those areas themselves), we instead targeted other areas where the structure is typically not used. Specifically, we targeted California and the Northeastern United States, which had especially low rates of use of this structure in an analysis of social media postings (Doyle, 2014). Second, we also revised the post-experiment questionnaire (as detailed below) to eliminate only those participants who were *certain* that they were familiar with Dialectal *Need* rather than those who were simply defaulting to a "yes" response.

Method

Participants. We recruited 160 participants from California and the northeastern United States using the targeted recruitment procedure. When participants signed up for the experiment, they were randomly assigned to either the *needs-exposure* condition or the *be-drop-exposure* condition. Figure 5 displays the location of the participants recruited for Study 2A.

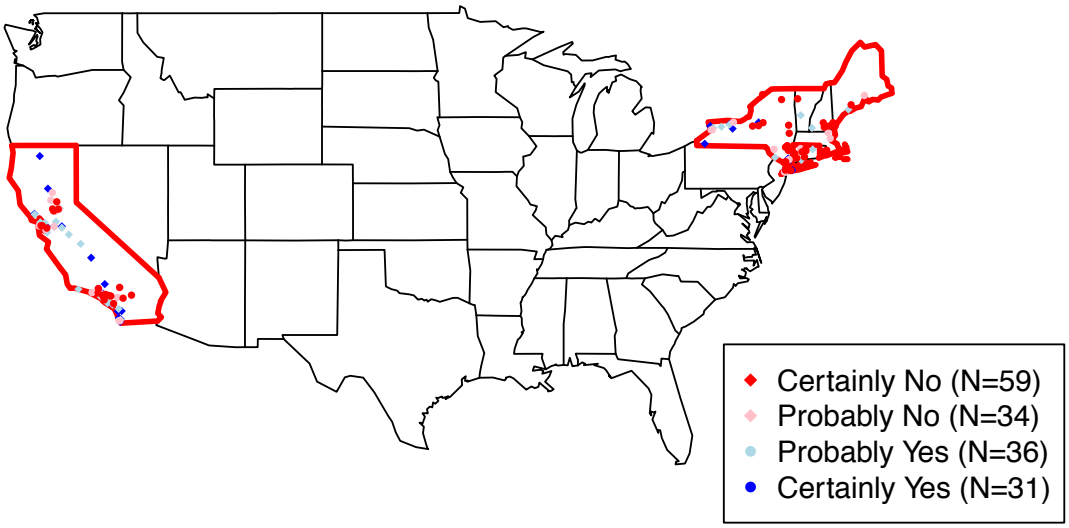

Figure 5. Map of target recruitment for Study 2A. Outlines indicate regions targeted for recruitment of participants unfamiliar with the Dialectal *Need* structure. Points indicate individual participants as a function of their previous experience with Dialectal *Need*, as self-reported on the post-experiment questionnaire.

We eliminated 22 participants because they reported they were not native speakers of

English, 27 participants because they reported they were certain they had heard the Dialectal *Need* structure before, and 10 participants because they incorrectly answered the "catch" question in the post-experiment questionnaire (i.e., they reported being certain that they saw *might could* sentences in the experiment), suggesting they might not have been attending to the questionnaire or answering it sincerely. These exclusions left 49 participants in the *needs*-exposure condition and 52 participants in the *be-drop*-exposure condition (101 total), all previously unfamiliar with Dialectal *Need*.

Materials. We constructed twenty new critical paragraphs for the *be*-drop condition and twenty for the *needs*-training condition. Paragraphs had the same format as in Experiment 1, again with 15% of all sentences containing the critical *be*-drop or Dialectal *Need* structure. To control for item effects, each *be*-drop paragraph was constructed to have the same context sentences and same participle as a corresponding *needs* paragraph. The words *needs*, *need*, and *will* never appeared in any of the items (in any structure) outside of these critical sentences. We also constructed additional filler paragraphs to maintain the same filler-to-item ratio as in Study 1. Appendix C presents the Study 2 materials.

We then used these items to construct twenty different lists. Ten of the lists were used in the *be*-drop-exposure condition. These ten lists constituted ten different possible stimulus orderings for the *be*-drop exposure condition. We constructed a single semi-random ordering of critical item positions and filler item positions with the constraint that there were never more than two critical or two filler items in a row. Then, using a Latin Square design we constructed five different assignments of individual stimulus items to these list positions such that all items were equally likely to appear in the exposure phase and training phase. The other five *be*-drop exposure lists were constructed by reversing these orders. The other ten lists were used in the

needs-exposure condition; each of these lists was exactly matched to a *be*-drop-exposure list except that the Dialectal *Need* version of each critical item was substituted for the *be*-drop version of that item. This procedure ensured that item identity was deconfounded from experimental phase (exposure phase or test phase), from serial position within each phase, from experimental condition (*be*-drop-exposure or *needs*-exposure), and from the condition x critical trial number interactions; thus, any differences across conditions or changes over the course of the experiment could not be attributed to any specific ordering of items.

We also revised the questionnaire to acquire a more sensitive measure of participants' dialectal experience. When faced with the two-alternative yes-or-no question used in Study 1, participants who were merely unsure whether they had heard Dialectal Need may have defaulted to a "yes" response under the assumption that they must have heard it at least once somewhere. In Study 2, we used a more detailed questionnaire in which all of the yes-no questions from the questionnaire in Study 1 were revised to a four-point scale of *certain no*, *probably no*, *probably yes*, and *certain yes*. In addition, we added questions querying where participants *thought* the Dialectal *Need* and modal-stacking structures might be used; these questions were included to assess participants' naïve beliefs in anticipation of a future cover-story manipulation about the putative location of the author of the paragraphs. Appendix D presents this revised questionnaire.

Procedure. The recruitment and experimental procedures were largely the same as in Study 1.

The main difference was that Study 2A used an exposure-test paradigm. The first 24 paragraphs (12 critical and 12 filler) constituted the exposure phase. For half of the participants—those randomly assigned to the *needs-exposure* condition—the 12 critical

paragraphs in this phase contained the Dialectal *Need* structure. For the other half of the participants—those in the *be-drop-exposure* condition—these 12 critical paragraphs contained *be*-drop sentences.

After the last item in the exposure phase was read, participants transitioned to the test phase. There was no indication to the participant that they were in a new phase of the experiment or that they should expect different categories of items. In the *test phase*, both groups of participants saw 8 critical paragraphs containing *be*-drop and 8 filler paragraphs.

Results

Using the same procedure as in Study 1, reading times were adjusted for length, screen position, and overall (log) serial position within the experiment. We performed the residualization simultaneously for Study 2A and for Study 2B below because they used the same items; doing so gave us the largest possible sample with which to adjust for the length and position effects in these items. We first excluded RTs over 2000 ms (affecting 0.2% of the data) or under 100 ms (affecting 1.6% of the data) or that were from paragraphs where at least one comprehension question was not answered correctly[6] (affecting 0.7% of the remaining data). As in Study 1, we then residualized the reading times on the fillers, excluded reading times with standardized residuals greater than 3 (affecting less than 0.1% of the remaining data), re-residualized the reading times, and used this final regression to calculate predicted and residual reading times for the critical items.

We first examine the exposure phase separately for each exposure group to establish whether participants were adapting to the structures provided in them as we expected. Then, we examine the consequences of this exposure manipulation for processing the sentences in the test phase.

***Needs-exposure* phase.** We analyzed the same regions of the Dialectal *Need* structure as in Study 1. Reading times at the critical disambiguating region and at the unambiguous word *needs* are plotted in Figure 6. For reference, we also plot reading times on the same materials from participants who were already familiar with Dialectal *Need*; these participants are described and analyzed in greater detail in Study 2B below.

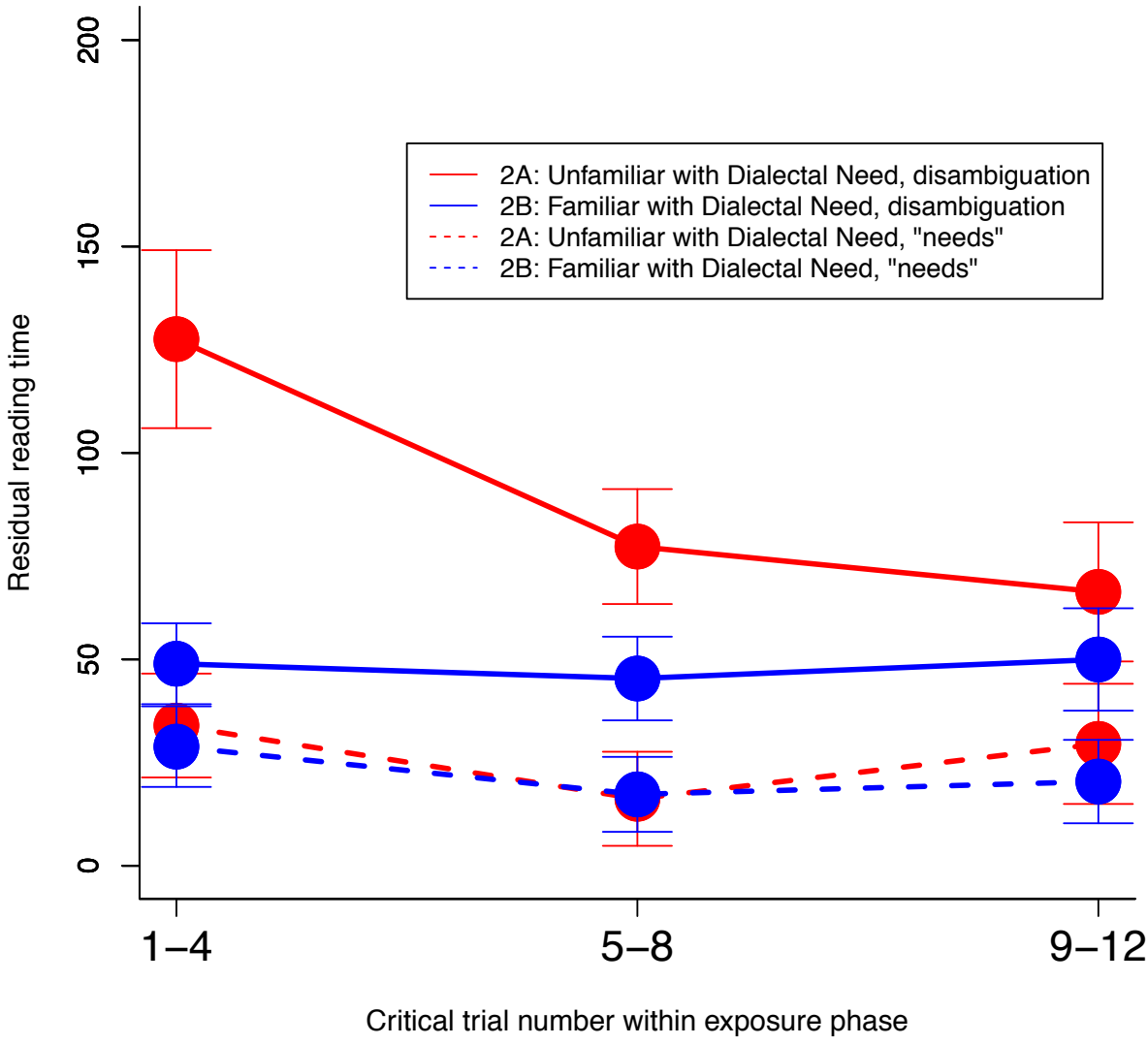

Figure 6. Residual reading times (corrected for word length, screen position, and task practice)

in *needs*-exposure phase of Study 2A and Study 2B as a function of participant group, sentence region, and critical trial number. Error bars represent ±1 standard error.

We performed a mixed-effects regression examining residual reading times in the *needs*-exposure phase as a function of sentence region (coded the same way as in Study 1), critical trial number, and their interaction. The results are summarized in Table 2.

Table 2
Fixed Effect Estimates for Linear Mixed Effects Regression of Residual Reading Times in Needs-*Exposure Phase of Study 2A.*

Fixed effect	$\hat{\beta}$	*SE*	*t*	*p*
Intercept (baseline residual reading time at unambigious *needs*)	25.80	9.11	2.83	< .01
Region: Participle (vs. *needs*)	7.05	10.90	0.65	.52
Region: Disambiguation (vs. *needs*)	63.00	14.83	4.27	< .001
Region: Spillover (vs. *needs*)	17.10	9.18	1.87	.06
Critical trial number	-1.04	2.26	-0.46	.65
Participle region (vs. mean) x critical trial number	0.13	2.67	0.05	.96
Disambiguation region (vs. mean) x critical trial number	-6.63	3.09	-2.15	< .05
Spillover region (vs. mean) x critical trial number	-2.24	2.67	-0.84	.40

Note. SE = standard error.

This regression provided evidence for adaptation to Dialectal *Need* analogous to that obtained in Study 1. For these participants, who were previously unfamiliar with Dialectal *Need*, reading times were especially elevated in the disambiguating region of Dialectal *Need* ($t = 4.27$, $p < .001$), relative to the baseline at the word *needs*. (Marginally elevated reading times also occurred in the spillover region, $t = 1.87$, $p = .06$). This difficulty at the disambiguation of Dialectal *Need* diminished over the course of the exposure phase ($t = -2.15$, $p < .05$). By comparison, there was no such decrease at the word *needs* ($t = -0.46$, p = .65) nor at the participle ($t = 0.05$, $p = .68$), suggesting that participants were adapting to the disambiguation of Dialectal

Need in particular. These patterns provide further evidence that participants unfamiliar with Dialectal *Need* can rapidly adapt to reading it.

***Be-drop-exposure* phase.** We next turn to whether participants in the *be*-drop-exposure condition similarly adapted to the *be*-drop sentences that they encountered in the exposure phase. The *be*-drop sentences had a different structure than the Dialectal *Need* sentences, and so the regions of interest were different. We examined three sentence regions (paralleling the analysis of Dialectal Need) in the critical *be*-drop sentences: (a) the unambiguous word *will*, (b) the disambiguating word after *will*, and (c) a spillover region consisting of the following word. For instance, in example (7), reproduced below as (8), the first region is *will,* the disambiguation region is *recycled*, and the spillover region is the word *because*. Figure 7 displays reading times in the unambiguous *will* region and the spillover region as a function of critical trial number. (As will be shown below, similar effects obtained at both the disambiguation itself and the spillover region; to simply the figure, we focus on the spillover region, where the effects were most robust.)

(8) The copier will recycled because it no longer works.

