Processing grammatical differences: Perceiving versus noticing

Lauren Squires

The Ohio State University

squires.41@osu.edu

Introduction

This chapter explores the relationship between English speakers' processing and awareness of morphosyntactic variability. Much sociolinguistic investigation has focused on speakers' explicit, metalinguistic knowledge of what linguistic differences exist and what their social meanings are. This chapter shifts the focus to how language variation is experienced in the moment in which it is encountered. I investigate this using an on-line experimental method to test participants' processing of grammatically variant sentences. The experiments test for both the *perceiving* of grammatical differences and the *noticing* of those differences, which I argue provide the foundation for the *knowledge* of variation and its social meaning.

Theoretical and methodological background

Sociolinguists have long been interested in the relative degree to which the production of sociolinguistic variation is conscious or unconscious. One often-used categorization is Labov's (1972) distinction between sociolinguistic *indicators*, *markers*, and *stereotypes*. Indicators and stereotypes represent ends of a continuum of sociolinguistic awareness: indicators are linguistic features that are correlated with social properties, but they are not deployed for stylistic purposes and speakers are not aware of their social correlation; stereotypes are features subject to highly conscious discussion and social evaluation. Markers are in between, being features that both

correlate with social properties and are deployed stylistically, yet speakers nonetheless may be unaware of their own stylistic uses of them (Johnstone & Kiesling, 2008). Labov's categorization focuses on the relation between sociolinguistic variation, social indexicality, and sociolinguistic awareness.

The levels of awareness articulated by the indicator/marker/stereotype categories are alternatively expressible as levels of *knowledge* held by speakers about the variation in question. I want to suggest that what *awareness* in these categories represents is *implicit v. explicit knowledge* of variation. With stereotypes, speakers *know* that the feature relates to a specific category of speaker; with indicators, speakers *do not know* that the feature relates to a specific category of speaker. Speakers have explicit knowledge of stereotypes that they can articulate and discuss (Silverstein, 1981; Preston, 1996), but speakers must also have some knowledge of indicators and markers to use them variably as part of their grammatical competence. But knowledge of indicators is *implicit*, not consciously articulable.

As used within most sociolinguistic research (as in, e.g., Silverstein, 1981; Preston, 1996, 2011; Mertz & Yovel, 2003), *awareness* seems to be a matter of the raising of internal *knowledge* to the surface of a speaker's consciousness, with a continuum of *awareness* representing a continuum from knowledge that is implicit to explicit. The construct of knowledge is in overt focus in work by Labov (1973) and Wolfram (1982), both of whom assess the degree to which speakers are aware of—as in, the extent to which they *know*—the patterns and constraints governing dialects other than those they have productive competence in. Awareness-as-knowledge is also foregrounded in the title of Preston's (1996) now-classic piece on sociolinguistic awareness, and his contribution to this volume, which ask "Whaddaya*know*?" (emphasis mine).

In contrast, we can think of a different sort of *awareness*—one that is centrally connected to subjective language experience, to in-the-moment language processing and production. One of the premises underlying this chapter is that knowledge of variation likely comes about from being exposed to linguistic differences, noticing them, and coming to understand patterns of their use in connection to social facts. While much research has investigated the state of speakers' knowledge about facts of variation (e.g., Johnstone & Kiesling, 2008; Campbell-Kibler, 2009; Staum Casasanto, 2009; Squires, 2013a), the process of speakers coming into that knowledge has gone relatively unexplored, particularly from a processing perspective (one recent exception is Docherty et al., 2013). While some beliefs about language variation may be accrued via metalinguistic stereotypes without direct experience (Carmichael, this volume; McGowan, this volume), it seems that implicit knowledge in the sense of "competence" is most likely structured from experience (see, e.g., Bresnan & Ford, 2010). Explicit knowledge ("awareness") should then emerge from aggregated experiences of in-the-moment awareness of linguistic differences—that is, the *noticing* of variation, and coming to understand it as linguistically and socially meaningful.

Differences within and among constructs like *knowledge, awareness,* and *noticing* have been discussed at length in the field of second language acquisition (see also Nycz, this volume). Schmidt (1990) addresses the role of *consciousness* in second language learning, first determining three ways that scholars have viewed *consciousness*: as awareness, as intention, and as knowledge. Further, three levels of awareness are distinguished: *perception, noticing,* and *understanding*. Things are frequently perceived without being noticed, Schmidt suggests, and are frequently noticed without being understood. Schmidt says,

> When reading, for example, we are normally aware of (notice) the content of what we are reading, rather than the syntactic peculiarities of the writer's style, the style of type in which the text is set, music playing on a radio in the next room, or background noise outside a window. However, we still perceive these competing stimuli and may pay attention to them if we choose. (Schmidt, 1990:132)

Further, Schmidt contends that noticing is requisite to understanding; there is no "subliminal" learning (though there may be subliminal perception).

It is natural to analogize other-dialect awareness to other-language awareness. If we do so, and take Schmidt's definitions seriously, then the question "How aware are speakers of sociolinguistic differences?" breaks down into three separate but related questions: How much do speakers *perceive* sociolinguistic difference? How much do speakers *notice* sociolinguistic difference? and How much do speakers *understand* sociolinguistic difference?

I want to think about these concepts as important to disentangle specifically for the burgeoning research field of "sociolinguistic perception," which investigates the relation between variation and perception (see Campbell-Kibler, 2010, this volume; also Drager & Kirtley, this volume). There is a drive within sociolinguistics to more robustly understand the connection between linguistic processing and the social meaning of linguistic forms, for instance in Preston's (2011; this volume) detail of how "language regard" might influence comprehension, or in Labov et al.'s (2011) development of a "sociolinguistic monitor" that tracks and stores frequencies of linguistic variables. Sociolinguistic perception research holds the promise of using rigorous experimentation to explore the cognitive structuring of knowledge about variation. It is important to consider the role of awareness in producing that knowledge (as many of the chapters in this volume do, notably those by Beck, Campbell-Kibler, Drager & Kirtley, McGowan, and Preston).

The present study represents an exploration into awareness by considering what happens *during the processing of language variation*. How do formal differences affect language comprehension? When are linguistic forms noticed as different? Are more-difficult-to-process forms foremost in speakers' awareness of difference? In particular, this study examines the relation between *perceiving* difference and consciously *noticing* it, positing that these processes are foundational to speakers developing *understanding* and *knowledge* of sociolinguistic variation through exposure to it.

Sociolinguistic processing *in-the-moment* is a relatively unexplored area of research (though see Loudermilk, 2013), and sociolinguistic perception research has in general focused on phonological variation and phonetic variables. The present chapter extends the inquiry into the on-line (that is, real-time) processing of grammatical variation, attempting to measure a) whether speakers perceive (morpho)syntactic difference, and b) whether they notice it. Psycholinguistic methods are well developed for measuring sentence processing (see more complete discussions in Squires, 2013a, 2013b, 2014), and the method used here is self-paced reading, wherein participants move through a sentence unit-by-unit at their own pace. Reading speed is taken to index language processing, with more-difficult linguistic units taking longer to read than less-difficult units (Just et al., 1982). What is "difficult" is typically that which is unexpected, due to a violation of grammatical constraints, low probability, or both. This method has been used to show that speakers' processing is sensitive to, for instance, agreement mismatches (Pearlmutter et al., 1999; Breadmore et al., 2013), semantic anomalies (De Vicenzi et al., 2003), and probabilistic facts about the occurrence of structural alternatives (Bresnan & Ford, 2010).

