BRIEF REPORT

Cues, quantification, and agreement in language comprehension

Darren Tanner 1,2,3 · Nyssa Z. Bulkes 1

Published online: 19 May 2015 © Psychonomic Society, Inc. 2015

Abstract We investigated factors that affect the comprehension of subject-verb agreement in English, using quantification as a window into the relationship between morphosyntactic processes in language production and comprehension. Event-related brain potentials (ERPs) were recorded while participants read sentences with grammatical and ungrammatical verbs, in which the plurality of the subject noun phrase was either doubly marked (via overt plural quantification and morphological marking on the noun) or singly marked (via only plural morphology on the noun). Both acceptability judgments and the ERP data showed heightened sensitivity to agreement violations when quantification provided an additional cue to the grammatical number of the subject noun phrase, over and above plural morphology. This is consistent with models of grammatical comprehension that emphasize feature prediction in tandem with cue-based memory retrieval. Our results additionally contrast with those of prior studies that showed no effects of plural quantification on agreement in language production. These findings therefore highlight some nontrivial divergences in the cues and mechanisms supporting morphosyntactic processing in language production and comprehension.

Darren Tanner and Nyssa Z. Bulkes contributed equally to this work.

- ☐ Darren Tanner dstanner@gmail.com
- Department of Linguistics, University of Illinois at Urbana-Champaign, 4080 Foreign Languages Building, MC-168, 707 S. Mathews Ave., Urbana 61801, IL, USA
- Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL, USA
- Neuroscience Program, University of Illinois at Urbana-Champaign, Urbana, IL, USA

Keywords Grammatical agreement · Language comprehension · P600 · ERP · Memory retrieval

Grammatical agreement is an integral part of the morphosyntax of many languages, including English, and as such, must routinely be processed during language use. For example, the production of well-formed sentences in English requires speakers to compute dependencies between discontinuous elements (e.g., a subject noun phrase [NP] and verb), sometimes at a great distance. Behavioral psycholinguistic research from speech production has repeatedly shown that this process can be error-prone (e.g., Bock & Miller, 1991; Brehm & Bock, 2013; Staub, 2009). Importantly, this line of research has also shown that examining the factors that influence error rates can provide a clear window into the cognitive mechanisms supporting the computation of grammatical relations, as well as identify the range and role of linguistic cues that influence how we compute agreement and discontinuous dependencies in language more generally.

Recent research has also highlighted the occasionally errorprone nature of agreement in comprehension, and specifically shown that agreement processing in comprehension can also be somewhat fragile and show gradient effects of grammaticality when a subject NP contains potentially ambiguous or misleading cues to the grammatical number of the subject (e.g., Tanner, Nicol, & Brehm, 2014; Wagers, Lau, & Phillips, 2009). In light of this, some important research goals are to better understand (1) the range of cues used in the processing of agreement dependencies (e.g., those that modulate error rates and sensitivity to agreement relations), and (2) the potential convergence or divergence in the cues and mechanisms used during language production and comprehension. Here, we use recordings of event-related brain potentials (ERPs) to investigate the cues used when



processing subject—verb agreement relations in comprehension, focusing specifically on the role of quantification as a cue to these dependencies. This will provide insight into the information used to establish discontinuous dependencies in language processing, and more broadly, help us to better understand the relationship between morphosyntactic processes in language production and comprehension.

Research on language production has shown that a primary determinant of agreement marking on verbs is the mental representation of the subject NP's number feature as a whole, which is a product of the grammatical number specification of the noun(s) in the subject NP and of conceptual number information from the message to be conveyed (e.g., Bock, Carreiras, & Meseguer, 2012; Eberhard, 1997, 1999; Humphreys & Bock, 2005). This line of research has been formalized in a model known as Marking and Morphing (MM; Eberhard, Cutting, & Bock, 2005). Although much of the empirical data for the MM model comes from studying interference effects from embedded nouns in complex subject NPs ("agreement attraction" effects: Bock & Miller, 1991), the model makes general predictions about how agreement is computed in production, where information from each noun in the sentence can potentially contribute meaningful information relevant to subject NP number representation and the resulting computation of verbal agreement. Two features of the model are important to the present study and warrant description. First, within MM, the mental representation of the entire subject NP's number feature is not discrete, but instead is continuously valued, such that some NPs can be "more plural" than others. In this way, the relative plurality of the entire NP determines the probability of producing a plural-marked verb.