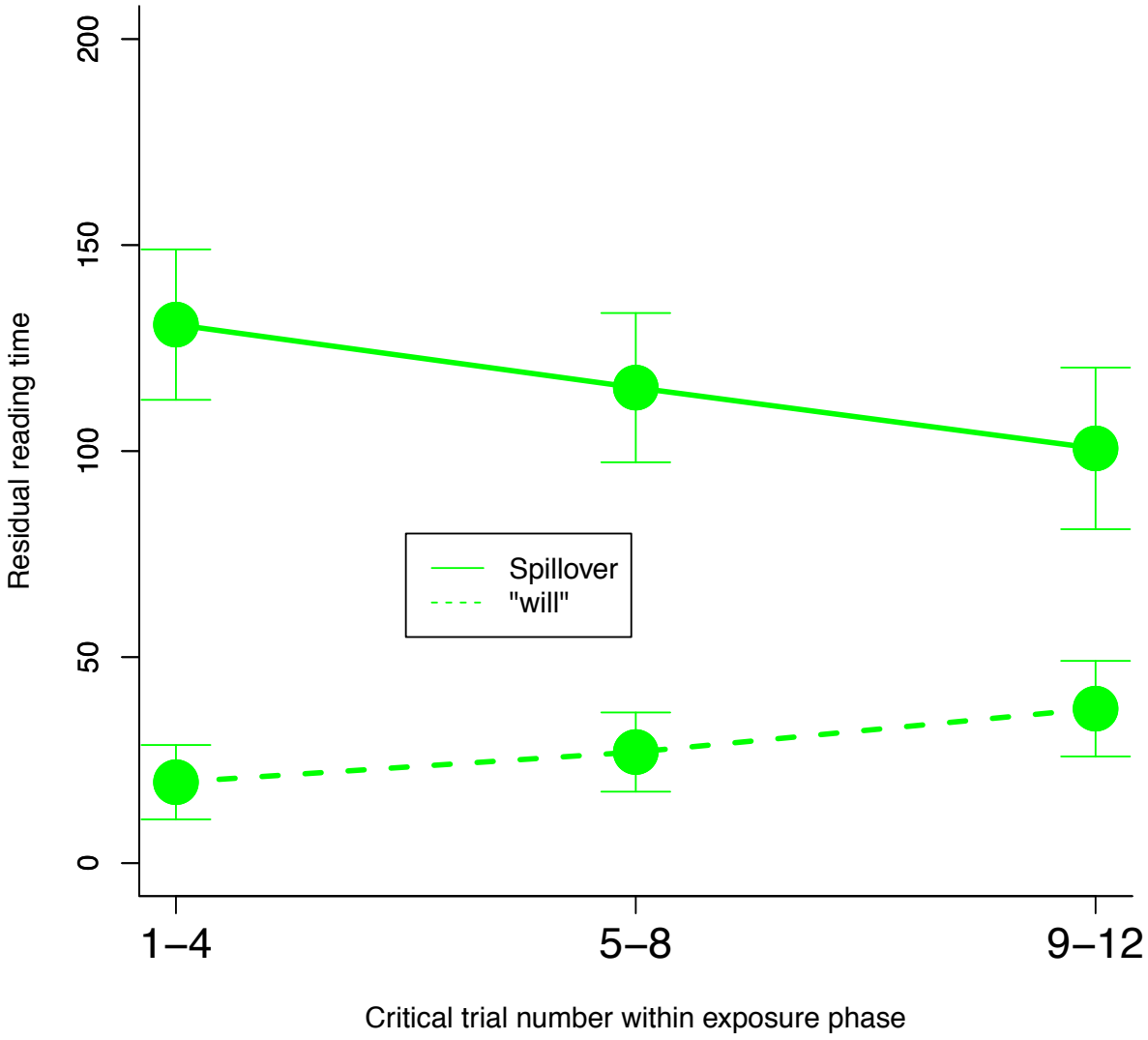

Figure 7. Residual reading times (corrected for word length, screen position, and task practice) in *be*-drop-exposure phase of Study 2A as a function of sentence region and critical trial number. Error bars represent ±1 standard error.

We compared these regions in a mixed-effects regression using two treatment-coded

contrasts: The first compared the disambiguation to the preceding unambiguous word *will*; the second compared the spillover region to the unambiguous *will*.

Table 3 summarizes the results. The unfamiliar *be*-drop structure was difficult for participants: Reading times were significantly elevated at the spillover region ($t = 5.57$, $p < .001$) and marginally so at the disambiguation ($t = 1.88$, $p = .06$) relative to the preceding region (the unambiguous word *will*). Note the fact that the difficulty occurred more robustly at the spillover region (rather than at the disambiguation itself) is common in self-paced reading and one of the reasons why, for example, studies on garden-paths often include one or two words of spillover in the disambiguation region (e.g., MacDonald et al., 1994; see also Smith & Levy, 2013).

Table 3
Fixed Effect Estimates for Linear Mixed Effects Regression of Residual Reading Times in Be-*Drop-Exposure Phase of Study 2A.*

Fixed effect	$\hat{\beta}$	*SE*	*t*	*p*
Intercept (baseline residual reading time at *will*)	28.12	8.49	3.31	< .001
Region: Disambiguation (vs. *will*)	30.99	16.49	1.88	.06
Region: Spillover (vs. *will*)	88.04	15.80	5.57	< .001
Critical trial number	2.05	2.42	0.85	.40
Disambiguation region (vs. *will*) x critical trial number	-5.65	2.87	-1.97	< .05
Spillover region (vs. *will*) x critical trial number	-7.23	2.96	-2.44	< .05

Note. SE = standard error.

As with Dialectal *Need*, participants appeared to adapt to this unfamiliar structure. The disambiguation of the *be*-drop structure showed changes in reading time above and beyond the baseline rate of change, suggesting adaptation to the *be*-drop structure in particular. Specifically, the difficulty at the spillover region of *be*-drop significantly diminished over the course of the experiment ($t = -2.44$, $p < .05$), as did the difficulty at the disambiguation ($t = -1.97$, $p < .05$). By

contrast, there was no such change at the unambiguous word *will* ($t = 0.85$, $p = .40$; indeed, the change was numerically in the opposite direction). These within-experiment changes in reading speed specifically at the locus of difficulty with the *be*-drop structure—relative to the previous, unambiguous *will* region—suggest that participants were adapting to the *be*-drop structure over the course of the training phase.

Test phase. The above analyses indicate[7] that participants exposed to Dialectal *Need* adapted to it over the course of the exposure phase, and participants instead exposed to *be*-drop adapted to that structure. The critical question was how these differences in exposure would affect reading of the *be*-drop sentences in the test phase: Would adaptation to Dialectal *Need* generalize to processing of *be*-drop, or would participants exposed specifically to *be*-drop have an advantage in processing that structure?

We examined residual reading times (displayed in Figure 8) in the critical sentences of the test phase, all of which involved *be*-drop. We examined reading times as a function of sentence region, critical trial number, and exposure condition (i.e., *needs*-exposure vs. *be*-drop-exposure), and their full factorial interaction. The regions of the *be*-drop sentences were compared using the same contrasts as above, and the other independent variables were centered.

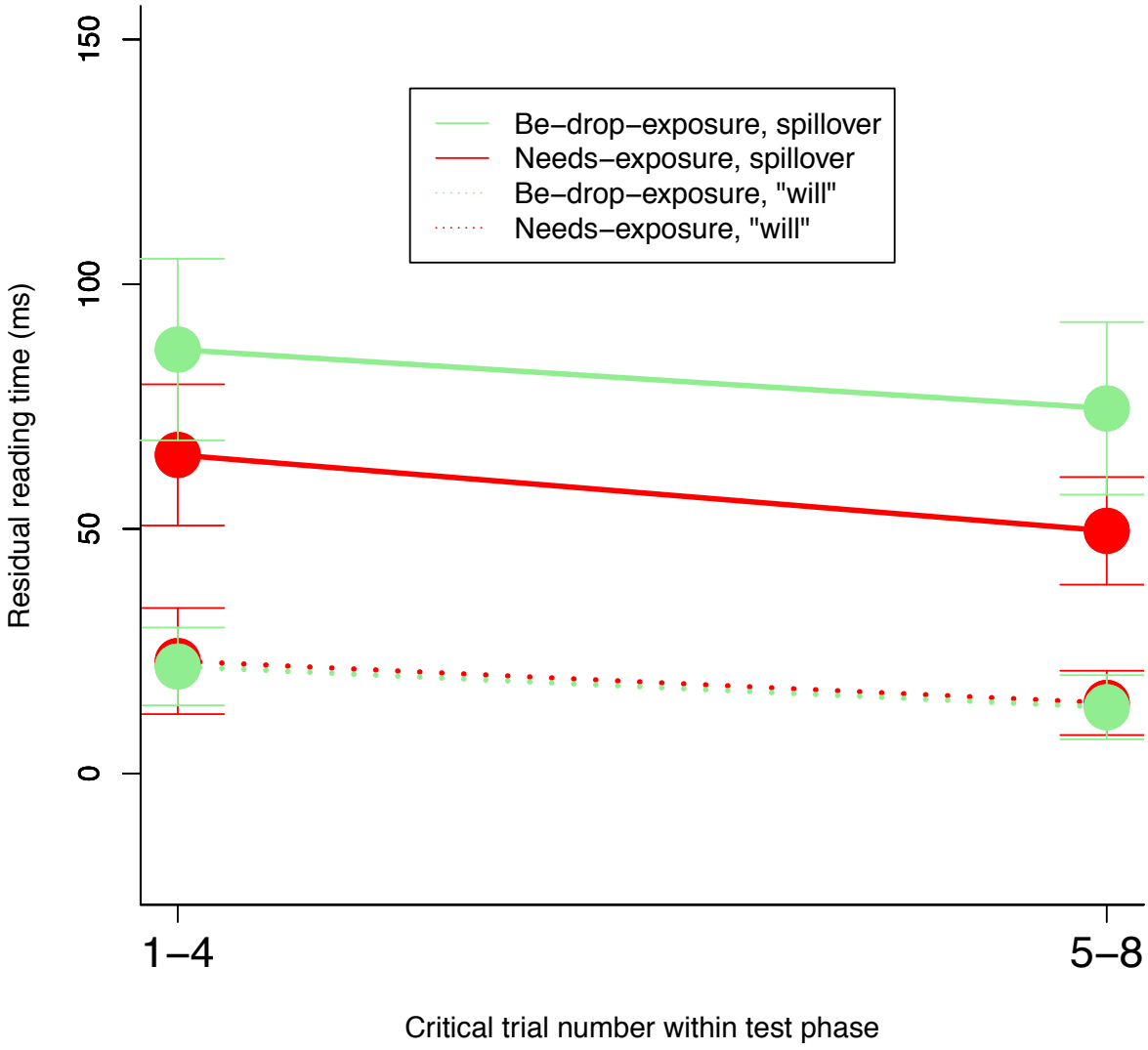

Figure 8. Residual reading time in *be*-drop sentences in Study 2A test phase as a function of exposure condition, sentence region, and critical trial number. Error bars represent ±1 standard error.

Table 4 summarizes the results. A main effect of Region indicated elevated reading times

for the *be*-drop sentences at the spillover region after the disambiguation ($t = 4.53$, $p < .001$), indicating that this region of the *be*-drop sentences continued to slow participants' reading (replicating the result from the *be-drop-exposure* phase, above). However, there was no effect of exposure condition, nor did exposure condition interact with any other variables (all *p*s>.20), indicating that the test *be*-drop sentences were not read differently by subjects previously exposed to *be*-drop itself versus Dialectal *Need*. In fact, the *needs*-exposure group was numerically faster.

Table 4
Fixed Effect Estimates for Linear Mixed Effects Regression of Residual Reading Times in Test Phase of Study 2A.

Fixed effect	$\hat{\beta}$	*SE*	*t*	*p*
Intercept (baseline residual reading time at *will*)	18.20	5.49	3.31	< .001
Region: Disambiguation (vs. *will*)	2.54	9.56	0.26	0.79
Region: Spillover (vs. *will*)	50.58	11.18	4.53	< .001
Critical trial number	-3.04	.03	-1.50	.13
Be-drop-exposure condition (vs. *needs*-exposure)	-0.99	10.00	-0.10	.92
Disambiguation region x critical trial number	2.59	3.11	0.83	.41
Spillover region x critical trial number	-2.40	3.34	-0.72	.47
Disambiguation region x *Be*-drop exposure condition	2.71	18.21	0.15	.88
Spillover region x *Be*-drop exposure condition	25.02	20.31	1.23	.22
Critical trial number x *Be*-drop exposure condition	-2.72	4.12	-0.66	.51
Disambiguation region x critical trial number x *Be*-drop exposure condition	-0.31	7.15	-0.04	.97
Spillover region x critical trial number x *Be*-drop exposure condition	1.38	6.81	0.20	.84

Note. SE = standard error.

Discussion

Study 2A appears to demonstrate generalization from one unfamiliar syntactic structure

to another. Participants exposed to one unfamiliar structure (Dialectal *Need*) were just as fast reading a different unfamiliar structure with a similar form (*be*-drop) as participants who had seen that latter structure throughout the entire experiment. This result is consistent with the hypothesis that participants are sensitive to the fact that dialectal differences tend to co-occur and thus they expected the unfamiliar dialect they were reading to contain other unfamiliar structures.

However, this conclusion depends on the lack of a difference between the two training conditions in the test phase. An alternate explanation of Study 2A involves the fact that reading times generally decline over the course of self-paced reading experiments, presumably as readers adapt to the self-paced reading task (Fine et al., 2010). Perhaps reading times had reached a floor at which we could no longer discern the effects of prior exposure. Some evidence *against* this account comes from the fact that reading times continued to decline over the course of the test phase; thus, it was not the case that reading times had already reached a floor by the start of the test phase. Nevertheless, to more directly address the possibility of a floor effect, we conducted Study 2B.

Study 2B

In Study 2B, we sought to demonstrate that the *be*-drop sentences in the test phase *could* cause difficulty for readers who had an equivalent number of self-paced reading trials—but who had no reason to expect unfamiliar structures. In particular, participants already familiar with Dialectal *Need* received the same Dialectal *Need* items as the *needs*-exposure condition in Study 2A. For existing users of Dialectal *Need*, *no* unfamiliar structure is being encountered during the exposure phase; consequently, they should have no reason to expect further unfamiliar structures in the test phase. (The *be*-drop structure is not actually part of the western Pennsylvania dialect and hence would be unfamiliar even to readers familiar with Dialectal *Need*.) Thus, if reading

times have not simply reached a floor, we predicted that the *be-drop* sentences in the test phase would be more unexpected—and thus read more slowly—for these subjects than for the corresponding *need-exposure* subjects in Study 2A, who had already encountered a structure unfamiliar to them (Dialectal *Need*). Or, if the lack of a difference in the test phase of Study 2A was merely a floor effect after many trials of self-paced reading, familiar subjects in Study 2B would exhibit equally low RTs since they have read the same sentences as the unfamiliar subjects in 2A.