Relevant to the questions of this chapter, two prior studies have productively used self-paced reading to study the processing of regional dialect variants. Kaschak and Glenberg (2004)

studied adults' "acquisition" of a novel dialect form: the [*need*+past participle] construction (e.g., *The dishes need washed*) common in the northern Midlands dialect area of the U.S. Participants, who were not speakers of the dialect, read sentences with the *need*+past participle construction more slowly than those with the standard construction (*The dishes need to be washed*). However, this effect was attenuated with multiple exposures to the pattern, and further research showed that participants were also able to generalize the construction to other structures (Kaschak, 2006).

Kaschak and Glenberg's research shows that speakers are sensitive to sentence structures that are not part of their own dialect: participants perceived the "oddity" of the [*need*+past participle] construction. However, participants also became less sensitive to the construction the more they encountered it, and possibly even learned its meaning and grammatical patterning. In Schmidt's (1990) terms, they came to *understand* the form. Because Kaschak & Glenberg do not report what their subjects thought about the sentences, we cannot assess whether they metalinguistically *noticed* the construction. Yet if Schmidt (1990) is correct, and understanding presupposes noticing, then participants should have also noticed the differences between sentences. What did participants come to understand [*need*+past participle] *as?* Did they (accurately) categorize it as a dialect form, or did they simply categorize it as an "error" and assume that the experiment involved making errorful sentences? Knowing what the participants *noticed*, and how they metalinguistically categorized what they noticed, might shed further light on the levels of awareness during the processing and acquisition of new grammatical forms.

In a study explicitly using self-paced reading to measure awareness during reading comprehension, Breadmore et al. (2013) consider self-paced reading times as a measure of "implicit awareness" of subject-verb agreement mismatches. To measure "explicit awareness,"

they used a post-experiment error correction task of the same subject-verb agreement mismatches. Comparing deaf and hearing children's performance, Breadmore et al. show that explicit and implicit awareness are not always aligned: deaf children did poorly at the error correction task, showing a lack of explicit awareness of agreement errors. But the deaf children's reading times *were* affected by the agreement errors, though this effect did not show up until one word later than for the hearing children.

Breadmore et al.'s results show that what is perceived does not always rise to consciousness, just as Schmidt (1990) suggests. Similar results have been found in recent neurolinguistic research, which has found that the brain may detect syntactic errors even when a listener does not consciously register them (Batterink & Neville, 2013). On the other hand, Hanulíková et al. (2012) show that brain responses to grammatical errors can be modulated by sociolinguistic perceptions of the speaker: knowing that a person speaking errorfully is a nonnative speaker makes the brain respond less strongly to those errors. Thus, perceiving linguistic differences might not always lead to noticing, and explicit awareness of differences—or the likelihood of them—may affect what is perceived.

The links between low-level automatic perception, conscious noticing, sociolinguistic differences, and social information are ripe for further investigation. The present chapter uses both an on-line behavioral measure and an off-line metalinguistic task to explore the link between perception and noticing, providing a basis from which to further explore the role of sociolinguistic processing in sociolinguistic knowledge and its acquisition.

Experiments overview

This chapter presents the results of a series of experiments testing participants' processing of subject-verb agreement variation (Pilot Experiment, Experiment 1, Experiment 2). The experiments tested adult English speakers' reading times in sentences containing [NP+*don't/doesn't*], where the combination of number on the subject noun and auxiliary verb form was variable. Some sentences contained standard agreement, some contained nonstandard agreement, and some contained what I will call "uncommon" agreement, explained below. Participants' word-by-word reading times were measured to assess the extent to which they automatically *perceived* the variant forms, and their off-line metalinguistic response to a post-experiment question was used to assess whether they had *noticed* the variant forms.

The agreement conditions are presented in (1a-d). The "nonstandard" variant, [SG+*don't*], is common across varieties of English ("Feature 171"; often called "invariant *don't*"). In the United States it is associated both descriptively and perceptually with lower social status (Squires, 2013a; Squires, 2014). In contrast, the "uncommon" form, [PL+*doesn't*], is not known to be a dialect variant. It was included in the sentences in order to provide a point of comparison between a nonstandard syntactic structure that participants would likely have encountered but probably do not use themselves, versus a structure that participants would likely not have encountered or use. This method enables probing the role of sociolinguistic experience in the perceiving and noticing of linguistic difference (see also chapters by Beck, Carmichael, Drager & Kirtley, and McGowan).

(1a) Standard (plural):	After eating, the	turtles	don't	walk	very	fast.
(1b) Standard (singular):	After eating, the	turtle	doesn't	walk	very	fast.
(1c) Nonstandard:	After eating, the	turtle	don't	walk	very	fast.
(1d) Uncommon:	After eating, the	turtles	doesn't	walk	very	fast.
		1	2	3	4	
		noun	don't	verb	verb+1	

In each experiment, participants read sentences one word at a time through a "moving window" self-paced reading task (Just et al., 1982). After reading each word, they pressed a button on an experimental response box to continue to the next word. The dependent measure was word reading time, from the appearance of a word on screen to the participants' button-press to advance to the next word.

There were four critical word regions within each sentence, labeled in (1): [1] the subject noun before *don't/doesn't* (noun region); [2] *don't/doesn't* (*don't* region); [3] the main verb following *don't/doesn't* (verb region); and [4] the word following the main verb (verb+1 region). Self-paced reading studies typically identify the strongest effects at the word *following* the introduction of a grammatical anomaly, with processing effects occasionally persisting even beyond (Just et al., 1982; Pearlmutter et al., 1999; Kaschak & Glenberg, 2004, 2006; Breadmore et al., 2013). Hence, the region expected to show the most consistent effect of agreement is the verb region [3]. The noun region [1] is shown in the figures below for purposes of comparison, but it is not included in the statistical analyses.

As a measure of *perceiving*, I hypothesized that participants' reading times would be fastest in the standard condition, slowest in the uncommon condition, and intermediate in the nonstandard condition. Standard agreement would be the most expected in this setting (a university lab) and for this modality (written), and it was also likely the most-used form of agreement for the participants (who were university students). In contrast to standard agreement, nonstandard agreement should also be unexpected. Yet because it is a common (and even stereotyped) dialect variant in the US, participants should have some knowledge of the nonstandard form, and therefore its effect should be less disruptive than that of uncommon agreement.

In addition to measuring reading times, I sought to group participants in each experiment by whether they consciously *noticed* the agreement differences or not, and to investigate whether differences in noticing corresponded to differences in perceiving. Schmidt (1990:132) suggests that "noticing can be operationally defined as availability for verbal report, subject to certain conditions." I used a post-experiment questionnaire question to ask participants to report if they "noticed anything interesting about the grammar of the sentences." Participant groupings were assigned based on the following criteria: if participants mentioned either *don't*, *doesn't*, subject-verb agreement, or expressions indexing any of these specifically, I considered them to be "aware" participants. If they did not mention any of these features, they were "unaware" participants. Though this method necessarily relies on participants' metalinguistic articulation, my coding criteria for being counted as "aware" attempted to minimize the importance of terminology. To summarize the experimental method: I consider participants' perception using a quantitative measure (reading time), and their awareness using a qualitative measure (metalinguistic report); I then investigate whether awareness is a predictor of perceptual behavior (similar to Nycz, this volume).