Second, number markedness plays an important role in the model: MM holds that plural nouns are overtly marked for number, whereas singular nouns are unmarked for number (i.e., they are neither overtly singular nor plural, but are default forms unspecified for number). This conception of number markedness has been supported by a range of findings from error rates and reaction times in production (e.g., Bock & Miller, 1991; Brehm & Bock, 2013; Eberhard, 1997; Staub, 2009), as well as from computational modeling (Eberhard et al., 2005). Relevant to the present study, Eberhard (1997) showed in a language production task that plural nouns embedded within a larger singular NP (e.g., The key to the cabinets . . .) influenced the number representation of the overall NP, indexed by participants' higher likelihood of producing erroneous plural nouns, relative to a condition with singular embedded nouns. Providing an additional cue to plurality on the embedded noun via quantification (e.g., The key to a few cabinets . . .) did not lead to an overall increase in plural marking on the verbs. In fact, there was a trend toward decreased agreement attraction in this condition, suggesting that participants' representations of the subject NP's number feature was not more plural in the presence of plural quantification than of unmarked determiners (i.e., the). Effects of quantification were evident for singular nouns, however, in which overt singular quantification (e.g., The keys to one cabinet) did cause an increase in erroneous singular verb marking over and above unquantified singular nouns (e.g., The keys to the cabinet). The conclusion drawn by Eberhard (1997; see also Eberhard et al., 2005) was that, given that they already bear a marked feature, plural NPs cannot become "more plural" by adding an additional cue through quantification. Thus, as opposed to singular quantification, plural quantification showed little effect on the processing of verbal agreement in language production.

One might ask, then, how markedness and number representation play out during language comprehension. Many general models of language processing have recently highlighted the interdependence of language production and comprehension processes, and in particular, have noted that language production mechanisms may play an important role in language comprehension via covert production and prediction (Dell & Chang, 2014; Pickering & Garrod, 2013; see also Gennari & MacDonald, 2009; MacDonald, 2013). Similar comprehension/production overlaps have also been proposed for agreement processes, in particular (e.g., Nicol, Forster, & Veres, 1997; Pearlmutter, Garnsey, & Bock, 1999; Severens, Jansma, & Hartsuiker, 2008; Thornton & MacDonald, 2003; Wagers et al., 2009). On the assumption that agreement processing mechanisms are largely shared across language comprehension and production, a salient hypothesis is that plural quantification should have no impact on the perception of agreement during language comprehension.

In contrast, Tanner and colleagues (2014) have recently proposed that, although not fully distinct, morphosyntactic processes in language comprehension and production might prioritize different cues. With specific regard to agreement, Tanner et al. suggested that agreement in production shows influences from numerous sources, but that a strong cue to verbal agreement is the number representation of the subject NP, in line with the MM model. Agreement in comprehension, however, relies more strongly on a combination of morphosyntactic predictions updated incrementally as linguistic input accrues over time, in tandem with cue-based memory retrieval mechanisms deployed when predictions fail or when verb-argument integration necessitates retrieval (see also Dillon, Mishler, Sloggett, & Phillips, 2013; Wagers et al., 2009). According to this approach, strong cues to an NP's number feature should facilitate the encoding and prediction of upcoming verbal agreement features. When predictions fail, these cues also provide retrieval targets for memory searches that seek out an appropriate agreement controller (see Lewis & Vasishth, 2005; McElree, Foraker, & Dyer, 2003; Van Dyke, 2007; Van Dyke & Johns, 2012; and Van Dyke & McElree, 2011, for other cue-based retrieval accounts of sentence