Method

The procedure was identical to the *needs*-exposure condition of Study 2A except that participants were from Ohio/western Pennsylvania and were not naïve to Dialectal *Need*. We recruited 60 participants, whose locations are depicted in Figure 9. Of these, we discarded data from 1 participant who was certain they were unfamiliar with Dialectal *Need* and from 4 participants who incorrectly answered the "catch" question on the questionnaire, leaving 55 participants.

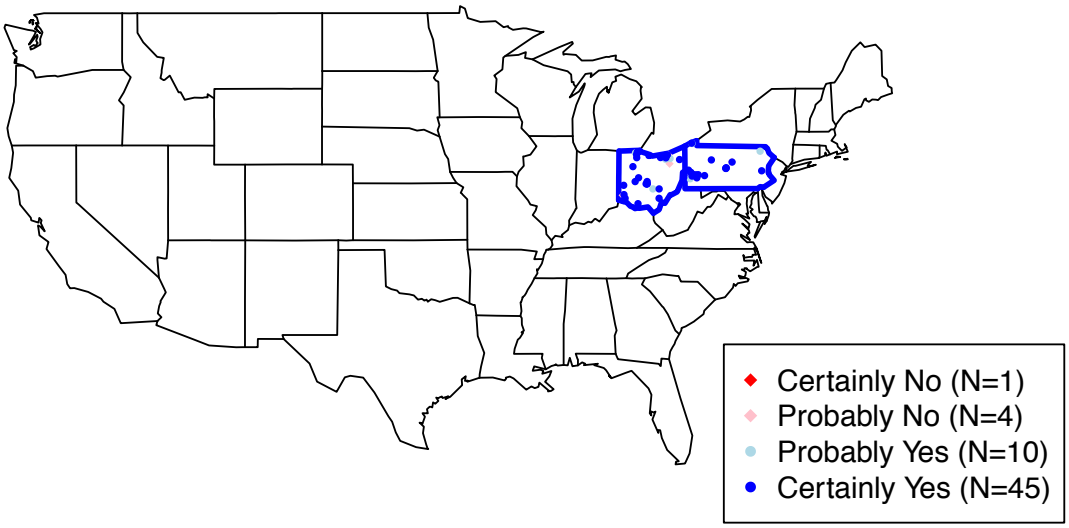

Figure 9. Map of target recruitment for Study 2B. Outlines indicate region targeted for recruitment of participants familiar with the Dialectal *Need* structure. Points indicate individual participants as a function of their previous experience with Dialectal *Need*, as self-reported on the post-experiment questionnaire.

Results

Exposure phase. We first examined reading times during the *needs*-exposure phase of Study 2B. These reading times are plotted above in Figure 6 in comparison to the *needs*-exposure subjects in Study 2A, who saw the same sentences but were unfamiliar with Dialectal *Need*. Like the unfamiliar participants (and replicating the pattern Study 1), the familiar participants showed some baseline differences across the regions in reading time, which might reflect differences in reading speeds due to the lexical or syntactic information conveyed in the regions or due to the word's serial position within the sentence.

Crucially, however, because participants in Study 2B were already familiar with Dialectal *Need*, we expected that a few additional encounters with Dialectal *Need* would not result in major changes in how they read this structure in particular. (And, indeed, participants in Study 1 who were already familiar with Dialectal *Need* did not show major changes in how they read Dialectal *Need* over the course of the experiment.) The mixed-effects regression of residual reading times, presented in Table 5, supported this prediction: There were no interactions between critical trial number and the specific sentence regions of the Dialectal *Need* structure (all *p*s >= .25); participants already familiar with Dialectal *Need* did not show any special changes in how they read the disambiguation of this structure as they progressed from 0 within-experiment exposures to 12 within-experiment exposures—presumably because these participants were *already* familiar with Dialectal *Need* even before the first critical trial.

Table 5
Fixed Effect Estimates for Linear Mixed Effects Regression of Residual Reading Times in Needs-*Exposure Phase of Study 2B.*

Fixed effect	$\hat{\beta}$	*SE*	*t*	*p*
Intercept (baseline residual reading time at *needs*)	22.23	6.93	3.21	< .01
Region: Participle (vs. *needs*)	-0.86	8.00	-0.11	.91
Region: Disambiguation (vs. *needs*)	25.90	7.14	3.63	< .001
Region: Spillover (vs. *needs*)	6.87	6.75	1.02	.31
Critical trial number	0.01	1.74	0.01	.99
Participle region x critical trial number	-2.26	1.96	-1.15	.25
Disambiguation region x critical trial number	0.48	2.18	0.22	.83
Spillover region x critical trial number	0.53	1.96	0.27	.79

Note. SE = standard error.

Test phase. The results from the exposure phase suggest that participants in Study 2A, who were already familiar with Dialectal *Need*, did not have to adapt to a new structure during the exposure phase. What consequences would this have for their subsequent encounter with a genuine new structure (*be*-drop) during the test phase?

We examined reading times on the critical *be*-drop sentences during the test phase, comparing the *needs*-exposure participants in Study 2B (who were already familiar with Dialectal *Need*) to the *needs*-exposure subjects already in Study 2A (who were not previously familiar with Dialectal *Need*). Figure 10 depicts reading times in both of these conditions.

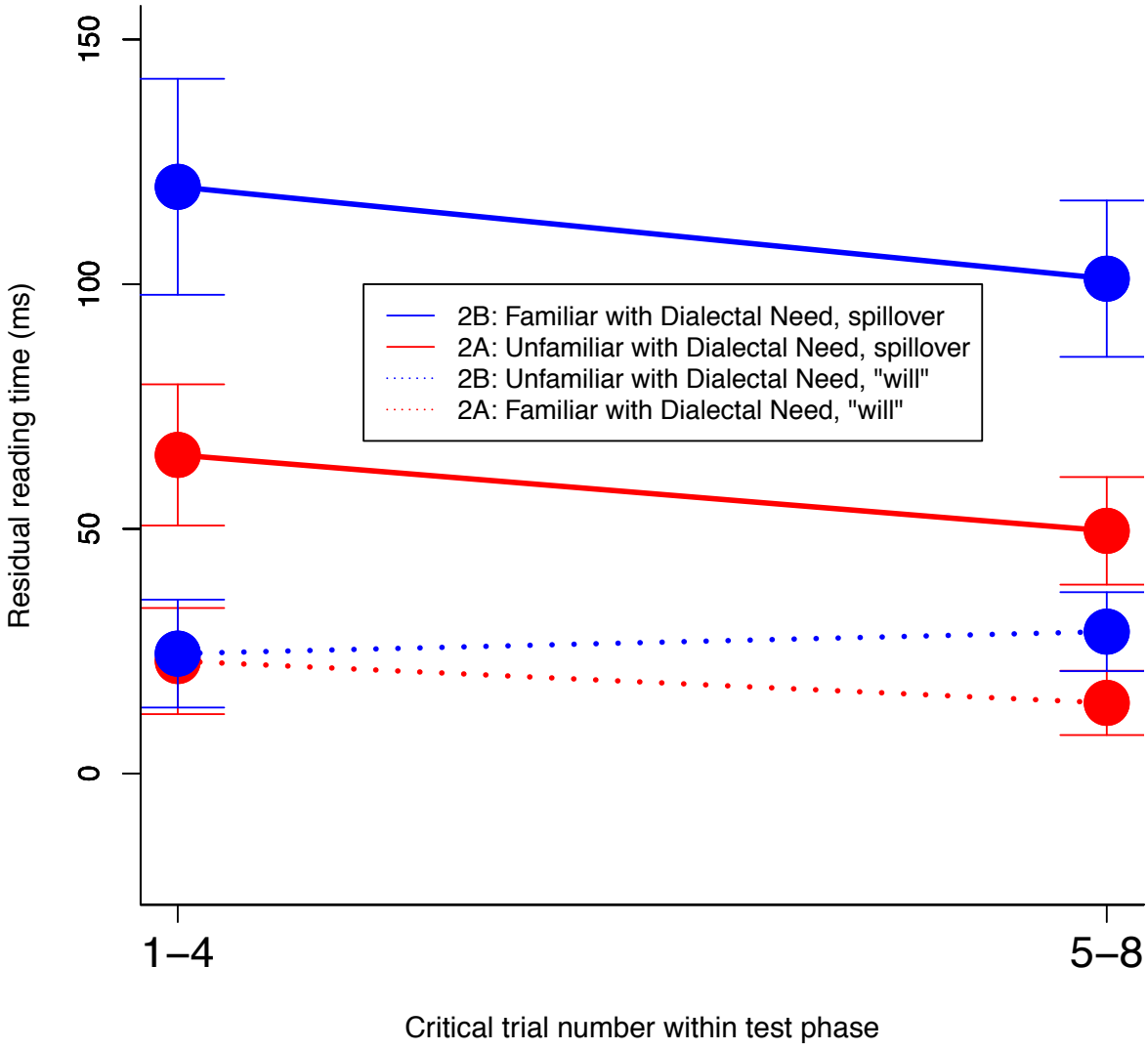

Figure 10. Residual reading times (corrected for word length, screen position, and task practice) in *be*-drop sentences in Study 2A and Study 2B test phase for participants initially exposed to Dialectal *Need*, as a function of prior familiarity with Dialectal *Need*, sentence region, and critical trial number. Error bars represent ±1 standard error.

We analyzed reading times as a function of participant group (familiar or unfamiliar; Helmert-coded), sentence region, and critical trial number, and their full factorial interaction. The results are presented in Table 6.

There was a main effect of region, with reading times significantly elevated at the spillover region overall ($t = 4.48$, $p < .001$), replicating the results presented above. This baseline difference across regions was not the result of primary interest as it might simply reflect region-to-region differences in lexical, syntactic, or other complexity. Rather, the crucial test was whether readers who had prior familiarity with Dialectal *Need* read this critical spillover region differently from those who had not. As predicted, a significant Participant Group x Sentence Region interaction indicated that reading times were especially inflated in the critical spillover region for participants *familiar* with Dialectal *Need* relative to *unfamiliar* participants, $t = 1.99$, $p < .05$. This effect cannot be attributed to overall differences between the groups in reading speed: There was no such familiarity effect at the word *will*, $t = 0.67$, $p = .50$, indicating that this effect of prior familiarity was localized to the disambiguation of the *be*-drop structure.

Table 6
Fixed Effect Estimates for Linear Mixed Effects Regression of Residual Reading Times in Test Phase for Needs-*Exposure Participants in Study 2A and Study 2B.*

Fixed effect	$\hat{\beta}$	*SE*	*t*	*p*
Intercept (baseline residual reading time at *will*)	22.61	6.40	3.54	< .001
Region: Disambiguation (vs. *will*)	5.65	10.48	0.54	0.59
Region: Spillover (vs. *will*)	61.12	13.64	4.48	< .001
Critical trial number	-0.78	2.40	-0.32	.75
Familiar with Dialectal *Need*	7.97	11.86	0.67	.50
Disambiguation region x critical trial number	4.85	3.76	1.29	.20
Spillover region x critical trial number	-4.38	3.78	-1.16	.25
Disambiguation region x familiarity	8.35	21.23	0.39	.70
Spillover region x familiarity	45.81	22.93	1.99	< .05
Critical trial number x familiarity	1.82	4.81	0.38	.70
Disambiguation region x critical trial number x familiarity	4.41	7.14	0.62	.54
Spillover region x critical trial number x familiarity	-2.81	8.10	-0.35	.73

Note. SE = standard error.

Discussion

In Study 2B, readers who were already familiar with Dialectal *Need* first encountered that structure, and then encountered the unfamiliar *be*-drop structure. These readers experienced difficulty in processing *be*-drop relative to readers who had been unfamiliar with Dialectal *Need* before the experiment. This result can be explained by the fact that comprehenders already familiar with Dialectal *Need* did not encounter a structure unfamiliar to them until they reached the critical *be*-drop sentences in the test phase. By contrast, the naïve subjects in Study 2A had already encountered an unfamiliar structure in the form of Dialectal *Need*. They may have inferred they were in a novel linguistic environment that was likely to contain other unfamiliar structures beyond Dialectal *Need*, thus facilitating processing of *be*-drop.

These significant results indicate the test phase in this procedure is indeed capable of detecting reliable effects of experience and that reading times have not simply reached a floor.

Therefore, it is unlikely that the non-significant difference between conditions in Study 2A was simply a floor effect after many self-paced reading trials (although it remains possible that, if the difference between conditions was of smaller magnitude in Study 2A than in 2B, it could have been obscured by a more mild or gradient floor effect). Rather, the results of Study 2A appear to reflect genuine syntactic generalization: Exposure to one unfamiliar structure (Dialectal *Need*) may have lead to greater expectations for—and facilitated processing of—other unfamiliar structures (or even unfamiliar linguistic elements, more generally), such as *be*-drop.

It is perhaps surprising that the familiar participants, whose first encounter with an unfamiliar structure in this experiment was the *be*-drop structure in the test phase, did not show adaptation to it (i.e., decreased reading times) over the course of the test phase. After all, participants in Study 2A *did* adapt to *be*-drop over the course of their initial encounter with it in the exposure phase. This discrepancy may reflect the prior input received in the experiment. In Study 2A, participants encountered *be*-drop from the beginning of the experiment and adapted rapidly. But, the familiar-group participants in Study 2B encountered *be*-drop only after a phase in which they received exclusively syntactic structures familiar to them, which provided de facto evidence that they were not encountering an unfamiliar dialect and should *not* expect unfamiliar structures. This behavior, if confirmed, would be consistent with our hypothesis that the language comprehension system is sensitive to syntactic contingencies and has less expectation for new syntactic structures when the other input suggests a familiar dialect. This pattern would also be consistent with results from other areas of language processing, such as phonetic adaption, in which exposure to familiar input blocks (or at least reduces) later adaptation to unfamiliar input (e.g., Kraljic, Samuel, & Brennan, 2008; Kraljic & Samuel, 2011). The potential effects of exposure to familiar input on subsequent adaptation are intriguing, but require

further investigation.