The three experiments differed in one crucial property: the ratio of nonstandard/uncommon sentences to standard sentences was decreased in each subsequent experiment. The Pilot Experiment found that the majority of participants were "aware" (discussed below). Subsequently, I sought a more balanced grouping of unaware and aware participants and hypothesized that fewer tokens of the grammatical variants would lead to fewer people noticing them. I therefore reduced the number of nonstandard/uncommon sentences in Experiments 1 and 2 to attempt to increase the number of unaware participants.

Pilot Experiment

The Pilot Experiment did not have investigating participant awareness as a main goal; rather, it was conducted simply to assess the processing of syntactic variation as part of a larger project, and it served as a pilot use of the "noticing" question in the post-experiment questionnaire. Full discussion of the methods and general results of this experiment appears in Squires (2014; Experiment 1).

Participants were exposed to 64 target sentences, which contained [NP+*don't/doesn't*], and 64 filler sentences, which were grammatically standard and did not contain *don't* or *doesn't*. After eight of the filler sentences, participants answered a yes/no comprehension question about the prior sentence, to ensure they were paying attention. The total proportion of non-standard sentences throughout the experiment was 25 percent (32 out of 128). Sentences were divided into four blocks; participants read 4 nonstandard and 4 uncommon sentences, along with 8 standard target sentences and 16 filler sentences, in each block. 45 participants received course credit for completing the experiment; the data of two participants were removed due to experiment error.

The majority of participants—33 out of 43—were classified as "aware" based on their post-experiment questionnaire responses. Figure 1 shows the mean response times (in milliseconds) across agreement conditions by both aware and unaware participants. As expected, participants' reading was slowed by both nonstandard and uncommon agreement relative to standard agreement, with the largest divergence being for uncommon agreement. Additionally (and unexpectedly), the aware participants were faster at reading overall than the unaware participants. Because of this difference, and because the relation between behavior and awareness is the major question here, for this chapter I analyze and report raw reading times rather than residual reading times (in Squires, 2014, I report the residual reading times). For all

experiments, observations with reaction times under 30 ms and over 2000 ms were removed as outliers prior to analysis.

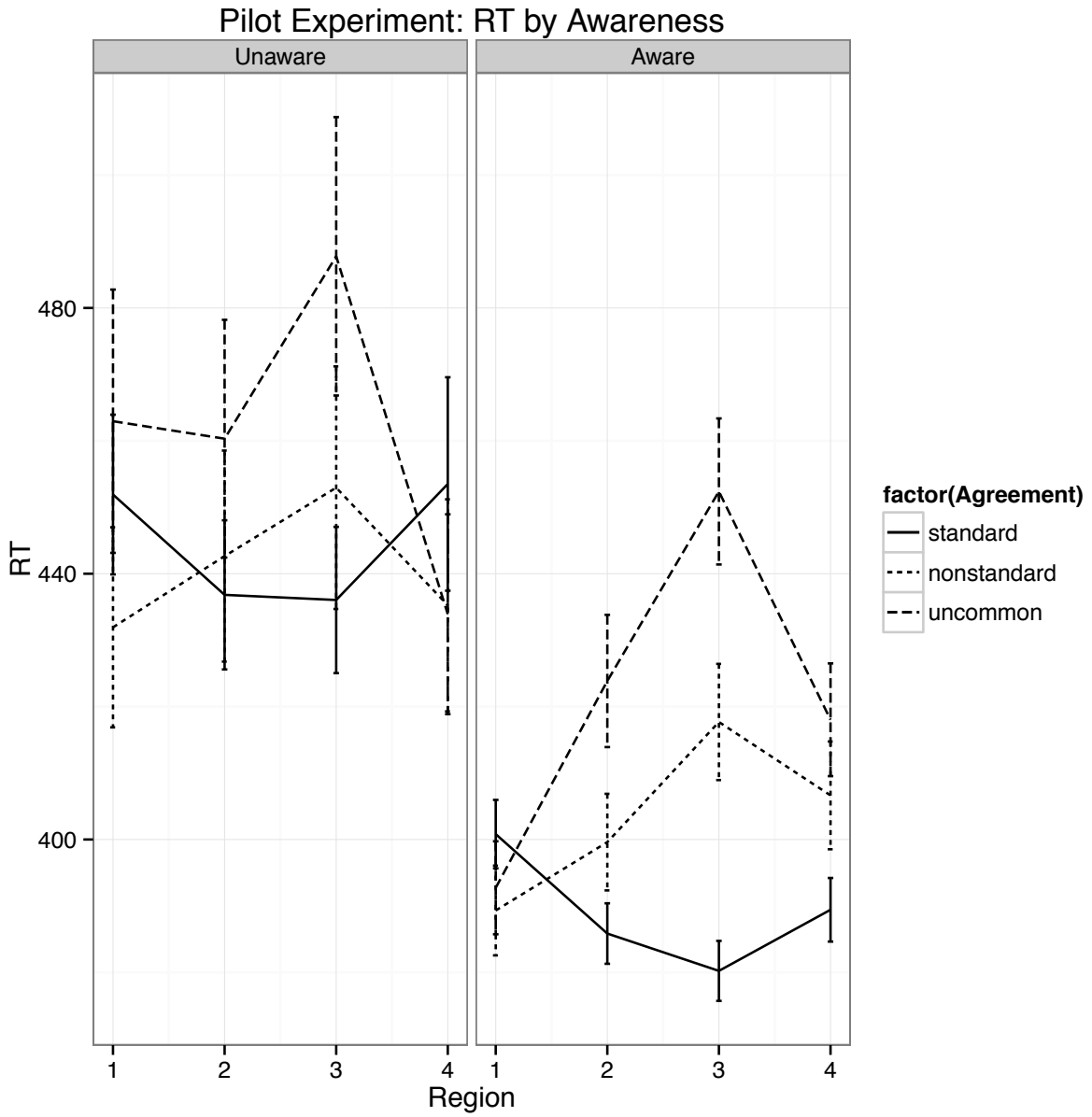

Figure 1. Pilot Experiment reaction times in milliseconds by agreement and participant awareness, across word regions (1 = noun, 2 = don't, 3 = verb, 4 = verb+1). Error bars represent the standard error of the mean.

To assess both agreement and awareness as predictors of reading time, I conducted statistical analysis using mixed-effects linear regression modeling with the {lmer4} package in R (Bates, Maechler & Bolker, 2011), estimating p-values with the {languageR} package (Baayen, 2010). Reading times for each word region were analyzed independently. Every model began with two terms: experimental block as a fixed effect, since reading times decreased over the course of the experiment, and a random intercept effect for experimental item, to account for variability across sentences/words. (I did not include random effects for subjects in order to better detect an effect of awareness, which is a subject-level variable.) I then followed a model comparison procedure in which I sequentially tested the effect of agreement condition, participant awareness, and the interaction of agreement and awareness. I retained a predictor when analysis of variance comparing the model with and without it showed a significant chi-square value at the $p<.05$ level. For all three experiments, the models I present are the best-fitting models according to this procedure.

Table 1 summarizes the regression models for the Pilot Experiment. The parameter estimates are interpretable as predicted increases or decreases in reading time for the listed factor levels, relative to the baseline factor levels of standard (for agreement) and unaware (for awareness).

Awareness was significant as a main effect at all three word regions, with aware participants reading faster than unaware participants. Agreement showed different effects across the regions. At *don't*, uncommon sentences took longer to read than standard sentences. At the verb, both nonstandard and uncommon sentences took longer than standard. At verb+1, there was an interaction between agreement and awareness: while there was not a main effect of

agreement, for the aware participants, uncommon agreement continued to slow them down (as shown in Figure 1). Aware participants seemed to have a more sustained response to the agreement differences.