comprehension). From this perspective, it is conceivable that multiple encodings of an NP's plural feature during incremental language comprehension (e.g., via both a plural quantifier and plural morphology on the noun itself) should create stronger expectations for a plural verb, leading to increased processing difficulty when the predictions fail. Similarly, when the prediction of verbal agreement features fails upon encountering a disagreeing verb, plural quantification should provide a stronger retrieval target cue than numberambiguous determiners such as "the," which occurs with both singular and plural nouns. The prediction/retrieval account thus holds that plurally quantified subject NPs should increase sensitivity to agreement mismatches at the verb, relative to unquantified NPs with number-neutral determiners, in contrast to the predictions made by representational accounts of number markedness from language production (Eberhard, 1997).

A third salient possibility is that quantification could decrease sensitivity to agreement violations. Cross-linguistic studies have shown that plural quantification of an NP can actually reduce the likelihood of plural agreement on a predicate. For example, plural numerical quantification of a noun in many languages co-occurs with singular marking on the noun itself, and with singular predicate agreement (Corbett, 2000). In these cases, plural agreement on nouns, verbs, and other predicates is redundant with the notional plurality clearly expressed via the quantifier. Some have proposed that typological patterns like this in language can arise from processing considerations, when languages' forms are shaped by processing demands (e.g., Hawkins, 2011; MacDonald, 2013). Indeed, some studies now show that redundant morphosyntax can lead to comprehension and processing difficulty in highly inflected languages (Caballero & Kapatsinski, in press; Harris & Samuel, 2011), and that multiple markings can be particularly costly when other informative cues are already present (Caballero & Kapatsinski, in press; Kail, 1989). Additionally, research in English from populations with constrained processing capacity has also shown that redundant lexical and grammatical marking can impair comprehension (e.g., children with language impairment: Montgomery, 2000) or may not be processed at all (e.g., nonnative language comprehension: VanPatten, 2004). Thus, processing economy considerations could lead to a reduced sensitivity to number agreement on a verb (where it is a redundant marking) when number is more strongly expressed in the NP itself (where it is semantically relevant).

Notably, behavioral methods of studying comprehension could be insensitive to subtle processing differences that are associated with quantificational cues to agreement, especially in such languages as English, in which agreement is obligatory and behavioral performance could be at ceiling. Using ERPs to investigate these issues can provide an important window into the processes underlying agreement comprehension, because ERPs have been shown to be sensitive to subtle

processing differences that go undetected in behavioral measures (e.g., McLaughlin, Osterhout, & Kim, 2004; Thierry & Wu, 2007; Tokowicz & MacWhinney, 2005). We investigated these hypotheses by recording neural activity during sentence comprehension in native English speakers. Our participants read sentences with verbs that were made either grammatical or ungrammatical by manipulating subject-verb number agreement, following subject NPs either that were doubly plural marked, via quantifiers (i.e., NPs with a plural quantifier and a plural-marked noun, e.g., Many cookies . . .), or that showed plural marking on the noun itself, but with a number-neutral determiner (e.g., The cookies . . .). Previous research using ERPs to study agreement comprehension has shown that violations elicit a large positive-going brainwave, beginning approximately 500 ms after stimulus onset with a parietal maximum (the P600 effect), which is sometimes (but not always) preceded by a negative-going wave, sometimes with a left anterior distribution (the LAN effect; e.g., Hagoort, Brown, & Groothusen, 1993; Osterhout & Mobley, 1995; see Molinaro, Barber, & Carreiras, 2011, for a review). Importantly, the P600 is very sensitive to subtle processing differences, as the amplitude of the positivity has been shown to vary in a graded fashion with perceived ungrammaticality and general processing difficulty, including in agreement studies (e.g., Nevins, Dillon, Malhotra, & Phillips, 2007; Tanner et al., 2014; see also Kaan & Swaab, 2003; Osterhout, Holcomb, & Swinney, 1994). Thus, for quantified versus unquantified sentences, the MM/markedness account predicts no differences in P600 effect size, the prediction/retrieval account predicts larger P600 effect sizes, and the redundancy account predicts smaller P600 effect sizes.