General Discussion

Language is variable at multiple levels: Talkers (and writers) differ in how they articulate particular speech sounds, what words they use, and what grammatical structures they put them in. Psycholinguistic theories must account for how comprehension often proceeds efficiently despite such variability. And, differences across varieties (including idio-, socio-, and dialects) imply that linguistic theory should account for the organization of the representations required for receptive competency in multiple varieties (e.g., Coseriu, 1955; Berthele, 2008).

We examined these questions in the domain of syntactic structures unique to a particular dialect, such as the Dialectal *Need* structure of Ohio and western Pennsylvania. In two studies, we investigated how readers process and generalize exposure Dialectal *Need*. In Study 1, readers unfamiliar with Dialectal *Need* initially encountered difficulty reading it, but rapidly adapted their processing to be more like readers already familiar with it. In Study 2A, exposure to Dialectal *Need* also facilitated other, similar unfamiliar structures. This generalization appears to have been driven by comprehenders' belief or inference that they were encountering a novel, unfamiliar dialect: In Study 2B, readers who were *already* familiar with the dialect that contains Dialectal *Need* did not expect other structures (not part of the western Pennsylvania dialect) simply on the basis of encountering Dialectal *Need*.

These results demonstrate that facilitatory effects of exposure in comprehension are not limited to known structures (see also Boland et al., 2015; Kaschak & Glenberg, 2004; Kaschak, 2006). Rather, even a structure that was previously completely unknown can be more rapidly processed after a few exposures. This ability is well-suited for a world in which syntactic differences among dialects and other environmental features (such as individual talkers; Kamide,

2012) may introduce entirely new structures to which comprehension must adapt.

Indeed, this incremental cumulative facilitation to novel structures closely resembles cumulative priming of known structures (Farmer, Monaghan, Misyak, & Christiansen, 2011; Farmer et al., 2014; Fine et al., 2010, 2013; Fine & Jaeger, 2016). This is expected if priming—like the learning of novel structures—is a form of implicit learning about syntax (Bock & Griffin, 2000; Chang et al., 2006; Dell & Chang, 2014; Fine & Jaeger, 2013; Jaeger & Snider, 2013). This process of implicit learning would allow comprehenders, even as adults, to continuously adapt their expectations about the distribution of syntactic structures that they will encounter; accurate expectations about upcoming material would, in turn, facilitate robust and efficient language understanding.

Another possible account of exposure effects on syntactic processing is that they reflect transient activation of structural representations (e.g., Pickering & Branigan, 1998; Traxler & Tooley, 2008). However, it is less clear how such mechanisms would apply to exposure to a structure with which one has no previous experience with whatever and of which one presumably has no representation (at least at the beginning of the experiment). Under such an account, the similarity in the effects of recent exposure for familiar and novel structures is a coincidence and requires two separate explanations. The implicit-learning hypothesis receives further support from the results of Study 2, which demonstrated that the effect of exposure to Dialectal *Need* exposure depends on comprehenders' pre-experiment knowledge about that structure. An effect of prior knowledge is natural in an account in which structural priming is part of an implicit learning process that is ongoing both pre- and within-experiment.

The results of Study 2 are also compatible with the additional claim that structural priming reflects the structure of the world, in which linguistic frequencies often covary. An

environment that introduces one unfamiliar structure—such as a writer with an unfamiliar dialect—is also more likely to contain others. Correspondingly, initial exposure to *needs*+participle also facilitated processing of other unfamiliar structures. This result is compatible with the principle that effective statistical learning requires knowledge of covariance between statistical distributions (Clark, 2013; Kleinschmidt & Jaeger, 2015; Qian, Jaeger, & Aslin, 2012); e.g., when encountering one unfamiliar structure, what *else* should be expected?

The present study leaves open on which dimensions this generalization occurs. We find that exposure to one unfamiliar structure also facilitated processing of another unfamiliar structure. However, because we tested only two unfamiliar structures, it is not clear whether comprehenders would similarly generalize to *all* unfamiliar structures. One possibility is that exposure to one unfamiliar structure would facilitate processing of *any* other unfamiliar structure; if, for instance, comprehenders simply came to expect unfamiliar structures (or unfamiliar linguistic input, even more broadly) in general. But, an alternate possibility is that exposure to one unfamiliar structure would generalize to *some* other structures but not to all; for instance, participants might generalize from Dialectal Need to *be*-drop but not to a less similar dialectal feature, such as positive *anymore* (Youmans, 1986). This type of learning would be expected under, among others, a *rational* or *ideal adapter* framework (e.g., Kleinschmidt & Jaeger, 2015) for comprehenders who had experience, for instance, that structures A and B tended to co-occur in the same environment but that structures A and C did not. (The same argument can be made about *types* of structures, such as omission alternations, as discussed in the introduction of Study 2.) Distinguishing these theoretical possibilities will be challenging: Ideal adapter models make precise predictions, but these predictions depend on good estimates of comprehenders' previous experience. Unlike for speech categories, for which larger cross-

dialectal quantitative databases are now available—allowing researchers to derive reasonable estimates of the language input that a "typical" comprehender has experienced—less is known about which syntactic structures actually *do* co-occur in the typical native American English speaker's experience of the world. Development of such databases is thus one way to advance this line of investigation. Even in the absence of direct experience with the co-occurrence of certain structures (or certain types of structures), comprehenders might generalize based on some (as of yet unknown) similarity metric across structures (as in exemplar-based accounts of syntactic presentations; Scha, Bod & Sima'An, 1999; Snider, 2008).

Perceived Grammaticality and Online Processing

One alternate account of participants' increasingly facilitated reading times for Dialectal *Need* and *be*-drop, and their generalization across structures, is that they are not learning these structures as grammatical structures per se but rather that participants are becoming more accommodating of unusual or "erroneous" input; for instance, by relaxing their criterion for what constitutes an acceptable utterance. This could be part of an explicit strategy or even a more implicit process of adjustment.

Some aspects of the present data seem to suggest this possibility does not fully account for the data. Specifically, participants were queried at the end of the experiment whether they thought the experiment contained unusual grammar (among other debriefing questions). First, among the participants previously unfamiliar with Dialectal *Need*, opinions were roughly even split in Study 1 between whether the experiment contained "unusual" grammar (47% yes, 53% no). (The majority of participants familiar with Dialectal *Need*—63%—reported no unusual grammar in Study 1; in Study 2, this question is less informative for probing interpretation of Dialectal *Need* because *all* participants see at least one unfamiliar structure: the "be"-drop

structure.) The fact that even many participants previously unfamiliar with Dialectal *Need* did not explicitly judge the grammar as unusual or suggests that the robust adaptation effect we observed across subjects is unlikely to be incumbent upon participants viewing the experiment as having unusual grammar.

Second, we conducted a post-hoc examination of whether participants' explicit perceptions of "unusual grammar" modulated their within-experiment adaptation in reading times. If participants' reading of Dialectal *Need* or *be*-drop sped up over the course of the experiment only because they had adopted a strategy of expecting and accepting "unusual" input, we would expect that only those participants who reported perceiving "unusual" grammar would show this adaptation. To test this prediction, we entered participants' perceptions of "unusual grammar" as an additional variable in the mixed effects regressions and allowed it to interact with sentence region and critical trial number.

Critically, participants' post-experiment grammaticality judgment did not in any case interact with the adaptation effect over the course of the experiment. In Study 1, for instance, the Disambiguation Region x Perceived Grammaticality x Critical Trial Number interaction was not significant, $t = -0.71$, $p = .48$, nor was the analogous interaction for the spillover region, $t = -0.15$, $p = .88$. The same was true in Study 2 both for the Dialectal *Need* structure and *be*-drop (all $ps > .10$). Instead, the critical interactions demonstrating adaptation remained significant even with the inclusion of the grammaticality judgment in all cases: for Dialectal *Need* in study 1 (Disambiguation Region x Familiarity x Critical Trial Number, $t = 2.80$, $p < .01$; Spillover Region x Familiarity x Critical Trial Number, $t = 2.80$, $p < .01$), for Dialectal *Need* in study 2A (Disambiguation Region x Critical Trial Number, $t = -2.28$, $p < .05$), and for *be*-drop in Study 2B (Spillover Region x Critical Trial Number, $t = -2.40$, $p < .05$). That is, participants' reading

times shifted over the experiment at the same rate regardless of whether or not they explicitly judged the grammar as unusual. These results are more consistent with an implicit learning account in which adaptation to a particular syntactic distribution implicitly proceeds apace regardless of comprehenders' explicit judgments.

Explicit judgments of grammaticality *did* relate to some aspects of participants' syntactic processing other than their within-experiment adaptation. Namely, in each study, there was an *overall* relationship, across trials, between perceived grammaticality and reading times in the critical region. For instance, in Study 1, a Disambiguation Region x Perceived Grammaticality interaction, $t = 2.26$, $p < .05$, indicated that participants who perceived the grammar as unusual read the disambiguation of Dialectal *Need* more slowly than those who did not. An analogous interaction also obtained in the spillover region, $t = 2.84$, $p < .01$, as well as in Study 2A for Dialectal *Need* in the disambiguation, $t = 2.73$, $p < .01$, and in Study 2A for *be*-drop in both the disambiguation ($t = 2.14$, $p < .05$) and spillover regions ($t = 4.54$, $p < .001$). This pattern suggests some potential link between participants' online syntactic processing and their perceptions of erroneous input. But, the direction of causality is left open: Slower processing times in the disambiguation region might make it more likely that participants perceive or remember the structure to be unusual, or the perception of ungrammaticality might cause additional slow down during reading. Crucially, however, in all cases the effect of perceived grammaticality was on the overall speed at which participants read the disambiguation, not the within-experiment adaptation that was of primary interest. Indeed, as noted above, the adaptation effects were not modulated by explicit judgments of grammaticality and remained significant even when controlling for those judgments.

This result is consistent with an implicit learning account, in which adaptation to a

distribution of syntactic structures proceeds implicitly regardless of comprehenders' explicit judgments, rather than a strategy for responding to input that the comprehenders perceived as unusual or erroneous. Nevertheless, the data here suggest further investigation is warranted into the relationship between metalinguistic perceptions of grammaticality and online language processing. More generally, fully disentangling adaptation to particular syntactic structures from comprehenders' general tolerance to "erroneous" or unusual input will likely acquire additional studies—for instance, by testing generalization to a wider variety of other unfamiliar or non-grammatical structures.

Other Caveats and Future Directions

Here, we report evidence that participants can rapidly adapt to reading unfamiliar syntactic structures and generalize that adaptation. However, there remain several important caveats for future studies to address. First, it is possible that participants' propensity to adapt the task was influenced by fact that they only needed to read for gist in order to answer the comprehension questions and were not tested on the grammatical details of the sentences. However, comprehension accuracy was still high; the mean accuracy on the comprehension questions was 95% both in Study 1 and in Study 2. Further, most reading is in fact done for gist rather than monitoring for grammatical detail (e.g., F. Ferreira & Patson, 2007; Graesser, Mills, & Zwaan, 1997), so our experiments examine adaptation under a comparatively ecological valid task.

Second, we tested participants' generalization to other structures using the *be*-drop structure, which is not actually an attested dialectal variant. The use of an artificial structure was motivated by the experimental control it offered—*be*-drop could be used in similar contexts as Dialectal *Need*, allowing us to hold constant the prior discourse and sentential context across

conditions—and by comparison with other experiments that have used these materials (e.g., Boland et al., 2015). In doing so, we were guided by work in other areas of language processing that have also used experimenter-created rules or materials, such as novel pseudo-dialectal vowel shifts (Weatherholtz, 2015), novel phonotactic rules (e.g., Dell, Reed, Adams, & Meyer, 2000) or nonce lexical items, such as *wug* or *blicket* (Berko-Gleason, 1958). (Note, too, that participants have no way to know that *be*-drop is an experiment creation rather than simply part of a dialect they simply have never encountered.) Nevertheless, it is possible that the fact that Dialectal *Need* is an attested grammatical feature and *be*-drop is not means that they differ on some property relevant to language processing, although it is not immediately clear what that property would be. It would be productive for future research to test generalization from one attested syntactic variant to another.

Comprehending Linguistic Variability

Our interest in the present study was in syntactic variability, but linguistic variability exists at multiple levels. Indeed, the variability of the linguistic signal has long been acknowledged and studied in phonetic and phonological perception, where it is known as the *lack of invariance* problem (Liberman, Cooper, Shankweiler, Studdert-Kennedy, 1967): Different talkers pronounce the same speech sound differently (Allen, Miller, & DeSteno, 2003; Hillenbrand, Getty, Clark, & Wheeler, 1995; Newman, Clouse, & Burnham, 2001), and even the same talker's productions vary as a function of circumstances such as external noise (Lombard, 1911); thus, there is no invariant signal that can be used to identify a particular speech sound. Work on speech perception has established that listeners are nevertheless able to succeed at speech perception by adapting to the characteristics of particular talkers (Goldinger, 1998; Johnson, 1990; Kraljic & Samuel, 2007; Pierrehumbert, 2002). For instance, experience with a

talker with non-standard realization of vowels (e.g., pronouncing *the wicked witch* as *the weckud wetch*) changes listeners' interpretations of subsequent input from that speaker (Maye et al, 2003; Weatherholtz, 2015; for review, see Weatherholtz & Jaeger, in press).

Further, there is preliminary evidence in speech perception that listeners can *generalize* this adaptation. Learning how a talker produces one consonant or vowel can change appropriate expectations about how they produce other, related sounds (Kraljic & Samuel, 2006; Maye et al., 2008; Weatherholtz, 2015). In addition to this generalization across linguistic elements, listeners also generalize across talkers. Importantly, this generalization does not appear to reflect insensitivity to talker identity but rather structured learning that respects whether or not the properties of one talker are likely to covary with the properties of another talker: Listeners *can* generalize experience a foreign or dialectal accent to a novel talker *if* the novel talker has the same accent or a similar accent, thus justifying the generalization (Baese-Berk, Bradlow, & Wright, 2013; Best et al., 2013; Bradlow & Bent, 2008; Sidaras et al., 2009; Weatherholtz, 2015; Ying et al., 2013); however, they do *not indiscriminately* generalize learning about one talker's speech categories to a different speaker (Eisner & McQueen, 2005; Kraljic & Samuel, 2006; Munson, 2011; Reinisch & Holt, 2014). Thus, adaptation of speech perception appears to reflect the structure of when generalization is or is not justified.