Word Region	Factor	Estimate	SE	t value	p-level
don't	intercept	525.03	10.99	47.78	<.001
	block	-37.42	3.04	-12.30	<.001
	agreement:nonstandard	12.02	8.31	1.45	
	agreement:uncommon	35.38	8.32	4.25	<.001
	awareness:aware	-44.67	8.20	-5.45	<.001
verb	intercept	546.90	11.81	46.30	<.001
	block	-47.86	3.27	-14.62	<.001
	agreement:nonstandard	32.71	8.95	3.65	<.001
	agreement:uncommon	68.06	8.95	7.61	<.001
	awareness:aware	-45.05	8.81	-5.12	<.001
verb+1	intercept	548.956	13.67	40.16	<.001
	block	-38.62	3.15	-12.26	<.001
	agreement:nonstandard	-16.11	18.34	-.88	
	agreement:uncommon	-18.56	18.41	-1.01	
	awareness:aware	-62.58	11.98	-5.23	<.001
	agreement:nonstandard X awareness:aware	32.22	20.80	1.55	
	agreement:uncommon X awareness:aware	46.36	20.86	2.22	<.05

Table 1. Pilot Experiment, summary of linear regression models (dependent variable is response time; models include random effect of items).

The finding that unaware participants were slower readers overall was surprising. It is intriguing given that much work on reading comprehension has shown correlations between poor reading skill and other cognitive and metacognitive processes, including "syntactic awareness" (Wagoner, 1983; Bowey, 1986; Gernsbacher, 1993; Nation & Snowling, 2000; Breadmore et al., 2013). Two unaware participants also reported having a reading disability, so there may be a relation between reading ability and sensitivity to these differences—however, these participants did no worse than others on the comprehension questions.

Because unaware participants had longer processing times overall, it seems that longer time spent processing a linguistic stimulus does not necessarily correlate with heightened

awareness of that stimulus, contrary to my intuition. Moreover, the participants who did not report noticing the manipulation seem to have been less strongly affected by it, or at least less consistently affected by it, supported by the interaction effect at the verb+1 region. To investigate possible differences between the two participant groups more closely, I conducted a supplemental analysis of the two groups independently, comparing separate regression models for the effect of agreement. For space purposes, Table 2 summarizes these results by simply reporting the significance of the agreement main effect for each group.

	don't		verb		verb+1	
	nonstandard	uncommon	nonstandard	uncommon	nonstandard	uncommon
Aware		<.001	<.001	<.001	<.01	<.01
Unaware				<.001		

Table 2. Pilot Experiment, supplemental analysis of the significance of agreement condition for aware participants (N=33) and unaware participants (N=10).

Agreement was a significant main effect at all word regions for the aware participants. However, for the unaware participants, agreement was only significant at the verb region, and then only in the uncommon condition. That the unaware participants did not show an effect of agreement at *don't* (the word introducing the grammatical variant) echoes Breadmore et al.'s (2013) findings, where subjects who were explicitly unaware were nonetheless implicitly aware of agreement mismatches, but the effect had a later onset. The failure of agreement to more robustly predict response times for unaware participants could be due to their data having wider variance due to longer total reading times, with that variance distributed across fewer observations.

Especially given that my measure of awareness—an off-line, metalinguistic self-report after the experiment—likely underestimates the number of participants truly noticing the

grammatical manipulation, the finding of differences between the two groups is quite compelling. Experiments 1 and 2 were designed to follow up on these initial results by investigating a) if the relationships between perceiving, noticing, and overall reading times could be replicated, and b) whether fewer tokens of non-standard sentences would lessen the overall noticing of the manipulation, to create a more balanced set of unaware versus aware participants.

Experiment 1

Experiment 1 differed from the Pilot Experiment in the proportion of nonstandard and uncommon sentences throughout the experiment. Experiment 1 lowered this proportion slightly, to 19 percent from 25 percent. In each of the four blocks in the experiment, participants saw 3 uncommon and 3 nonstandard sentences, along with 10 standard target sentences and 16 fillers sentences (so that throughout the experiment there were 24 total non-standard sentences out of 128). 36 participants received undergraduate course credit for their participation. Three participants' data were removed from analysis because they reported not having English as their native or most-fluent language.

18 participants were coded as "aware" and 15 were coded as "unaware." This presents a slightly more balanced grouping than in the Pilot Experiment—yet the majority of participants still reported noticing agreement differences. Figure 2 shows that the general results for Experiment 1 are similar to those for the Pilot Experiment. Aware participants were again faster readers overall, and the main effect of awareness was significant for all word regions, as shown by the statistical analyses summarized in Table 3.