Method

Participants

The participants were 24 monolingual, native American-English speakers. All were right-handed (Oldfield, 1971), had normal or corrected-to-normal vision, and reported no history of neurological impairment or use of psychoactive medication. The data from four participants were excluded due to excessive artifacts in the raw EEGs, resulting in 20 participants' data being included in the final analysis (12 females, eight males; ages 18 to 23 years, M=20 years). All participants provided informed consent and received a small amount of cash.

Materials

A total of 120 sentence frames were created in a 2×2 design, crossing the factors Grammaticality (grammatical, ungrammatical) and Subject NP Quantification (quantified, unmarked



determiner). The sentences contained a simple plural NP subject that either was overtly quantified (with *some* or *many*; 50 % each quantifier) or contained a determiner unmarked for number (*the*). Grammaticality was manipulated by varying the agreement status of the verb (e.g., *Many/The cookies taste/*tastes the best when dipped in milk*). The critical verbs were never in sentence-final position. The four versions of each sentence frame were distributed across four lists using a Latin square design, such that each list contained 30 items per condition, and no participant saw two versions of the same item.

To ensure that we had sufficient statistical power to detect small modulations of the P600, we included a filler condition in which we expected to observe graded P600 effects. A set of 90 sentence frames was adapted from Osterhout, Holcomb, and Swinney (1994) in order to examine the effects of verb transitivity bias on verb subcategorization processing with intransitive verbs (i.e., verbs disallowing NP complements), transitively biased verbs (i.e., verbs that occur more frequently with NP than sentential complements), and fully transitive verbs (e.g., The captain agreed/heard/*bought the crew was unhappy during the voyage). The sentences were distributed across three experimental lists, with 30 sentences in each list. These three lists were counterbalanced across the four lists from the agreement experiment, resulting in 12 total list combinations. Because the results from the filler conditions were consistent with previous work (Osterhout et al., 1994), we do not report them here.

An additional 75 filler sentences were included so as to avoid potential participant bias within the experimental context for each of the verbs in the subcategorization experiment. These filler items included 25 sentences with transitive verbs in transitive frames, resulting in grammatical sentences (e.g., *The dentist forced the patient to brush more often*); 25 sentences with intransitive verbs in transitive frames, resulting in ungrammatical sentences (e.g., *My neighbor agreed the book on my shelf); and 25 sentences with transitive-biased verbs in sentence frames that varied in their grammaticality (12 ungrammatical, 13 grammatical), where ungrammaticalities were created via word order or other violations after the direct object (e.g., *The judge heard the plaintiff during the his testimony). The resulting lists each had 285 sentences, of which 55 % were grammatical and 45 % were ungrammatical.

Procedure

Each participant was tested in a single session, lasting around 2.5 h. Upon arriving in the lab, participants were asked to complete a language background questionnaire and an abridged version of the Edinburgh Handedness Inventory. Each participant was seated in a comfortable chair in front of a computer monitor and was randomly assigned to an experimental list. Participants were asked to try to relax and to minimize blinks and movements during stimulus presentation. Each trial

consisted of the following series of events. A fixation cross appeared onscreen for 300 ms, followed by a blank screen for 200 ms. Sentences were presented one word at a time, with each word remaining onscreen for 300 ms and a blank screen for 200 ms between words. Sentence-final words appeared with a period, followed by a blank screen displayed for 950 ms. A "Good/Bad?" screen appeared following each sentence, to which participants were asked to respond by pushing a button on a response pad corresponding to the correct answer. Participants were instructed to press "good" for sentences that they felt were well-formed and semantically coherent, and "bad" for sentences that they felt were ungrammatical or did not make sense. The experiment was preceded by six practice sentences, with two syntactically and one semantically ill-formed sentence. The "good" response hand was counterbalanced across participants. Between each trial, the word "Ready?" appeared onscreen, during which time the participants were encouraged to blink or take breaks whenever necessary. Participants pressed a button on the response pad to begin the next trial at their own pace. The experiment was broken into five blocks, with participants being given a short break between blocks.