These findings from adaptation in speech perception appear remarkably similar, thus far, to those related to adaptation in syntactic processing. Other recent studies have suggested that syntactic expectations can be rapidly updated even as an adult for familiar structures (Fine et al., 2010, 2013; Wells et al., 2009; Thothathiri & Snedeker, 2008; Tooley et al., 2009; see also similar work on lexical processing, e.g., Creel et al., 2008; Metzing & Brennan, 2003; Yildirim et al., in press). And, there is preliminary evidence that these syntactic expectations are

conditioned, at least in part, on talker identity (Kamide, 2012). Further, the present experiments, and those by Kaschak and colleagues (Kaschak and Glenberg, 2004; Kaschak, 2006), demonstrate that comprehenders can both adapt to, and generalize from, exposure to entirely unfamiliar structures. Paralleling phonetic adaptation, comprehenders can generalize from an encounter with one structure to the same structure in a new syntactic context (Kaschak, 2002) or to a different, but similar structure (the present Study 2; see also ongoing work by Boland, Kaschak, and colleagues, first presented as Boland et al., 2015). These parallels tentatively suggest that more general processes or principles of adaptation may underlie adaptation at both the syntactic level and phonetic and phonological levels.

However, one open question is whether generalization in syntactic adaptation is subject to some of the same structured limits as generalization in phonetic adaptation. If this property also holds for syntactic adaptation, comprehenders who have encountered an unfamiliar structure should generalize less, or not generalize at all, to a structure that is less similar to the first structure than *be*-drop is to Dialectal *Need*. This prediction remains to be tested.

Representing and Learning Dialects and Languages

In the present study, we examined how comprehenders adapt to a relatively simple dialectal difference in which they needed to learn only one novel syntactic structure. In other cases, however, there may be far more to learn. In particular, research on bi- and multi-lingualism has characterized how language users represent and manage competencies in multiple *languages,* rather than dialects within a single language. This research has established that languages are not represented fully independently; a recently-learned second language can alter processing of the first (Dussias, 2004; Dussias & Sagarra, 2007 Van Hell & Dijkstra, 2002; for a recent review, see Pajak et al., 2016). If even a wholly different language affects native-language

processing, it is perhaps unsurprising that exposure to a dialect of the *same* language affects processing of other L1 varieties (including the standard).

Indeed, an emerging perspective (Pajak et al., 2016) is that language learning exists on a continuum from learning an entire second language (for which only a little existing linguistic knowledge is relevant) to adapting to an individual talker of one's own dialect (for which almost all existing knowledge is relevant). The present experiment presents the intermediate ground of learning a wholly *unfamiliar* (dialectal) syntactic structure within the learner's *native language*. However, there has been little cross-talk between research on implicit adaptation during L1 processing and L2/L*n* acquisition, and the implications of bilingualism research is often not considered in research on L1 processing. We argue that the similarity—or lack thereof—between within-language adaptation and L2/L*n* acquisition merits further investigation. The paradigms and methods used here present one way that this work may be carried out.

Targeted Participant Recruitment for Language Processing Studies

We provide further evidence that Web-based, crowdsourced data collection paradigms (e.g., Buhrmester, Kwang, & Gosling, 2012) can replicate patterns observed in laboratory studies of reading time, including those involving within-experiment changes in reading behavior (see also Craycraft & Jaeger, 2015; Fine & Jaeger, 2016; Fine et al., 2013).

Further, we introduce a geographically targeted web-based recruitment method. This method allowed us to recruit participants who varied in dialectal features relevant to structural priming and sentence processing. We used a combination of explicit queries (asking participants to self-report their location and their experience with different dialects) and implicit queries (querying the geolocation information in participants' web browsers) to identify relevant participants. This method was effective at identifying participants with the desired background

especially in Study 2, where nearly 90% of the participants recruited met our geographic/dialectal inclusion criteria. This suggests that these targeted recruitment methods can be especially effective when combined with methods that leverage large datasets (e.g., social media posts; Doyle, 2014) to identify the distribution of dialectal features. Our results validate these methods as identifying dialectal differences relevant to language processing: In both studies, the differences we identified in prior familiarity with Dialectal *Need* had consequences for participants' online reading time behavior.

Finally, we were able to successfully query participants' dialectal experience in our post-experiment questionnaire. One concern might have been that participants would not accurately report their dialectal background, either out of genuine lack of awareness or out of the demand characteristics of the questionnaire (e.g., a desire to affirm familiarity with all presented structures). Although we do not claim to have completely eliminated such issues, the use of a "catch" question (i.e., asking whether participants "noticed" a structure that did not actually appear in the experiment) may help mitigate these concerns. Again, the questionnaire is validated by the results of both Study 1 and Study 2, in which answers to the debriefing questionnaire predicted differences in participants' online reading of the relevant dialectal structure.

These methodological advances open the door to other comparisons relevant to structural priming. For example, future studies could use web-based recruitment and debriefing to identify participants with and without exposure to other geographically-conditioned syntactic structures, such as positive *anymore* (Youmans, 1986) or modal stacking (Di Paolo, 1998), as well as geographically conditioned variation in *relative* structural frequency (e.g., Bresnan & Ford, 2010). In addition, these methods may be useful for examining non-syntactic dialectal

differences, such as phonetic, phonological, or lexical differences.

Conclusion

Together, these results support a view in which adaptation to newly encountered syntactic structures is a process of implicit learning about a syntactic distribution. This learning includes not only the relative frequencies of structures within that distribution, but also entirely new structures, as well as other features (e.g., the presence of other unfamiliar structures) that may covary with the encountered structures. This claim is analogous to proposals that both cumulative effects of exposure to *known* structures and trial-to-trial facilitation of those structures (i.e., structural priming) be attributed to an implicit learning process (e.g., Fine & Jaeger, 2013; Fine et al., 2013; see also Dell & Chang, 2014), although it is not yet clear whether these phenomena truly reflect the *same* processes and representations (for discussion, see also Traxler & Tooley, 2008; Fine & Jaeger, 2016). More broadly, these results support the emerging view that the speed and accuracy of language processing depends at least in part on rapidly adapting to the present linguistic environment.

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*, 544-552. doi: 10.1121/1.1528172

Arai, M., & Keller, F. (2013). The use of verb-specific information for prediction in sentence processing. *Language and Cognitive Processes*, *28*, 525–560. doi:10.1080/01690965.2012.658072

Arai, M., & Mazuka, R. (2014). The development of Japanese passive syntax as indexed by structural priming in comprehension. *The Quarterly Journal of Experimental Psychology*, *67*, 60-78. doi: 10.1080/17470218.2013.7940454

Arai, M., van Gompel, R.P.G., & Scheepers, C. (2007). Priming ditransitive structures in comprehension. *Cognitive Psychology*, *57*, 218-250. doi: 10.1016/j.cogpsych.2006.07.001

Baese-Berk, M.M., Bradlow, A.R., & Wright, B.A. (2013). Accent-independent adaptation to foreign accented speech. *The Journal of the Acoustical Society of America*, *133*, EL174–EL180. doi:10.1121/1.4789864

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). lmer4: linear mixed-effects models using Eigen and S4. R package version 1.1-8. http://CRAN.R-project.org/package=lme4.

Bender, E.M. (2000). Syntactic variation and linguistic competence: The case of AAVE copula absence [Unpublished doctoral dissertation]. Stanford University.

Berko-Gleason, J. (1958). The child's learning of English morphology. *Word*, *14*, 150–177

Bernolet, S., & Hartsuiker, R.J. (2010). Does verb bias modulate syntactic priming? *Cognition*, *114*, 455-461. doi: 10.1016/j.cognition.2009.11.005

Berthele, R. (2008). Dialekt-Standard Situationen als embryonale Mehrsprachigkeit.

Erkenntnisse zum interlingualen Potenzial des Provinzlerdaseins. In K. J. Mattheier & A. Lenz (Eds.), Dialektsoziologie / Dialect Sociology / Sociologie du Dialecte. Sociolinguistica, Volume 22 (pp. 87-107). Tübingen: Niemeyer.

Best, C.T., Shaw, J.A., & Clancy, E. (2013). Recognizing words across regional accents: The role of perceptual assimilation in lexical competition. In F. Bimbot, C. Cerisara, C. Fougeron, G. Gravier, L. Lamel, F. Pellegrino, & P. Perrier (Eds.), *INTERSPEECH 2013: 14th Annual Conference of the International Speech Communication Association* (pp. 2128-2132). Lyon, Paris: International Speech Communication Association.

Bock, J K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, *18*, 355-387.

Bock, J.K., & Griffin, Z.M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General*, *129*, 177-192. doi: 10.1037//0096-3445.129.2.177

Boland, J., de los Santos, G., Carranza, J., & Kaschak, M. (2015, March). Self-paced reading time as a measure of learning novel constructions. Poster presented at CUNY 2015 Sentence on Human Sentence Processing.

Bolinger, D. 1972. *That's that*. The Hague: Mouton.

Boston, M.F., Hale, J.T., Vasishth, S., Kliegl, R. (2011). Parallel processing and sentence comprehension difficulty. *Language and Cognitive Processes*, *26*, 301-349. doi: 10.1080/01690965.2010.492228

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

Branigan H.P., Pickering, M.J., & McLean, J.F. (2005). Priming prepositional-phase attachment

during comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 468-481. doi: 10.1037/0278-7393.31.3.468

Bresnan, J. (2007). Is syntactic knowledge probabilistic?: Experiments with the English dative alternation. In S. Featherston & W. Sternefeld (Eds.), *Roots: Linguistics in search of its evidential base* (pp. 77-96). Berlin: Walter de Gruyter.

Bresnan, J., & Ford, M. (2010). Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language*, *86*, 168-213. doi: 10.1353/lan.0.0189

Bresnan, J., & Hay, J. (2008). Gradient grammar: An effect of animacy on the syntax of give in New Zealand and American English. *Lingua*, *118*, 245–259. doi:10.1016/j.lingua.2007.02.007

Chang, F., Dell, G.S., & Bock, J.K. (2006). Becoming syntactic. *Psychological Review*, *113*, 234-272. doi: 10.1037/0033-295X.113.2.234

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*, 181-204. doi: 10.1017/S0140525X12000477

Craycraft, N., & Jaeger, T.F. (2015, March). Adapting syntactic expectations in the face of changing cue informativity. Poster presented at CUNY 2015 Sentence on Human Sentence Processing, Los Angeles, CA.

Coseriu, E. (1956). *La geografía lingüística*. Univ. de la Republica, Facultad de Humanidades y Ciencias, Inst. de Filologia, Departamento de Lingüistica.

Creel, S.C., Aslin, R.N., & Tanenhaus, M.K. (2008). Heeding the voice of experience: The role of talker variation in lexical access. *Cognition*, *106*, 633-664. doi: 10.1016/j.cognition.2007.03.013

Dell, G.S., & Chang, F. (2014). The P-chain: Relating sentence production and its disorders to

comprehension and production. *Philosophical Transactions of the Royal Society B*, *369*, 20120394. doi:10.1098/rstb.2012.0394

Dell, G.S., Reed, K.D., Adams, D.R., & Meyer, A.S. (2000). Speech errors, phonotactic constraints, and implicit learning: A study of the role of experience in language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1355-1367. doi: 10.1037//0278-7393.26.6.1355

DeLong, K.A., Urbach, T.P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, *8*, 1117–1121. doi:10.1038/nn1504

Demberg, V., & Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition*, *109*, 193-210. doi:10.1016/j.cognition.2008.07.008

Di Paolo, M. (1989). Double modals as single lexical items. *American Speech*, *64*, 195– 224.

Dikker, S., & Pylkkänen, L. (2013). Predicting language: MEG evidence for lexical preactivation. *Brain and Language, 127*, 55–64. doi:10.1016/j.bandl.2012.08.004

Doyle, G. (2014). Mapping dialectal variation by querying social media. *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics*

Dussias, P.E. (2004). Parsing a first language like a second: The erosion of L1 parsing strategies in Spanish-English bilinguals. *International Journal of Bilingualism*, *8*, 355–371. doi:10.1177/13670069040080031001

Dussias, P.E., & Sagarra, N. (2007). The effect of exposure on syntactic parsing in Spanish–English bilinguals. *Bilingualism: Language and Cognition*, *10*, 101-116. doi:10.1017/S1366728906002847

Eisner, F., & McQueen, J.M. (2005). The specificity of perceptual learning in speech processing. *Perception & Psychophysics*, *67*, 224-238.

Elsness, J. (1984). *That* or zero? A look at the choice of object clause connective in a corpus of American English. *English Studies*, *65*, 519-533. doi: 10.1080/00138388408598357

Farmer, T.A., Monaghan, P., Misyak, J.B., & Christiansen, M.H. (2011). Phonological typicality influences sentence processing in predictive contexts: Reply to Staub, Grant, Clifton, and Rayner (2009). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 1318-1325. doi: 10.1037/a0023063

Farmer, T.A., Fine, A.B., Yan, S., Cheimariou, S., & Jaeger, T.F. (2014). Syntactic expectation adaptation in the eye-movement record. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Meeting of the Cognitive Science Society* (pp. 2181-2186). Austin, TX: Cognitive Science Society.

Federmeier, K.D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, *44*, 491–505. doi:10.1111/j.1469-8986.2007.00531.x

Ferreira, F., & Clifton, C., Jr. (1986). The independence of syntactic processing. *Journal of Memory and Language*, *25*, 348-368. doi: 10.1016/0749-596X(86)90006-9

Ferreira, F., & Patson, N. (2007). The "good enough" approach to language comprehension. *Language and Linguistics Compass*, *1*, 71-83. doi: 10.1111/j.1749-818x.2007.00007.x

Ferreira, V.S., Bock, K., Wilson, M.P., & Cohen, N.J. (2008). Memory for syntax despite amnesia. *Psychological Science*, *19*, 940-946. doi: 10.1111/j.1467-9280.2008.02180.x

Fine, A.B., & Jaeger, T.F. (2013). Evidence for implicit learning in syntactic comprehension. *Cognitive Science*, *37*, 578-591. doi: 10.1111/cogs.12022

Fine, A.B., & Jaeger, T.F. (2016). The role of verb repetition in cumulative syntactic priming in

comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. doi: 10.1037/xlm0000236

Fine, A.B., Jaeger, T.F., Farmer, T.A., & Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. *PLoS ONE*, *8*, e77661. doi: 10.1371/ journal.pone.0077661

Fine, A.B., Qian, T., Jaeger, T.F., & Jacobs, R.A. (2010). Is there syntactic adaptation in language comprehension? In *CMCL '10 Proceedings of the 2010 Workshop on Cognitive Modeling and Computational Linguistics* (pp. 18-26). Stroudsburg, PA: Association for Computational Linguistics.