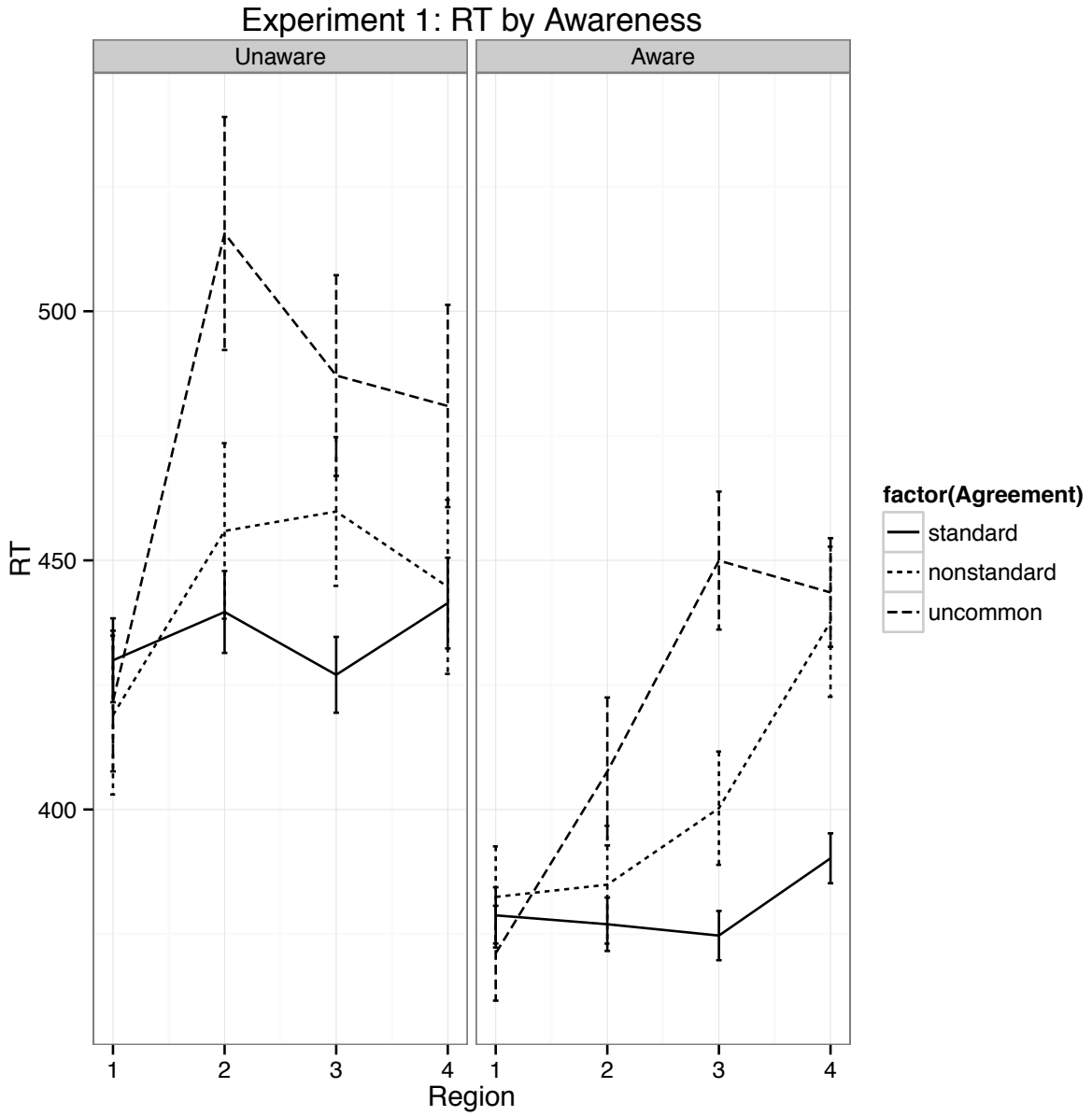

Figure 2. Experiment 1 reaction times in milliseconds by agreement and participant awareness, across word regions (1 = noun, 2 = don't, 3 = verb, 4 = verb+1). Error bars represent the standard error of the mean.

Word Region	Factor	Estimate	SE	t value	p-level
don't	intercept	561.17	11.68	48.04	<.001
	block	-46.40	3.74	-12.41	<.001
	agreement:nonstandard	11.72	11.01	1.06	
	agreement:uncommon	51.64	11.03	4.68	<.001

	awareness:aware	-72.68	8.40	-8.66	<.001
verb	intercept	538.72	13.06	41.26	<.001
	block	-45.00	4.32	-10.42	<.001
	agreement:nonstandard	28.84	12.71	2.27	<.05
	agreement:uncommon	68.36	12.71	5.38	<.001
	awareness:aware	-50.84	7.50	-6.78	<.001
verb+1	intercept	533.73	25.38	21.03	<.001
	block	-33.43	9.13	-3.66	<.001
	awareness:aware	-41.04	7.86	-5.22	<.001

Table 3. Experiment 1, summary of linear regression models (dependent variable is response time; models include random effect of items).

At the *don't* region, uncommon agreement again led to longer reading times. At the verb, both nonstandard and uncommon agreement took longer to read than standard. However, there was no main effect of agreement at the verb+1 region in this experiment, nor was there an interaction with awareness.

To probe further the differences between the groups, I again conducted a supplemental analysis treating the two participant groups independently, summarized in Table 4. While aware participants were affected by agreement in all three word regions, unaware participants were not affected at the verb+1 region (just as in the Pilot Experiment). And, at the verb region, only the uncommon sentences were significantly slower than standard for unaware participants, whereas for the aware participants, both nonstandard and uncommon sentences were (also as in the Pilot).

	don't		verb		verb+1	
	nonstandard	uncommon	nonstandard	uncommon	nonstandard	uncommon
Aware		<.01	<.01	<.001		<.05
Unaware		<.001		<.001		

Table 4. Experiment 1, supplemental analysis of the significance of agreement condition for aware participants (N=18) and unaware participants (N=15).

Experiment 1 replicated the Pilot Experiment in three ways. First, unaware participants were slower readers than aware participants. Second, for unaware participants only, sentences in the nonstandard condition never produced significantly longer reading times than those in the standard condition. Third, the difference between agreement conditions was not significant at the verb+1 region for unaware participants, but it was for aware participants. Both experiments suggest some qualitative difference in the experiencing of variation between those participants who later reported noticing the manipulation and those who did not. They seem to confirm that there is a distinction between what Broadmore et al. (2013) call *implicit* and *explicit awareness*, or what I am calling *perceiving* and *noticing.* Both participant groups *perceived* the agreement differences, but somewhat differently; and, they did not equally *notice* it.

Importantly, for both groups of participants, the form that caused the greatest disruption was the "uncommon" form [PL+*doesn't*]. What about the nonstandard pattern, which did not reach significance as different from standard for the unaware group? It could be that unaware participants did not perceive these because they had greater implicit knowledge of them—they may have been speakers of dialects who use the nonstandard form. If this were the case, we might expect some consistency in the demographic makeup of the unaware participants, reflecting similar dialect backgrounds. However, to the extent that there was social heterogeneity among participants (most of whom were White, and all of whom were university students), this was no more the case for the unaware groups than the aware groups.

Unaware participants read overall more slowly than aware participants, and so perhaps sensitivity to variation is tied to more general language processing or reading comprehension skills, which vary at an individual level. For instance, because this experiment is in the reading mode, it might be the case that the unaware participants were poorer readers than the aware

participants, indexed by their slower reading times. Reading skills have been linked by researchers to a variety of other cognitive abilities, including comprehension monitoring (Wagoner, 1983) and syntactic awareness (Bowey, 1986; Nation & Snowling, 2000).

One of the factors known to affect comprehension is the ability to suppress cues that are irrelevant or contradictory to the comprehension task (Gernsbacher, 1993). That is, when information is activated that "gets in the way" of comprehension, poor readers are worse at suppressing that information than good readers are, which makes comprehension more difficult. If variability in verb forms were considered a type of information that one needed to suppress in order to continue reading the sentence adeptly, we would expect poor readers to be *more* affected by agreement differences than good readers. That is, we would expect the slower readers to show more sensitivity to the agreement differences, and perhaps even more conscious awareness of them, since they would be less able to recover from the mismatches. However, it seems that the slower readers were less perceptive of the differences in the first place.

Perhaps unaware participants had a more difficult time processing the sentences in the task, and therefore had less processing energy to commit to formal differences that did not affect content (see Schmidt, 1990). Or, perhaps they were simply less likely to notice differences in general because of a lower degree of metalinguistic awareness or lower ability to monitor their own comprehension. If either of these were true, performance on the comprehension questions in the experiment should be worse for unaware than for aware participants. In the Pilot Experiment, unaware participants did average a lower percentage of correct comprehension question responses than aware participants (78% versus 87%), but this amounts only to a one-question difference in average accuracy between groups. In Experiment 1, average accuracy was equivalent between groups (88%). Generally speaking, then, failing to register awareness did not

align with poor comprehension. All participants seemed to be paying attention to the task and comprehending the sentences, regardless of the degree to which they perceived or noticed the agreement differences.

Experiment 2 sought to make agreement even less salient by reducing sharply the number of nonstandard/uncommon sentences during the experiment. In their investigation of the (ING) variable, Labov et al. (2011) suggest that the "sociolinguistic monitor" works as a logarithmic function, being extremely sensitive to the first few tokens of a socially marked variant and tapering off afterward. Kaschak & Glenberg (2004) and Kaschak (2006) also found that participants' sensitivity to dialect structures dissipated, representing adjustment or adaptation to the initially-unexpected forms. It is unclear whether speakers' tendency to register conscious awareness of what they perceive, however, is modulated by the number of variant tokens to which they are exposed during an experimental session. Experiment 1 used fewer non-standard tokens than the Pilot Experiment and had a higher proportion of unaware participants. Experiment 2 reduces the number further, attempting to further mitigate participants' noticing of the manipulation, to increase the number of unaware participants and continue probing the differences between groups described above.