Data acquisition and analysis

Continuous EEG was recorded from 28 tin scalp electrodes mounted in an elastic cap (Electro-cap International), in accordance with the extended 10–20 system (Jasper, 1958; FP1, FP2, F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6, T7, C3, Cz, C4, T8, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, O1, Oz, and O2). Eye movements were monitored with electrodes placed below the left eye (referenced offline to FP1) and at the outer canthus of each eye (referenced offline in a right-to-left bipolar montage). Trials characterized by excessive eye movements, drift, alpha, or other artifacts were rejected; all data from any participant with greater than 30 % rejected trials in any single experimental condition were excluded from analysis, resulting in the rejection of four participants (see above). In the remaining 20 participants, an average of 11.5 % of trials were excluded across the conditions. Scalp electrodes were referenced online to the left mastoid and re-referenced offline to the algebraic mean of activity over the left and right mastoids. Impedances at scalp and mastoid sites were held below 5 k Ω , and impedances were kept below 15 k Ω for the ocular electrodes.

The EEG was amplified with a BrainAmpDC bioamplifier system (Brain Products, Gilching, Germany) and digitized with a 1000-Hz sampling rate and an online analog 0.016- to 250-Hz bandpass filter. All offline processing was carried out with EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014) toolboxes. A 0.1- to 30-Hz bandpass filter (-6-dB cutoff, 12-dB/octave roll-off) was applied to the continuous EEG offline, and data were downsampled to 200 Hz. ERPs were averaged time-locked to the



onset of the critical word (underlined above) for each participant over each electrode in each condition, relative to a 200-ms prestimulus baseline. All artifact-free trials were included in the analyses.

ERPs were quantified using mean amplitudes within time windows chosen a priori (300-500 and 500-800 ms), corresponding to the LAN and P600 time windows, respectively. Repeated measures analyses of variance (ANOVAs) were computed within each time window separately for midline (Fz, Cz, Pz) and lateral sites. For midline sites, Grammaticality, Quantification, and Anteriority were included as factors. Data for lateral sites were grouped into four regions of interest: left frontal (F7, F3, FC5, FC1), right frontal (F4, F8, FC2, FC6), left posterior (CP5, CP1, P7, P3), and right posterior (CP2, CP6, P4, P8). For lateral sites, Grammaticality, Quantification, Anteriority (anterior, posterior), and Hemisphere (right, left) were included as factors. The Greenhouse-Geisser correction for inhomogeneity of variance was applied to all repeated measures with more than one degree of freedom in the numerator. In such cases, the corrected p value is reported. Sensitivity to agreement violations in the quantified and unquantified conditions from the behavioral acceptability judgment task was quantified using A statistics (Zhang & Mueller, 2005).

Results

Behavioral results are reported in Table 1. Overall, participants were more sensitive to the agreement violations in the quantified than in the unquantified condition, indicated by a statistically significant difference in A scores (quantified minus unquantified mean difference=.042, SE=.012, 95 % CI: .017–.066), t(19)=3.580, p=.002.

The ERP results showed a similar heightened sensitivity to agreement violations following quantified relative to unquantified NPs. Grand mean waveforms comparing the brain responses to grammatical and ungrammatical verbs are depicted in Figs. 1 and 2 for the quantified and unquantified conditions, respectively. In the 300- to 500-ms time window, statistical analysis showed no significant effects or interactions, indicating no reliable LAN effect. Statistics for the 500- to 800-ms window are presented in Table 2. We found a significant effect of grammaticality that showed a parietal maximum (Grammaticality×Anteriority interaction),

 Table 1
 Acceptability judgment results

Condition	Grammatical Proportion Correct	Ungrammatical Proportion Correct	A statistic	
Quantified	.95 (.01)	.87 (.04)	.946 (.02)	
Unquantified	.93 (.01)	.79 (.05)	.905 (.02)	

Standard errors are in parentheses.

typical of P600 effects. Over lateral sites, the three-way interaction between grammaticality, anteriority, and hemisphere indicated that the P600 effect had a right hemisphere bias and extended farther into anterior regions over the right than over the left hemisphere.