Finegan, E., & Biber, D. (2001). Register variation and social dialect variation: The register axiom. In P. Eckert & J. R. Rickford (Eds.), *Style and sociolinguistic variation* (pp. 235-267). Cambridge: Cambridge University Press.

Finley, S. (2011). Generalization to novel consonants in artifical grammar learning. In L. Carlson, C. Hoelscher, & T. F. Shipley (Eds.), *Proceedings of the 33rd annual conference of the Cognitive Science Society* (pp. 3268–3273). Austin, TX: Cognitive Science Society.

Finley, S., & Badecker, W. (2009). Artificial language learning and feature-based generalization. *Journal of Memory and Language*, *61*, 423–437. doi:10.1016/j.jml.2009.05.002

Gebhart, A.L., Aslin, R.N., & Newport, E.L. (2009). Changing structures in midstream: Learning along the statistical garden path. *Cognitive Science*, *33*, 1087–1116. doi:10.1111/j.1551-6709.2009.01041.x

Goldinger, S.D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105*, 251–279. doi:10.1037/0033-295X.105.2.251

Graesser, A.C., Mills, K.K., & Zwaan, R.A. (1997). Discourse comprehension. *Annual Reviews of Psychology*, *48*, 163-189. doi: 10.1146/annurev.psych.48.1.163

Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In *NAACL '01: Proceedings of the Second Meeting of the North American Chapter of the Association for Computational Linguistics on Language Technologies* (pp. 1-8). Stroudsburg, PA: Association for Computational Linguistics.

Hillenbrand, J., Getty, J.A., Clark, M.J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America*, *97*, 3099-3111. doi: 10.1121/1.411872

Hofmeister, P., Jaeger, T.F., Arnon, I., Sag, I.A., & Snider, N. (2013). The source ambiguity problem: Distinguishing the effects of grammar and processing on acceptability judgments. *Language and Cognitive Processes*, *28*, 48–87. doi:10.1080/01690965.2011.572401

Jaeger, T.F. (2006). Redundancy and reduction in spontaneous speech. [Unpublished doctoral dissertation]. Stanford University.

Jaeger, T.F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology*, *61*, 23-62. doi: 10.1016/j.cogpsych.2010.02.002

Jaeger, T.F., & Snider, N.E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*, *127*, 57-83. doi: 10.1016/j.cognition.2012.10.013

Johnson, K. (1990). The role of perceived speaker identity in F0 normalization of vowels. *The Journal of the Acoustical Society of America*, *88*, 642-654. doi:10.1121/1.399767

Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, *20*, 137-194.

Kamide, Y. (2012) Learning individual talkers' structural preferences. *Cognition*, *124*, 66-71. doi: 10.1016/j.cognition.2012.03.001

Kamide, Y., Altmann, G.T.M., & Haywood, S.L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, *49*, 133–156. doi:10.1016/S0749-596X(03)00023-8

Kaschak, M.P., & Glenberg, A.M. (2004). This construction needs learned. *Journal of Experimental Psychology: General*, *133*, 450-467. doi: 10.1037/0096-3445.133.3.450

Kaschak, M.P. (2006). What this construction needs is generalized. *Memory & Cognition*, *34*, 368-379.

Kaschak, M.P., Kutta, T.J., & Jones, J.L. (2011). Structural priming as implicit learning: Cumulative effects and individual differences. *Psychonomic Bulletin & Review*, *18*, 1133-1139. doi: 10.3758/s13423-011-0157-y

Kim, S., & Mauner, G. (2006, March). Priming vs. contingency learning accounts of structural priming effects in comprehension. Poster presented at CUNY Conference on Human Sentence Processing, New York, NY.

Kleinschmidt, D.F., & Jaeger, T.F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review*, *122*, 148-203. doi: 10.1037/a0038695

Kraljic, T., & Samuel, A.G. (2006). Generalization in perceptual learning for speech. *Psychonomic Bulletin & Review*, *13*, 262–268. doi:10.3758/BF03193841

Kraljic, T., & Samuel, A.G. (2007). Perceptual adjustments to multiple speakers. *Journal of Memory and Language*, *56*, 1-15. doi: 10.1016/j.jml.2006.07.010

Kraljic, T., & Samuel, A.G. (2011). Perceptual learning evidence for contextually-specific representations. *Cognition*, *121*, 459-465. doi: 10.1016/j.cognition.2011.08.015

Kraljic, T., Samuel, A.G., & Brennan, SE. (2008). First impressions and last resorts: How

listeners adjust to speaker variability. *Psychological Science*, *19*, 332-338. doi: 10.1111/j.1467-9280.2008.02090.x

Kuperberg, G., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition, and Neuroscience*, *31*, 32-59. doi: 10.1080/23273798.2015.1102299

Kuperman, V., Dambacher, M., Nuthmann, A., & Kliegl, R. (2010). The effect of word position on eye-movements in sentence and paragraph reading. *The Quarterly Journal of Experimental Psychology*, *63*, 1838-1857. doi: 10.1080/17470211003602412

Labov, W. (1969). Contraction, deletion, and inherent variability of the English copula. *Language*, *45*, 715–762.

Levelt, W.J.M., & Kelter, S. (1982). Surface form and memory in question answering. *Cognitive Psychology*, *14*, 78-106. doi: 10.1016/0010-0285(82)90005-6

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*, 1126-1177. doi: 10.1016/j.cognition.2007.05.006

Liberman, A.M., Cooper, F.S., Shankweiler, D.P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*, 431–461. doi:10.1037/h0020279

Lombard, E. (1911). Le signe de l'elevation de la voix. *Ann. Maladies Oreille, Larynx, Nez, Pharynx*, *37*, 101-119.

MacDonald, M.C. (1999). Distributional information in language comprehension, production, and acquisition: Three puzzles and a moral. In B. MacWhinney (Ed.), *The emergence of language* (pp. 177-196). Mahwah, NJ: Lawrence Erlbaum Associates.

MacDonald, M.C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, *4*, 226. doi:10.3389/fpsyg.2013.00226

MacDonald, M.E., Pearlmutter, N.J., & Seidenberg, M.S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*, 676-703.

Maye, J., Aslin, R.N., & Tanenhaus, M.K. (2008). The weckud wetch of the wast: Lexical adaptation to a novel accent. *Cognitive Science*, *32*, 543–562.

McRae, K., Spivey-Knowlton, M.J., & Tanenhaus, M.K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, *38*, 283–312. doi:10.1006/jmla.1997.2543

Metzing, C., & Brennan, S.E. (2003). When conceptual pacts are broken: Partner-specific effects on the comprehension of referring expressions. *Journal of Memory and Language*, *49*, 201-213. doi: 10.1016/S0749-596X(03)00028-7

Munson, C. (2011). Perceptual learning in speech reveals pathways of processing [Unpublished doctoral dissertation]. University of Iowa.

Murray, T.E., Frazer, T.C., & Simon, B.L. (1996). *Needs* + past participle in American English. *American Speech*, *71*, 255-271.

Newman, R.S., Clouse, S.A., & Burnham, J.L. (2001). The perceptual consequences of within-talker variability in fricative production. *The Journal of the Acoustical Society of America*, *109*, 1181-1196. doi: 10.1121/1.1348009

Norris, D., McQueen, J.M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, *47*, 204–238. doi:10.1016/S0010-0285(03)00006-9

Pajak, B., Fine, A.B., Kleinschmidt, D.F., & Jaeger, T.F. (2016). Learning additional languages as hierarchical probabilistic inference: insights from L1 processing. *Language Learning*. Advance online publication. doi: 10.1111/lang.12168

Pajak, B., & Levy, R. (2014). The role of abstraction in non-native speech perception. *Journal of*

Phonetics, *46*, 147-160. doi: 10.1016/j.wocn.2014.07.001

Pickering, M.J., & Branigan, H.P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, *39*, 633-651. doi: 10.1006/jmla.1998.2592

Pickering, M.J., & Ferreira, V.S. (2008). Structural priming: A critical review. *Psychological Bulletin*, *134*, 427-459. doi: 10.1037/0033-2909.134.3.427

Pierrehumbert, J. (2002). Word-specific phonetics. In C. Gussenhoven & N. Warner (Eds.), *Laboratory phonology VII* (pp. 101-139). Berlin: Mouton de Gruyter.

Qian, T., Jaeger, T.F., & Aslin, R.N. (2012). Learning to represent a multi-context environment: More than detecting changes. *Frontiers in Psychology*, *3*, 228. doi: 10.3389/fpsyg.2012.00228

Reinisch, E., & Holt, L.L. (2014). Lexically guided phonetic retuning of foreign-accented speech and its generalization. *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 539–555. doi:10.1037/a0034409

Reitter, D., Keller, F., & Moore, J.D. (2011). A computational cognitive model of syntactic priming. *Cognitive Science*, *35*, 587-637. doi: 10.1111/j.1551-6709.2010.01165.x

Rohdenburg, G. (1996). Cognitive complexity and increased grammatical explicitness in English. *Cognitive Linguistics*, *7*, 149-182. doi: 10.1515/cogl.1996.7.2.149

Roland, D., Elman, J.L., & Ferreira, V.S. (2006). Why is *that*? Structural prediction and ambiguity resolution in a very large corpus of English sentences. *Cognition*, *98*, 245-272. doi: 10.1016/j.cognition.2004.11.008

Scha, R., Bod, R., & Sima'An, K. (1999). A memory-based model of syntactic analysis: data-oriented parsing. *Journal of Experimental & Theoretical Artificial Intelligence*, *11*, 409-

440. doi: 10.1080/095281399146481

Sidaras, S., Alexander, J.E.D., & Nygaard, L. C. (2009). Perceptual learning of systematic variation in Spanish-accented speech. *The Journal of the Acoustical Society of America*, *125*, 3306-3316. doi: 10.1121/1.3101452

Smith, R., & Levy, R. (2013). The effect of word predictability on reading time is logarithmic. *Cognition*, *128*, 302-319. doi; 10.1016/j.cognition.2013.02.013

Snider, N. (2008). An exemplar model of syntactic priming [Unpublished doctoral dissertation]. Stanford University.

Staub, A., & Clifton, C. (2006). Syntactic prediction in language comprehension: Evidence from *either...or*. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *32*, 425–436. doi:10.1037/0278-7393.32.2.425

Tagliamonte, S., & Smith, J. (2005). No momentary face! The *zero* 'complementizer' in English dialects. *English Language and Linguistics*, *9*, 289-309. doi: 10.1017/S1360674305001644

Tagliamonte, S., Smith, J., & Lawrence, H. (2005). No taming the vernacular! Insights from the relatives in northern Britain. *Language Variation and Change*, *17*, 75–112.

Tenny, C. (1998). Psych verbs and verbal passives in Pittsburghese. *Linguistics*, *36*, 591– 597.

Thothathiri, M., & Snedeker, J. (2008). Give and take: Syntactic priming during spoken language comprehension. *Cognition*, *108*, 51-68. doi: 10.1016/j.cognition.2007.12.012

Tooley, K.M., & Bock, K. (2014). On the parity of structural persistence in language production and comprehension. *Cognition*, *132*, 101-136. doi: 10.1016/j.cognition.2014.04.002

Tooley. K.M., Traxler, M.J., & Swaab, T.Y. (2009). Electrophysiological and behavioral evidence of syntactic priming in sentence comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 19-45. doi: 10.1037/a0013984

Tottie, G., & Rey, M. (1997). Relativization strategies in Earlier African American Vernacular English. *Language Variation and Change*, *9*, 219-247. doi: 10.1017/S0954394500001885

Traxler, M.J. (2008). Lexically independent priming in online sentence comprehension. *Psychonomic Bulletin & Review*, *15*, 149-155. doi: 10.3758/PBR.15.1.149

Traxler, M.J., & Tooley, K.M. (2008). Priming in sentence comprehension: Strategic or syntactic? *Language and Cognitive Processes*, *23*, 609-645. doi: 10.1080/01690960701639898

Van Hell, J.G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review*, *9*, 780–789. doi:10.3758/BF03196335

Wasow, T., Jaeger, T.F., & Orr, D.M. (2011). Lexical variation in relativizer frequency. In H.J. Simon & H. Wiese (Eds.), *Expecting the unexpected: Exceptions in grammar* (pp. 175-195). Berlin: Walter de Gruyter.

Weatherholtz, K. (2015). Perceptual learning of systemic cross-category vowel variation [Unpublished doctoral dissertation]. The Ohio State University.

Weatherholtz, K., & Jaeger, T. F. (in press). Speech perception and generalization across talkers and accents. For *Oxford Research Encyclopedia of Linguistics*.

Weatherholtz, K., Seifeldin, M., Kleinschmidt, D. F., Kurumada, C., & Jaeger, T. F. (2016). Speech perception as probabilistic inference under uncertainty based on social-indexical knowledge. Manuscript under review.

Weiner, E. J., & Labov, W. (1983). Constraints on the agentless passive. *Journal of Linguistics*, *19*, 29. doi:10.1017/S0022226700007441

Wells, J.B., Christiansen M.H., Race, D.S., Acheson, D.J., & MacDonald, M.C. (2009).

Experience and sentence processing: Statistical learning and relative clause comprehension. *Cognitive Psychology*, *58*, 250-271. doi: 10.1016/j.cogpsych.2008.08.002

Wolfram, W. (1974). The relationship of white Southern speech to Vernacular Black English. *Language*, *50*, 498-527. doi: 10.2307/412221

Yildirim, I., Degen, J., Tanenhaus, M.K., & Jaeger, T.F. (in press). Talker-specificity and adaptation in quantifier interpretation. *Journal of Memory and Language*.

Ying, J., Shaw, J.A., & Best, C.T. (2013). L2 English learners' recognition of words spoken in familiar versus unfamiliar English accents. In F. Bimbot, C. Cerisara, C. Fougeron, G. Gravier, L. Lamel, F. Pellegrino, & P. Perrier (Eds.), *INTERSPEECH 2013: 14th Annual Conference of the International Speech Communication Association* (pp. 2108-2112). Lyon, Paris: International Speech Communication Association.

Youmans, G. (1986). Any more on anymore: Evidence from a Missouri dialect survey. *American Speech*, *61*, 61–75.