Experiment 2

Experiment 2 was identical to Experiment 1 except that only one nonstandard and one uncommon sentence occurred in each experimental block, making for only 8 total non-standard sentences throughout the experiment (6 percent). 36 participants received extra credit for participating. Three participants' data were removed due to experiment error, and three participants' data were removed because they reported not having English as their native or

most-fluent language. Even with the small amount of non-standard tokens in this experiment, a majority of participants reported noticing agreement: 16 were coded as "aware" and 14 were coded as "unaware."

As can be seen in Figure 3 and Table 5, the results for this experiment were somewhat different from those of the Pilot and Experiment 1. There was no main effect of awareness on reading times at any region; overall, aware participants were not faster readers than unaware participants. However, there was an interaction effect between awareness and agreement at both the *don't* and verb+1 regions.

At *don't*, there was a main effect of both nonstandard and uncommon agreement, but the effect for the uncommon condition was stronger for the aware group than the unaware group. (This is visually apparent in Figure 2—compare region [2] between the groups.) At the verb, both nonstandard and uncommon agreement were slower than standard, and there was no interaction with awareness. At verb+1, there was a main effect of uncommon agreement, and an interaction effect such that only the aware participants were slowed by nonstandard agreement.

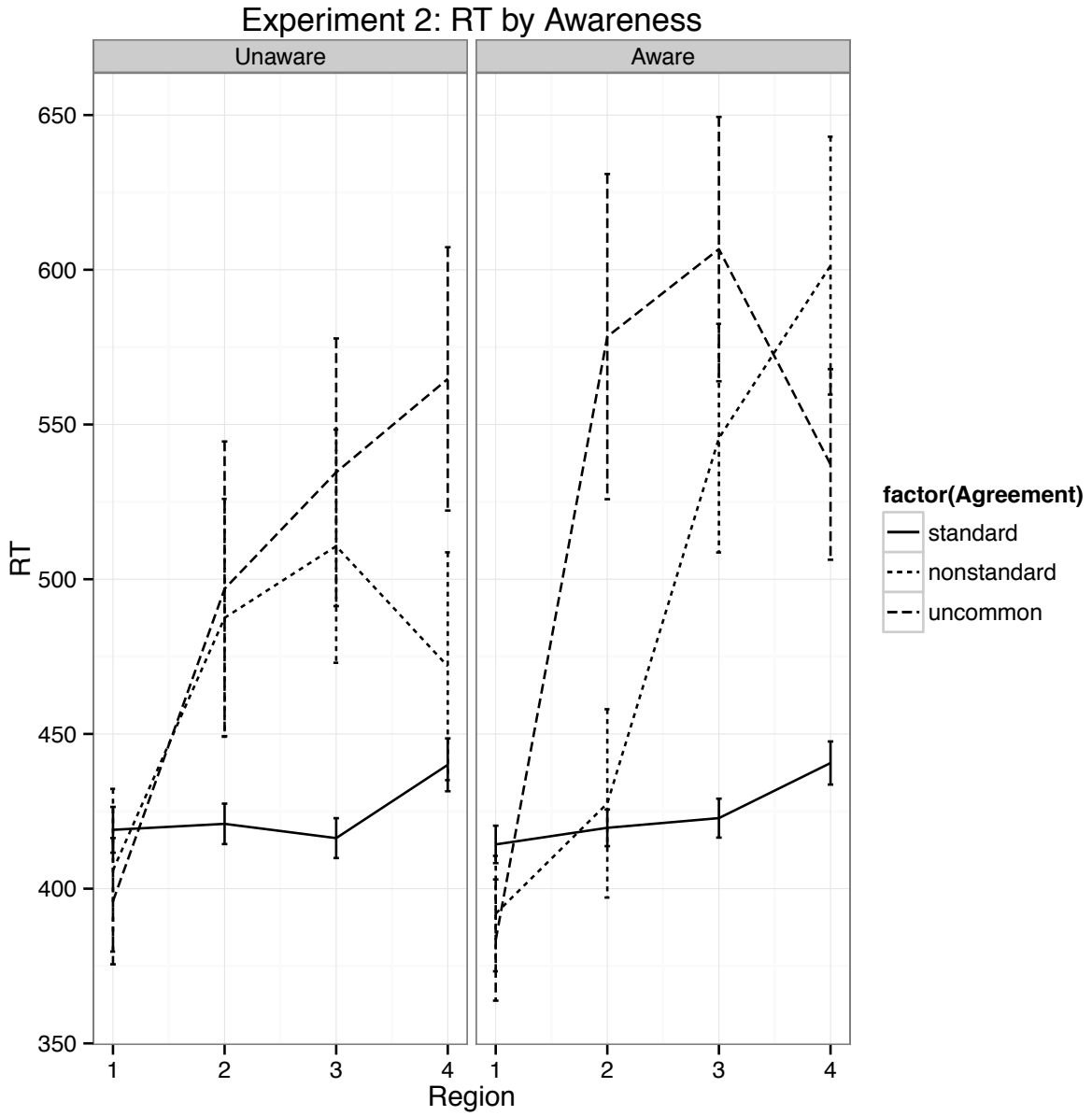

Figure 3. Experiment 2 reaction times in milliseconds by agreement and participant awareness, across word regions (1 = noun, 2 = don't, 3 = verb, 4 = verb+1). Error bars represent the standard error of the mean.

Word Region	Factor	Estimate	SE	t value	p-level
don't	intercept	518.32	14.94	34.70	<.001
	block	-38.92	5.06	-7.70	<.001
	agreement:nonstandard	66.61	30.98	2.15	<.05
	agreement:uncommon	74.49	30.98	2.40	<.05

	awareness:aware	-1.29	9.70	-.13	
	agreement:nonstandard x awareness:aware	-58.77	37.69	-1.56	
	agreement:uncommon x awareness:aware	85.76	37.82	2.27	<.05
verb	intercept	507.01	11.37	44.58	<.001
	block	-34.86	4.10	-8.50	<.001
	agreement:nonstandard	109.43	18.95	5.77	<.001
	agreement:uncommon	154.40	19.10	8.08	<.001
verb+1	intercept	520.20	33.51	15.52	<.001
	block	-31.56	11.93	-2.65	<.01
	agreement:nonstandard	41.44	59.75	.69	
	agreement:uncommon	123.44	59.09	2.09	<.05
	awareness:aware	-.47	10.22	-.05	
	agreement:nonstandard x awareness:aware	122.15	40.57	3.01	<.01
	agreement:uncommon x awareness:aware	-27.17	39.47	-.69	

Table 5. Experiment 2, summary of linear regression models (dependent variable is response time; models include random effect of items).

The supplemental independent group analyses confirm different agreement effects, presented in Table 6. For aware participants, agreement was significant at all three regions. At *don't*, the uncommon sentences were slower than standard, and at the verb, reading was slower for both nonstandard and uncommon sentences. At verb+1, aware participants were still slowed by nonstandard sentences. For unaware participants, at *don't* and the verb, nonstandard and uncommon sentences were slower than standard. Once again, agreement was not a significant predictor for unaware participants at the verb+1 region.

	don't		verb		verb+1	
	nonstandard	uncommon	nonstandard	uncommon	nonstandard	uncommon
Aware		<.001	<.001	<.001	<.01	
Unaware	<.01	<.01	<.001	<.001		

Table 6. Experiment 2, supplemental analysis of the significance of agreement condition for aware participants (N=16) and unaware participants (N=14).