Crucially, the Grammaticality×Quantification interaction was fully significant over midline electrodes and neared significance over lateral sites. Figure 3 depicts this interaction over parietal electrode sites, where the P600s were largest. Brain responses were more positive-going in the quantified than in the unquantified ungrammatical condition, whereas there was little difference between the brain responses in the two grammatical conditions. Figure 4 further depicts P600 effect sizes with ungrammatical-minus-grammatical difference waves. P600 effects were larger in the quantified (mean amplitude difference over midline electrodes=3.383 μ V, SE=0.77) than in the unquantified condition (mean=1.578 μ V, SE=0.63), indicating heightened sensitivity to ungrammaticalities and increased processing difficulty in this condition.

Discussion

The results of this study clearly show that overt quantification influenced the perception of subject-verb agreement violations in English. Both behavioral (sentence acceptability judgment accuracy) and electrophysiological measures (P600 amplitude) converged on the finding that sensitivity to agreement-violating verbs increases when the verb is preceded by a subject NP with overt plural quantification. Our results align most closely with the prediction/retrieval account of agreement comprehension outlined in the introduction (Dillon et al., 2013; Tanner et al., 2014; Wagers et al., 2009). According to this approach, morphosyntactic sensitivity—indexed here by P600 amplitude and acceptability judgment accuracy—is a function of both predictive and retrieval cue quality. In the context of the present study, quantified noun phrases contained stronger cues to NP plurality than did NPs with a number-ambiguous determiner (the), both for prediction and retrieval.

First, in terms of prediction, plural quantifiers provide an immediate predictive cue to upcoming verb number, even before a plural noun is encountered. For example, in English, many plural quantifiers (including those in the present study) allow for noun ellipsis, such that a plural quantifier and verb can cooccur with no intervening noun (e.g., *Many were unhappy*). Because the word following a plural quantifier could in fact be a plural-marked verb, comprehenders can immediately begin anticipating an upcoming plural verb before encountering a noun, providing a temporal advantage in prediction over unmarked determiners. Additionally, plural quantifiers cooccurring with plural nouns provide a double-encoding of the

¹ We thank an anonymous reviewer for pointing this out.



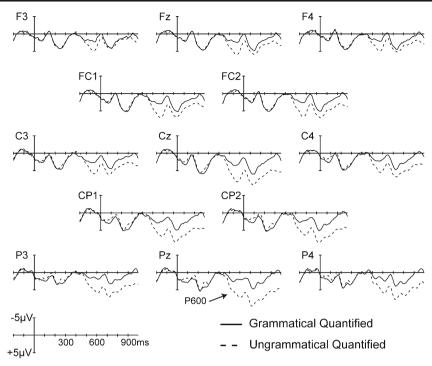


Fig. 1 Grand mean ERP waveforms for grammatical (solid lines) and ungrammatical (dashed lines) verbs in the quantified condition over 13 representative electrodes. Negative voltage is plotted up. The waveforms depict 200 ms of prestimulus and 1,000 ms of poststimulus activity; each

tick mark represents 100 ms. The vertical calibration bar shows $\pm 5\,\mu V$. These and all subsequent waveforms were filtered with a 15-Hz low-pass filter for presentation purposes only

subject NP's plural number feature. This double encoding and temporal advantage can lead to stronger expectations for upcoming plural verbal agreement features, relative to NPs with determiners that are unmarked for number, which in turn can lead to stronger processing responses when these expectations are violated.