Footnotes

[1]A second syntactically valid interpretation of this string is one in which *photocopier needs* is a compound noun and *recycled* is the matrix verb. This interpretation is largely nonsensical, and it is unlikely that participants considered this interpretation much or at all. Moreover, even if participants did initially interpret the sentence in this very unlikely way, the compound-noun interpretation would be also ruled out at the same sentence region that rules out the modifier interpretation. That is, at the disambiguation region, Dialectal *Need* becomes the *only* interpretation possible. Thus, reading time at the disambiguation region is still a valid measure of whether participants considered Dialectal *Need,* regardless of whether readers previously considered the modifier interpretation or the unlikely and nonsensical compound-noun interpretation.

[2]These two variables were crossed with an additional cover story manipulation—whether the fictive company was putatively located in Denver or Pittsburgh. This manipulation had no effect. As we detail in Appendix E, participants in our studies generally have little knowledge about the geographic distribution of the Dialectal Need feature; thus, it is unsurprising that the putative location of the writer would affect comprehension. Consequently, we omit further discussion of this variable.

[3]One of the comprehension questions for the item with the participle *edited* had low mean accuracy (52%), which suggested this question may have been misleading. For this paragraph, we counted the item as correctly understood as long as participants correctly answered the other

question.

[4]The critical sentences showed the same pattern of results even if we did not regress out overall serial position, except that we observed a main effect of critical trial number on reading time across all participant groups and sentence regions, consistent with the idea that this reflects adaptation to the self-paced reading task. Controlling for serial position during residualization has the advantage that it allows clearer visualizations of the effects of interest, since sentence regions that are not differentially affected by *syntactic* adaptation are now expected to show no effect of serial position.

[5]The use of Dialectal *Need* in Pennsylvania English has been argued to actually be more similar, in both its meaning and in the constraints on its use, to Standard American English *need* + gerund (Doyle, 2008; Murray et al., 1996), as in *This car needs washing*. However, participants in our experiment who are unfamiliar with Pennsylvania English do not have access to that knowledge.

[6]Four out of 40 critical comprehension questions and two filler questions had unusually low accuracy rates (less than 80% accurate), suggesting that something about these questions may have been misleading to the participants. Since these questions were unlikely to reflect participants' actual comprehension, we did not exclude reading times if participants answered one of these questions incorrectly.

[7]Inspection of the change across critical trial numbers in Figures 6 and 7 might suggest that

the Dialectal *Need* and *be*-drop structures differed somewhat in the rate at which reading times adapted to them (as suggested by an anonymous reviewer). To assess this possibility, we performed a new linear mixed-effects regression using the reading times in the critical region of each of the two structures (the disambiguation region of Dialectal *Need* and the spillover region of *be*-drop). The Structure x Critical Trial Number interaction in this regression was not significant ($t = -0.69$, $p = .49$), providing no significant evidence that the rate of adaptation was significantly differ for *be*-drop than for Dialectal *Need*.

Appendix A

Self-paced reading materials and comprehension questions for Study 1. The locations of the line breaks when the stimuli were displayed to the participants are indicated with a / for the critical items. (We did not have any particular constraints on the location of the line breaks in the filler items.)

Practice items

1. I might be in a little late tomorrow as I'm dropping off my car for service. Please direct / any urgent calls to Tara.

Q1. Will this person be in early tomorrow? N

Q2. Should calls be directed to Tara? Y

2. A treat to brighten your work day: There are homemade brownies in the kitchen. Come get one while they last!

Q1. Are the brownies in the kitchen? Y

Q2. Are the brownies homemade? Y

3. How's everything going? Did you ever watch the movies I loaned you? Hope you liked them as much as I did!

Q1. Did the e-mail writer loan someone books? N

Q2. Did this person hate the movies? N

4. Think I might have left the heat up too high when I left for work this morning. When you get home, could you check the thermostat? Thanks!

Q1. Is this person worried about the thermostat? Y

Q2. Was the heat too low? N

Critical items

1. Why haven't you finished getting Order 437 ready? The / complete shipment needs boxed before our delivery truck / goes out this afternoon. I need you to do that ASAP.

Q1. Is the delivery truck going out today? Y

Q2. Is Order 437 ready? N

2. I got the manuscripts you forwarded to me. It looks / like we're ready to go with the magazine article. The / science book needs edited because there are still some / typos. Let me know when it's done.

Q1. Is the magazine article unfinished? N

Q2. Does the book have typos? Y

3. Not quite ready yet for my part of the big speech. The / computer presentation needs downloaded to my Mac first. / Otherwise, I'm set.

Q1. Has the presentation been downloaded yet? N

Q2. Is this person completely ready for the speech? N

4. I'll probably be working late tonight. Whenever you get / home, would you go ahead and get dinner started? The / pork roast needs cooked for 90 minutes and ought to get / an early start. Thanks and love you.

Q1. Is this person working late? Y

Q2. Are they cooking a roast? Y

5. I'm certainly excited to have a new award to display. / The display case needs polished before we add the / trophy. Let's do that this afternoon.

Q1. Will the display case be polished this afternoon? Y

Q2. Is there a new award? Y

6. Thanks for organizing the end-of-year party. I / understand that most of the prep work is done. The /whole kitchen needs cleaned after the party as well. / Can we get a volunteer for that, too?

Q1. Is the preparation for the party mostly done? Y

Q2. Is the party celebrating the middle of the year? N

7. Heard back from the consultant. The biggest issue she / noted is that our storefronts don't grab the eye. The / store entrance needs decorated if we want to capture / shoppers' attention. I think we can fix that.

Q1. Did the storefronts grab shoppers' attention? N

Q2. Did they think the problem can be fixed? Y

8. I don't want to focus on new software features until / we've finished this year's worldwide release. The / current software needs translated before we add / anything new. Let's make that a priority.

Q1. Is the worldwide release finished? N

Q2. Will translation be a priority? Y

9. Hi, I'm doing a report on our Web sales, but I only / have the first quarter numbers, which are out of date. / The sales report needs updated with the latest data. / Can you help me with that?

Q1. Is the report about Web sales? Y

Q2. Does the writer have the latest sales numbers? N

10. Anyone else keep hearing that clattering sound? I think / the vent in the corner is getting a little loose. The / metal plate needs tightened if that's the case. It's / getting pretty annoying.

Q1. Is it a fan that's loose? N

Q2. Is the vent too tight? N

Filler items

1. Sorry, it's gonna take me a while to get to your maintenance request. Swamped with e-mails right now and can't start on any new issues. Maybe tomorrow.

Q1. Does this person have a lot of e-mails? Y

Q2. Will the maintenance request be completed today? N

2. Had a great time in Hawaii, but happy to be home now. We got stuck in O'Hare for 5 hours. Horrible way to spend an evening!

Q1. Did this person go to California? N

Q2. Was the person stuck in O'Hare for five hours? Y

3. Not going to be able to make it into work today due to the snow. I'll be available on email/phone from home. See you tomorrow.

Q1. Did rain keep this person from getting into work? N

Q2. Is this person still available over email? Y

4. I received your request about replacing about the microwave in the break room. I'm afraid that's not in the current budget. The old one still works, so we'll have to stick with it.

Q1. Does the old microwave work? Y

Q2. Is the microwave being replaced? N

5. It's that time of year again! My daughter is selling Girl Scout cookies, so here's your chance to get your Thin Mints. I'll put the order sheet in the break room.

Q1. Is the daughter selling magazine subscriptions? N

Q2. Is the order sheet in the break room? Y

6. As a reminder, the fax machine is for business use only. Personal accounts have been discontinued. Please direct any questions to Margie.

Q1. Can the fax machine be used for personal use? N

Q2. Should questions be directed to Dave? N

7. Tomorrow's staff meeting has been moved from 9 to 11 AM. Sarah has a dentist appointment early in the morning. Apologies for the late notice.

Q1. Has the meeting time been changed? Y

Q2. Does Sarah have a doctor's appointment? N

8. Will you be home this afternoon? The delivery people just called me here at work. They can deliver the new sofa if you'll be around to let them in.

Q1. Is a refrigerator going to be delivered? N

Q2. Is the delivery scheduled for the morning? N

9. This is the last call for feedback on the candidate we interviewed today for accounting. If you have comments, please get them to me by the close of business today. Thanks for your input.

Q1. Was there a job interview? Y

Q2. Are the comments due today? Y

10. Has anyone else been getting Java errors on their computer? It's been happening to me since this morning. What can I do about this?

Q1. Is this person having computer problems? Y

Q2. Did the problems start this morning? Y

Appendix B

Post-experiment questionnaire for Study 1.

Please answer the following questions about the experiment you just completed.

1. Did you think there was anything odd about the sentences you read? Y N

2. Did you think that some of the sentences had unusual grammar to them? Y N

3. Did you see any sentences with structures like *The car needs washed*? Y N

4. Did you see any sentences with structures like *I might could pass this test*? Y N

5. Outside of this experiment, have you heard people say things like *The car needs washed*?

Y N

6. There are some areas where people say things like *The car needs washed*.

Do you know where? Y N

7. Have you heard that people in Denver say *The car needs washed*? Y N

8. Have you heard that people in Pittsburgh say *The car needs washed*? Y N

9. What do you consider your native language? _____

10. What is the first language you ever learned? _____

10b. If you learned two first languages simultaneously, what was the other one?

If you did not learn two first languages, please enter: none

_____

11. Where (city, state, country) did you learn English? _____

12. Where (city, state, country) do you live now? _____

13. Do you regularly interact with anyone (family member, friend, classmate) from Pittsburgh?

Y N

14. How much time have you spent in Pittsburgh? (choose one)

None

Less than a day

Less than a week

Less than a month

Less than a year

Less than 10 years

More than 10 years

15. If yes, how old were you the first time you went? _____

Appendix C

Self-paced reading materials and comprehension questions for Studies 2A and 2B. The locations of the line breaks when the stimuli were displayed to the participants are indicated with a / . (We did not have any particular constraints on the location of the line breaks in the filler items.) As described in Footnote 6, a handful of comprehension questions were not used in determining which trials to include; these questions are denoted with a *.

Critical *be*-drop items

1. Yes, I have everything for Steve's going-away party. / A dessert will baked by Monica this afternoon. Please / remind me if I forget to pick it up.

Q1. Is it a welcoming party? N

Q2. Is there a dessert? Y

2. The current draft of the press release is attached, but / please note that this version is not for distribution. / The draft will approved by corporate before it goes out / to customers. This is for internal use only.

Q1. Can the draft be distributed? N

Q2. Is this a draft of a television commercial? N

3. I hope everyone saw that 10th Avenue is shutting down / between Washington and Jefferson the rest of this week. / The neighborhood will paved soon and the city is / starting there. Green Street should be the best / alternate route during this time.

Q1. Are the streets being paved? Y

Q2. Is the best alternate route Green Street? Y

4. We're not completely ready yet for today's big speech. / The presentation file will downloaded onto my Mac in / another few minutes. Then I'll be set.

Q1. Have all the files been downloaded yet? N

Q2. Are they completely ready for the speech? N

5. I'll probably be working late tonight. Whenever you get / home, would you go ahead and get dinner started? If the / pork roast will cooked for 90 minutes, I'm sure I'll be / home before it's ready.

Q1. Is this person working late? Y

Q2. Are they cooking a roast? Y

6. Are the renovations going to include the first floor as / well? I know the warehouse will remodeled soon due to / the increasing inventory demands, but what about the / front area? Let me know who I ought to be making my / case to.

Q1. Are the inventory demands increasing? Y

Q2. Is something being renovated? Y

7. I agree; we've lost too many repeat clients. But, I / think our retention rate will improved by the changes / Scott is making in customer service. This has to be a / high priority.

Q1. Has the company lost some clients? Y

Q2. Is this a high-priority issue? Y

8. I checked with Jessica about the fourth floor copier. / It just will recycled because it can't be fixed. / We'll have to purchase a new one.

Q1. Does the copier work? N

Q2. Is the copier on the tenth floor? N

9. Can you get more info from Dan on the website issue he / reported? I understand that other people have not seen / it. In the mean time, the report will verified by other / members of the dev team. Thank you for your attention.

Q1. Is the problem with the website? Y

*Q2. Have other people seen the problem? N

10. The sink was filled with dirty mugs again this morning. / Right now, the entire kitchen will washed only weekly. / If you use a mug in the mean time, please rinse it out / and return it to the dishwasher.

Q1. Was the sink clean this morning? N

*Q2. Do the mugs belong in the cabinet? N

11. I like the new print ads' design, but the text is too / long. The text will trimmed quite a bit by the editorial / staff. I'll let you know when the next draft is / available.

*Q1. Does this person like the layout? Y

Q2. Should the advertisement be longer? N

12. Sales to other businesses are clearly our weakness. / Even while our other sales are up, those are down. If / that business will raised soon, we'll be looking great. / Otherwise, I'm not sure we can justify our presence in / that sector.

Q1. Have sales to businesses increased? N

Q2. Have other sales increased? Y

13. Why haven't you finished getting Order 437 ready? The / shipment itself will boxed by the warehouse staff, but / you have to handle the paperwork. The delivery truck / goes out this afternoon, so I need you to do this ASAP.

Q1. Is the delivery truck going out today? Y

Q2. Is Order 437 ready? N

14. I got the manuscripts you forwarded to me. It looks / like we're ready to go with the magazine article. The / science book will edited by our staff because there are / still some typos.

Let me know if you approve.

*Q1. Is the magazine article unfinished? N

Q2. Does the book have typos? Y

15. I'm certainly excited to have a new award to display. / The display case will polished this afternoon. Let's / put the trophy at the top of it.

Q1. Will the display case be polished this afternoon? Y

Q2. Is there a new award? Y

16. Thanks for organizing the end-of-year party. I / understand that most of the prep work is done. The / whole kitchen will cleaned by volunteers after the / party. You don't have to worry about that.

Q1. Is the preparation for the party mostly done? Y

Q2. Is the party celebrating the middle of the year? N

17. Heard back from the consultant. The biggest issue she / noted is that our storefronts don't grab the eye. The / store entrance will decorated more brightly in the / future to capture shoppers' attention. I think it can / fix the problem.

Q1. Did the storefronts grab shoppers' attention? N

Q2. Did they think the problem can be fixed? Y

18. I don't want to focus on new software features until / we've finished this year's worldwide release. The most / recent version will translated by the end of March. / Supporting that translation work is our top priority.

Q1. Is the worldwide release finished? N

Q2. Will translation be a priority? Y

19. Hi, I'm doing a report on our Web sales, but I only / have the first quarter numbers, which are

out of date. / The report will updated with the latest data once I / receive them. That might take a few days.

Q1. Is the report about Web sales? Y

Q2. Does the writer have the latest sales numbers? N

20. Anyone else keep hearing that clattering sound? I think / the vent in the corner is getting a little loose. The / metal plate will tightened by Facility Services. I hope / that fixes the problem.