Experiment 2 did not replicate the general difference in reading times between aware and unaware participants. While there may indeed be relations between reading ability, reading skill, or comprehension skill and perception/noticing of morphosyntactic differences, Experiment 2 did not confirm this, and more research is needed on the matter. Yet the interaction effects do indicate some differences in processing between the groups—with aware participants affected more strongly.

Experiment 2 also sought to test whether reducing the number of non-standard sentences in the experiment would reduce the number of participants noticing the agreement differences. This did not happen: the majority of participants reported noticing the manipulation, just as in the prior experiments. With only 8 non-standard tokens throughout the experiment, most participants nonetheless registered explicit awareness of the variation such that they remembered at the end of the experiment that they had encountered it. I discuss this in the section below.

General discussion

These experiments explored the relationship between the perceiving of linguistic variation, as measured by an on-line temporal measure, and the conscious noticing of that variation, as measured by participants' off-line self-reports. The results are consistent enough to support the idea that perceiving and noticing are usefully considered separate cognitive processes, as suggested by second language acquisition researchers such as Schmidt (1990). That is, while participants could be divided based on their noticing of differences, both groups nonetheless showed perception of the differences. The degree of processing disruption from unexpected agreement—including its duration and its magnitude—differed across groups,

though. Aware participants' behavior was more consistent across experiments, whereas unaware participants' behavior was noisier, particularly on either side of the verb region (in the *don't* and verb+1 regions).

Together, the results of the experiments permit three generalizations about the processing behavior of aware versus unaware groups. First, the effect of agreement lasted longer for aware participants than unaware participants; unaware participants were never significantly affected by agreement in the verb+1 region. Second, being affected by nonstandard agreement entailed also being affected by uncommon agreement, whose effect was generally larger. There was only one word region in one experiment for one participant group (Experiment 2, verb+1, aware participants) in which this generalization did not hold. Third, even very few tokens of variant grammatical patterns triggered metalinguistic awareness in the majority of participants; markedly reducing the proportion of non-standard tokens did not commensurably reduce the proportion of unaware participants (particularly from Experiment 1 to Experiment 2). The differences between groups—especially given the relatively unsophisticated method used here for measuring awareness—do suggest that perceiving and noticing are different components of sociolinguistic processing.

How do perceiving and noticing relate to knowledge? Sociolinguistic knowledge is the foundation for sociolinguistic perception, in that what we have implicit knowledge of delimits our perceptual expectations and adjustments (Beck, this volume; Staum Casasanto, 2009). The present study suggests that all kinds of grammatical anomalies are not equally perceived, and that this may be an outcome of the structure of participants' linguistic knowledge. I tested participants' reactions to two different kinds of grammatical "variants": one that they have likely experienced as spoken by real speakers and which is a social stereotype (SG+*don't*), and one that

they are unlikely to have experienced as systematically connected to real speakers (PL+*doesn't*). As expected, across experiments, the effect of agreement was most consistent at the verb—the word immediately following the introduction of variant agreement with *don't/doesn't*. In this region, the uncommon sentences always had the longest average reading times, and these were significantly longer than those for standard sentences for both aware and unaware participants. Reading times for the nonstandard sentences were slower than standard but faster than uncommon. The logical explanation for this gradience is that participants have some experience with the "nonstandard" pattern that makes it slightly less unexpected than the "uncommon" pattern, and patterns previously experienced are easier to process than patterns not experienced.

The implication is that grammatical forms that one is exposed to, even when they are different from one's own production baseline, are stored in memory—not passed over or discarded (Kaschak & Glenberg, 2004; Kashak, 2006). Of course, at some level, this must logically be the case in order to explain how people make social judgments based on dialect forms that they don't themselves control. The present experiments provide empirical evidence, though, that knowledge of these forms is implicitly active during processing, rather than only activated when a task or social situation evokes overt social stereotypes (see Campbell-Kibler, this volume).

As compared to the "nonstandard" pattern, exposing participants to the "uncommon" pattern is more akin to traditional sentence processing research which investigates participants' detection of syntactic errors, without considering whether these might be dialect forms or not (see discussion in Squires, 2013b). Participants may have been perceiving the nonstandard forms as "variants" but the uncommon forms as "errors." I think this is a central problem for sociolinguistic perception research moving forward: When linguistic differences are perceived,

and even moreso when they are noticed, how are they categorized? Is there a distinction between perception as *variant* and perception as *error*? If so, how do linguistic or social circumstances shift the categorization of incoming tokens? Is one more likely to lead to *noticing* and *understanding* than another? Is one more likely to be kept active in memory than another? These processes must be at the heart of the behavioral responses we see regarding sociolinguistic experience, social evaluations not least among them.

Though the present study did not investigate these questions specifically, the qualitative responses participants gave do speak to them. When participants noticed agreement differences, how did they describe them? What exactly did they notice? In Table 7, I present a list of strings used by "aware" participants across all three experiments. Note that a string's number of occurrences does not necessarily line up with number of participants who mentioned it, since some responses include the same word more than one time (especially *don't* and *doesn't*). (There are 67 total aware participants.)

String	**Tokens**
doesn't	39
don't	37
verb	32
correct	29
agree	25
subject	22
grammar	13
noun	12
grammatical	11
singular	8
wrong	7
plural	7
tense	6
proper	7

Table 7. Strings used by "aware" participants in describing the sentences.

Immediately interesting is the fact that *don't* was mentioned almost as many times as *doesn't*. These words each appeared in equal numbers of standard and non-standard constructions across the experiments, but because the uncommon form caused the greatest degree of processing difficulty, I expected *doesn't* to be noticed more often or more strongly than *don't*. This was not the case, which perhaps speaks to the preexisting salience of the nonstandard *don't* pattern. A few participants also attempted to re-create the "incorrect" sentences in their responses. Eight of these included a SG+*don't* pattern (only half with a full NP as in the experiment; half used a singular pronoun), while only two included a PL+*doesn't* pattern (both with a full NP). Though this analysis can only be qualitative, these responses suggest that participants were applying existing knowledge of SG+*don't* during the memory task of articulating what they had noticed. These forms were more accessible because they were activated by the nonstandard sentences, whereas PL+*doesn't* did not activate existing knowledge.

Though linguistic knowledge may have played a role in what participants noticed, their responses do not contain clear articulations about social meaning in the sense of stereotypes about individuals or speaker groups who might have produced these forms. As Table 4 shows, the strings "correct," "grammatical," "wrong," and "proper" occurred in several participants' responses. This is unsurprising given the social stigmatization of the SG+*don't* pattern and the artificiality of the PL+*doesn't* pattern. It is somewhat surprising that these terms were as close as participants came to articulating social judgments of the sentences. That is, none of their words indicate that they perceived the patterns as being dialect variants rather than as errors, though this may simply reflect the fact that in the US, dialect variants are commonly ideologized as errors (Preston, 1996). If participants perceived the nonstandard sentences as forms that English speakers might plausibly produce—as sociolinguistic variants rather than errors—this was not

evident in their self-reports. It is a non-trivial task for future work to understand the cognitive processing that moves one from the perception of linguistic forms to their interpretation as socially meaningful (if these can even be separated) (see also Squires, 2014).