Q1. Is it a fan that's loose? N

Q2. Is the vent too tight? N

Critical Dialectal *Need* items

1. Yes, I have everything for Steve's going-away party. / The dessert needs baked towards the end of the / afternoon, though. Please remind me if I forget.

Q1. Is it a welcoming party? N

Q2. Is there a dessert? Y

2. The current draft of the press release is attached, but / please note that this version is not for distribution. / The draft needs approved before it goes out to / customers. This is for internal use only.

Q1. Can the draft be distributed? N

Q2. Is this a draft of a television commercial? N

3. I hope everyone saw the news that 10th Avenue is closing / between Washington and Jefferson the rest of this week. / The neighborhood needs paved soon and the city is / starting there. The best alternate route during this / time is Green Street.

Q1. Are the streets being paved? Y

Q2. Is the best alternate route Green Street? Y

4. We're not completely ready yet for today's big speech. / Your file still needs downloaded to my Mac first. / Otherwise, I'm set.

Q1. Have all the files been downloaded yet? N

Q2. Are they completely ready for the speech? N

5. I'll probably be working late tonight. Whenever you get / home, would you go ahead and get dinner started? The / pork roast needs cooked for about 90 minutes, so / getting a head start would be helpful. Thanks and love / you.

Q1. Is this person working late? Y

Q2. Are they cooking a roast? Y

6. Are the renovations going to include the first floor as / well? The warehouse needs remodeled if we want to / remain on top of the increasing inventory demands. Please let / me know who I ought to be making my case to.

Q1. Are the inventory demands increasing? Y

Q2. Is something being renovated? Y

7. I agree; we've lost too many repeat clients. Our / customer service needs improved to boost up our / retention rate. This has to be a high priority.

Q1. Has the company lost some clients? Y

Q2. Is this a high-priority issue? Y

8. I checked with Jessica about the fourth floor copier. / It just needs recycled because it can't be fixed. We'll / have to purchase a new one.

Q1. Does the copier work? N

Q2. Is the copier on the tenth floor? N

9. Can you get more info from Dan on the website issue he / reported? I understand that other

people have not seen / it. The report needs verified before I take this issue / any further. Thank you for your prompt attention.

Q1. Is the problem with the website? Y

*Q2. Have other people seen the problem? N

10. The sink was filled with dirty mugs again this morning. / The kitchen needs washed daily at the close of / business, and I'm not always going to be the person who / does it. If you use a mug, please rinse it out and / return it to the dishwasher.

Q1. Was the sink clean this morning? N

*Q2. Do the mugs belong in the cabinet? N

11. I like the new print ads' design, but the text is too / long. The text needs trimmed quite a bit before the ad / is ready. Let me know when the next draft becomes / available.

*Q1. Does this person like the layout? Y

Q2. Should the advertisement be longer? N

12. Sales to other businesses are clearly our weakness. / Even while our other sales are up, those are down. That / business needs raised if we want to justify our / presence in that sector. What does Michael think?

Q1. Have sales to businesses increased? N

Q2. Have other sales increased? Y

13. Why haven't you finished getting Order 437 ready? The / complete shipment needs boxed before our delivery truck / goes out this afternoon. I need you to do that ASAP.

Q1. Is the delivery truck going out today? Y

Q2. Is Order 437 ready? N

14. I got the manuscripts you forwarded to me. It looks / like we're ready to go with the

magazine article. The / science book needs edited because there are still some / typos. Let me know when it's done.

*Q1. Is the magazine article unfinished? N

Q2. Does the book have typos? Y

15. I'm certainly excited to have a new award to display. / The display case needs polished before we add the / trophy. Let's do that this afternoon.

Q1. Are they polishing the display case this afternoon? Y

Q2. Is there a new award? Y

16. Thanks for organizing the end-of-year party. I / understand that most of the prep work is done. The / whole kitchen needs cleaned after the party as well. / Can we get a volunteer for that, too?

Q1. Is the preparation for the party mostly done? Y

Q2. Is the party celebrating the middle of the year? N

17. Heard back from the consultant. The biggest issue she / noted is that our storefronts don't grab the eye. The / store entrance needs decorated if we want to capture / shoppers' attention. I think we can fix that.

Q1. Did the storefronts grab shoppers' attention? N

Q2. Did they think the problem can be fixed? Y

18. I don't want to focus on new software features until / we've finished this year's worldwide release. The most / recent version needs translated before we can add / anything new. Let's make that a priority.

Q1. Is the worldwide release finished? N

Q2. Are they making translation a priority? Y

19. Hi, I'm doing a report on our Web sales, but I only / have the first quarter numbers, which are out of date. / The report needs updated with the latest data. Can you / help me with that?

Q1. Is the report about Web sales? Y

Q2. Does the writer have the latest sales numbers? N

20. Anyone else keep hearing that clattering sound? I think / the vent in the corner is getting a little loose. The / metal plate needs tightened if that's the case. It's / getting pretty annoying.

Q1. Is it a fan that's loose? N

Q2. Is the vent too tight? N

Filler items

1. Sorry, it's gonna take me a while to get to your maintenance request. Swamped with e-mails right now and can't start on any new issues. Maybe tomorrow.

Q1. Does this person have a lot of e-mails? Y

Q2. Is the maintenance request going to be completed today? N

2. Had a great time in Hawaii, but happy to be home now. We got stuck in O'Hare for 5 hours. Horrible way to spend an evening!

Q1. Did this person go to California? N

Q2. Was the person stuck in O'Hare for five hours? Y

3. Not going to be able to make it into work today due to the snow. I'll be available on email and phone from home. See you tomorrow.

Q1. Did rain keep this person from getting into work? N

Q2. Is this person still available over email? Y

4. I received your request about replacing the microwave in the break room. I'm afraid that's not in the current budget. The old one still works, so we'll have to stick with it.

Q1. Does the old microwave work? Y

Q2. Is the microwave being replaced? N

5. It's that time of year again! My daughter is selling Girl Scout cookies, so here's your chance to get your Thin Mints. I'll put the order sheet in the break room.

Q1. Is the daughter selling magazine subscriptions? N

Q2. Is the order sheet in the break room? Y

6. As a reminder, the fax machine is for business use only. Personal accounts have been discontinued. Please direct any questions to Margie.

Q1. Can the fax machine be used for personal use? N

Q2. Should questions be directed to Dave? N

7. Tomorrow's staff meeting has been moved from 9 to 11 AM. Sarah has a dentist appointment early in the morning. Apologies for the late notice.

Q1. Has the meeting time been changed? Y

*Q2. Does Sarah have a doctor's appointment? N

8. Are you going to be home this afternoon? The delivery people just called me here at work. They can deliver the new sofa if you'll be around to let them in.

Q1. Is a refrigerator going to be delivered? N

*Q2. Is the delivery scheduled for the morning? N

9. This is the last call for feedback on the candidate we interviewed today for accounting. If you have comments, please get them to me by the close of business today. Thanks for your input.

Q1. Was there a job interview? Y

Q2. Are the comments due today? Y

10. Has anyone else been getting Java errors on their computer? It's been happening to me since

this morning. What can I do about this?

Q1. Is this person having computer problems? Y

Q2. Did the problems start this morning? Y

11. The network changes I've discussed in my previous e-mails are effective tomorrow. Restarting your computer should be enough for the changes to take effect. If you have any further troubles, please contact me.

Q1. Must the computers be restarted? Y

Q2. Do the changes take effect next month? N

12. A new Windows update was released this morning. This update addresses security vulnerabilities. When your computer prompts you to perform the update, please do so at your earliest opportunity.

Q1. Should the update be done as soon as possible? Y

Q2. Is the update for Windows? Y

13. I've got two tickets for tonight's hockey game that I won't be able to use. $200 for the pair. First come, first serve.

Q1. Are there four tickets? N

Q2. Are the tickets for a basketball game? N

14. I'm pleased to share the final version of our new TV commercial. The link is below. The ad is going to begin airing next week. Just wanted to pass this along to the whole team.

Q1. Is the ad still being worked on? N

Q2. Is the ad going to begin airing today? N

15. Missed you when I was handing out paychecks this morning. You can pick yours up from my office after lunch. See you soon.

Q1. Did this person get the paycheck? N

Q2. Should the paycheck be picked up before lunch? N

16. The silver Honda Civic in the parking lot has its lights on. Anyone know whose car that is? You might want to turn them off if that's yours.

Q1. Is the car silver? Y

Q2. Is the car a Civic? Y

17. As you may have heard, heavy snow is forecast beginning this afternoon. The office is going to close at 1 PM to allow everyone to travel home safely. If you have work that can be continued at home, please plan to do so. Safe travels.

Q1. Is the office closing because of rain? N

Q2. Is the office closing at 1? Y

18. Leftovers from today's sales team meeting are in the kitchen. Pizza, pop, cookies, and more. Hurry before they're all gone!

Q1. Are the leftovers from a sales meeting? Y

Q2. Do the leftovers include pizza? Y

19. Our windows are being washed today, starting on the north side of the building. We don't expect the process to cause any major disruptions. If you do anticipate any issues, please let me know ASAP.

Q1. Are the windows being washed? Y

Q2. Are major disruptions expected? N

20. I'm missing my keys, and I know I had them with me this morning. I'm guessing they're somewhere around the office. Has anyone seen them? If you could keep an eye out for them, I'd appreciate it.

Q1. Is the person missing a wallet? N

Q2. Did the person have their keys this morning? Y

Appendix D

Post-experiment questionnaire for Studies 2A and 2B.

Please answer the following questions about the experiment you just completed.

1. Did you think there was anything odd about the sentences you read?

Certain NO Probably NO Probably YES Certain YES

2. Did you think that some of the sentences had unusual grammar to them?

Certain NO Probably NO Probably YES Certain YES

3. Did you see any sentences with structures like *The car needs washed*?

Certain NO Probably NO Probably YES Certain YES

4. Did you see any sentences with structures like *I might could do my homework*?

Certain NO Probably NO Probably YES Certain YES

5. Did you see any sentences with structures like *Gas is so expensive anymore*?

Certain NO Probably NO Probably YES Certain YES

6. Did you see any sentences with structures like *I bought me a house*?

Certain NO Probably NO Probably YES Certain YES

7. Outside of this experiment, have you heard people say things like *The car needs washed*?

Certain NO Probably NO Probably YES Certain YES

8. Outside of this experiment, have you heard people say things like *I might could do my homework*?

Certain NO Probably NO Probably YES Certain YES

For the next two questions, please think about what you would say in a conversation with a friend. We are interested with what YOU would actually SAY, no matter what you think is "correct" or what you "should say."

9. Would you use a sentence like *The car needs washed* in a conversation with a friend?

Certain NO Probably NO Probably YES Certain YES

10. Would you use a sentence like *I might could do my homework* in a conversation with a friend?

Certain NO Probably NO Probably YES Certain YES

11-16. For each of the following cities, we'd like your judgment on whether you think a person in the city would say a sentence like *The car needs washed*. For these questions, we just mean the AVERAGE person in the city, not anyone you know in particular.

The car needs washed:

Boston	Would say	Wouldn't say
Denver	Would say	Wouldn't say
Minneapolis	Would say	Wouldn't say
Nashville	Would say	Wouldn't say
Pittsburgh	Would say	Wouldn't say
Seattle	Would say	Wouldn't say

17-22. For each of the following cities, we'd like your judgment on whether you think a person in the city would say a sentence like *I might could do my homework*. For these questions, we just mean the AVERAGE person in the city, not anyone you know in particular.

I might could do my homework:

Boston	Would say	Wouldn't say
Denver	Would say	Wouldn't say
Minneapolis	Would say	Wouldn't say
Nashville	Would say	Wouldn't say

Pittsburgh	Would say	Wouldn't say
Seattle	Would say	Wouldn't say

Please answer the following questions to help us learn more about your language background.

23. What do you consider your native language? _____

24. What is the first language you ever learned? _____

24b. If you learned two first languages simultaneously, what was the other one?

If you did not learn two first languages, please enter: none

_____

25. Where (city, state, country) did you learn English? _____

26. Where (city, state, country) do you live now? _____

Appendix E

As detailed in Footnote 2, Study 1 included a manipulation of the putative origin of the stimulus materials (a company in Denver versus a company in Pittsburgh) that had no effect on participants' reading times, neither among participants already familiar with Dialectal *Need* nor among participants previously unfamiliar with it. One possible reason why this variable was irrelevant is that comprehenders do not have much knowledge about the specific geographic distribution (even if they may have encountered the structure itself).

To test this possibility, we used the post-experiment questionnaire in Study 2 to examine participants' naïve beliefs about the distribution of the Dialectal *Need* feature. We asked participants to report whether an individual in each of six different United States cities was likely or unlikely to say a sentence such as *The car needs washed*. The results of these questions are displayed in Table E1 separately for each level of self-reported participant prior familiarity with Dialectal *Need*.

Table E1
Percentage of Participants Who Judged an Average Person in Particular Cities Would Use Dialectal Need *as a Function of Participant Familiarity with Dialectal* Need.

	Self-Reported Prior Exposure to Dialectal *Need*			
City	Certainly No	Probably No	Probably Yes	Certainly Yes
Boston	12%	56%	54%	80%
Denver	0%	18%	35%	60%
Minneapolis	6%	43%	43%	83%
Nashville	24%	59%	47%	78%
Pittsburgh	2%	32%	41%	92%
Seattle	4%	7%	27%	53%

These questions indicate that participants generally did not have strong expectations about the geographic distribution of Dialectal *Need*. Except among participants who were certain that they had heard Dialectal *Need* before, participants guessed that participants would be more likely to use Dialectal *Need* in Nashville or Minneapolis than in Pittsburgh, which is

inconsistent with the sociolinguistic evidence (Doyle, 2014; Murray et al. 1996). Conversely, the excluded participants who were certain they had heard Dialectal *Need* were (correctly) most apt to identify Pittsburgh as a city in which it would be used, but they also (incorrectly) predicted a high probability of usage in other cities (Boston and Minneapolis) in which Dialectal *Need* is not prevalent (Doyle, 2014; Murray et al., 1996). In sum, it seems that participants—even those who are previously familiar with the Dialectal *Need* structure as a whole—do not have accurate expectations about where it is most likely to be used. That is, knowledge of the existence of dialectal structures does not imply knowledge of their geographic distribution. This pattern likely explains why the cover story manipulation in Study 1 did not have any effect: If participants do not know where Dialectal *Need* is used, the putative origin of the stimulus materials is irrelevant to them.