I want to close with a few methodological comments. The experiments presented here have several limitations. First, they used written stimuli rather than spoken stimuli, which is the modality in which nonstandard grammatical forms are more likely to be both experienced and expected. Writing rigidly codifies forms, perhaps making nonstandard and uncommon agreement even more salient than they would be in natural spoken language (see also Meek and Choksi, this volume). The experiments did find the expected differences between the processing of the uncommon and nonstandard forms, but it would be fruitful to complement this work by investigating the on-line perception of spoken grammatical variation. Second, the self-report method of assessing what was noticed during the experiments results in highly variable information, uncontrolled for factors related to general metalinguistic awareness, verbal ability, reading skill, etc. If nothing else, I hope to have shown that we shouldn't be satisfied to think that behaviorally sensitive experimental behavior tells the whole story: what speakers do with the information they have perceived differs, and including such information in our interpretation of results might be informative.

Related is the more general issue of awareness and studies of sociolinguistic perception. Within sociolinguistics, there has been something of a privileging of what is "unaware," visible in the priority of eliciting the most vernacular speech and the most automatic social beliefs. There is also a privileging of "unaware" processing in psycholinguistics, which takes distractor or filler items as an indispensable component of experimental methodology, and seeks the most finely-tuned instruments to measure the most automatic behavioral responses or, even better,

non-behavioral (neural) responses. Against this backdrop, what are we to make of the fact that the majority of the participants in my experiments reported noticing the experimental manipulation? Probing speakers' perception and knowledge of socially-varying structures in particular may inevitably raise awareness, since what is perceptually salient is that which is different from what is expected—and our research questions are fundamentally about difference. But in the wild, too, sociolinguistic information may be processed in just this way: by perceiving difference, taking notice of it, and figuring out where it fits in with what is already known about language and the people who speak it.

Acknowledgments

Thank you to Anna Babel for motivating and editing this volume; thanks to Anna, an anonymous reviewer, and Kathryn Campbell-Kibler for comments on this chapter. For experimental and analytical assistance I owe Amanda Boomershine, Damon Tutunjian, Ruth Friedman, Kelsy Hernandez, and Sydney Watsek. I also thank Benjamin Munson, Kevin McGowan, Jen Nycz, Katie Carmichael, and John Rickford for their feedback and discussion at the 2013 LSA Annual Meeting. Finally, I'm grateful to Julie Boland and the participants in our 2013 LSA Linguistic Institute course, for valuable questions and rumination on matters related to this content. All weaknesses of this chapter are, naturally, my responsibility.

References

Baayen, R.H. (2010). languageR: Data sets and functions with "Analyzing Linguistic Data: A practical introduction to statistics." R package version 1.0. http://CRAN.R-

project.org/package=languageR.

Bates, D., Maechler, M. & Bolker, B. (2011). lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-39. http://CRAN.R- project.org/package=lme4.

Batterink, L. & Neville, H.J. (2013). The human brain processes syntax in the absence of conscious awareness. *The Journal of Neuroscience*, 33(19): 8528-8533.

Bowey, J.A. (1986). Syntactic awareness in relation to reading skill and ongoing comprehension monitoring. *Journal of Experimental Child Psychology* 41: 282-299.

Breadmore, H.L., Krott, A. & Olson, A.C. (2013). Agreeing to disagree: Deaf and hearing children's awareness of subject-verb number agreement. *The Quarterly Journal of Experimental Psychology*.

Bresnan, J. & Ford, M. (2010). Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language* 86(1): 168-213.

Campbell-Kibler, K. (2009). The nature of sociolinguistic perception. *Language Variation and Change 21*(01): 135-156.

Campbell-Kibler, K. (2010). Sociolinguistics and perception. *Language and Linguistics Compass* 4: 377-389.

De Vicenzi, M., Job, R., Di Matteo, R., Angrilli, A., Penolazzi, B., Ciccarelli, L. & Vespignani, F. (2003). Differences in the perception and time course of syntactic and semantic violations. *Brain and Language* 85: 280-296.

Docherty, GJ., Langstrof, C. & Foulkes, P. (2013). Listener evaluation of sociophonetic variability: Probing constraints and capabilities. *Linguistics* 51(2): 355-380.

"Feature 158: Invariant *don't* for all persons in the present tense." In: B. Kortmann and K. Lunkenheimer, eds. (2011). The electronic World Atlas of Varieties of English

[eWAVE]. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://www.ewaveatlas. org/parameters/158.

Gernsbacher, M.A. (1993). Less skilled readers have less efficient suppression mechanisms. *Psychological Science* 1993(4): 294-299.

Hanulíková, A., Van Alphen, P.M., Van Goch, M.M., & Weber, A. (2012). When one person's mistake is another's standard usage: The effect of foreign accent on syntactic processing. *Journal of Cognitive Neuroscience* 24(4): 878-887.

Johnstone, B. & Kiesling, S. (2008). Indexicality and experience: Exploring the meanings of /aw/-monophthongization in Pittsburgh. *Journal of Sociolinguistics* 12(1): 5-33.

Just, M.A., Carpenter, P.A. & Woolley, J.D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General* 111(2): 228-238.

Kaschak, M. P. (2006). What this construction needs is generalized. *Memory & Cognition 34*(2): 368-79.

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

Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia, PA: University of Pennsylvania Press.

Labov, W. (1973). Where do grammars stop? In R.W. Shuy (ed), *Report of the Twenty-Third Annual Round Table Meeting on Linguistics and Language Studies,* 43-88. Washington, DC: Georgetown University Press.

Labov, W., Ash, S., Ravindranath, M., Weldon, T., Baranowski, M. & Nagy, N. (2011). Properties of the sociolinguistic monitor. *Journal of Sociolinguistics* 15(4): 431-463.

Loudermilk, B.C. (2013). *Cognitive Mechanisms in the Perception of Sociolinguistic Variation*. University of California, Davis: PhD dissertation.

Mertz, E. & Yovel, J. (2003). Metalinguistic awareness. In J.-O. Ostman, J. Veschueren, J. Blommaert, & C. Bulcaen (eds.), *The Handbook of Pragmatics.*

Nation, K. & Snowling, M.J. (2000). Factors influencing syntactic awareness skills in normal and poor comprehenders. *Applied Psycholinguistics* 21: 229-241.

Pearlmutter, N. J., Garnsey, S. M. & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language* 41: 427-456.

Preston, D. (1996). Whaddayaknow: The modes of folk linguistic awareness. *Language Awareness* 5: 40-74.

Preston, D. (2011). The power of language regard – discrimination, classification, comprehension, and production. *Dialectologia* (Special Issue II): 9-33.

Schmidt, R.W. (1990). The role of consciousness in second language learning. *Applied Linguistics* 11: 129-158.

Silverstein, M. (1981). The limits of awareness. *Texas Working Papers in Linguistics* 84: 1-30.

Squires, L. (2013a). Limited bidirectionality in sociolinguistic perception. *Journal of Sociolinguistics* 17(2): 200-237.

Squires, L. (2013b). Talker specificity and the perception of grammatical variation. *Language and Cognitive Processes*. Ahead-of-print (DOI: 10.1080/01690965.2013.804193).

Squires, L. (2014). Knowledge, processing, evaluation: Testing the perception of English subject-verb agreement variation. *Journal of English Linguistics* 42(2): 144-172.

Staum Casasanto, L. (2009). Experimental investigations of sociolinguistic knowledge. Stanford, CA: Stanford University PhD dissertation.

Wagoner, S.A. (1983). Comprehension monitoring: What it is and what we know about it. *Reading Research Quarterly* 18(3): 328-346.

Wolfram, W. (1982). Language knowledge and other dialects. *American Speech* 57(1): 3-18.