Second, in terms of retrieval, multiple markings of plurality in the subject NP provide stronger disambiguating cues when

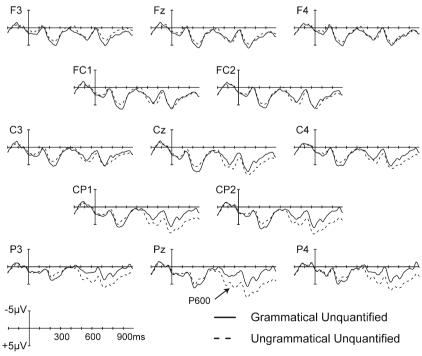


Fig. 2 Grand mean ERP waveforms for grammatical (solid lines) and ungrammatical (dashed lines) verbs in the unquantified condition over 13 representative electrodes



Table 2 Results from the omnibus ANOVA in the 500- to 800-ms time window

Contrast	df	F Value	MSE	p
Midline				
Gram	1, 19	16.837	21.921	<.001
Quant	1, 19	<1	9.939	.496
Gram×Quant	1, 19	5.393	9.059	.031
Gram×Ant	2, 38	32.705	2.014	<.001
Quant×Ant	2, 38	<1	1.858	.688
$Gram \times Quant \times Ant$	2, 38	1.852	1.781	.188
Lateral				
Gram	1, 19	11.946	13.622	.003
Quant	1, 19	<1	5.728	.577
Gram×Quant	1, 19	4.222	6.804	.054
Gram×Ant	1, 19	20.692	2.708	<.001
Quant×Ant	1, 19	<1	1.175	.471
Gram×Hem	1, 19	1.433	0.754	.246
Quant×Hem	1, 19	<1	0.624	.513
Gram×Ant×Hem	1, 19	6.188	0.215	.022
Quant×Ant×Hem	1, 19	<1	0.108	.702
$Gram \times Quant \times Ant$	1, 19	2.928	1.452	.103
$Gram \times Quant \times Hem$	1, 19	1.784	0.216	.197
$Gram \times Quant \times Ant \times Hem$	1, 19	<1	0.173	.574

Significant and near-significant results are bolded. Gram=grammaticality; Quant=quantification status; Ant=anteriority; Hem=hemisphere

violations of agreement predictions trigger retrieval processes for possible controllers in working memory. Work in cue-based models of comprehension has shown that both processing difficulty and the likelihood of correctly retrieving the intended target are partly a function of the overlap in retrieval cues between those to the item triggering the retrieval and those to potential candidates held in memory (e.g., Dillon et al., 2013; Fedorenko, Gibson, & Rohde, 2006; Gordon, Hendrick, & Johnson, 2001; Gordon, Hendrick, & Levine, 2002; Lewis & Vasishth, 2005; Lewis, Vasishth, & Van Dyke, 2006; see also Badecker & Kuminiak, 2007). Even in the absence of potentially irrelevant intervening NPs, this content-addressable

mechanism should ensure that the more explicit the encoding of agreement features, the more likely it is that a dependency will be resolved without error. Thus, although the simple NPs in our stimuli did not contain potentially interfering attractor nouns, the double-encoding of number in quantified NPs provides a stronger disambiguating cue to the subject NP's number feature than do NPs with unmarked determiners. That is, if retrieval processes are initiated upon encountering an agreement-violating verb, stronger agreement mismatch effects would be expected when the plurality of the subject NP is more explicitly marked for number (see also Severens et al., 2008).

Although the present results cannot unambiguously distinguish whether prediction, retrieval, or a combination of the two led to the enhanced sensitivity to disagreeing verbs in the quantified condition, we have some reason to believe that the predictability of agreement features may have played a stronger role in shaping our results than retrieval accuracy. Notably, research using the speed-accuracy trade-off paradigm has shown an advantage for accessing the most recently processed item in a linguistic string when verb-argument integration is necessary. In such cases, full retrievals may not be necessary, since the most recently processed item can remain in the focus of attention (McElree, 2006; McElree et al., 2003). In the present study, the head noun of the subject NP and verb were always linearly adjacent, such that information from the noun could have remained in the focus of attention when processing the verb, potentially obviating the need for retrieval. Thus, although prediction and retrieval mechanisms can work together in shaping processing behavior, the most parsimonious explanation of our results is that the prediction and encoding of agreement features via quantification led to the heightened sensitivity to agreement violations on the verb.

Note that our findings go beyond previous accounts of prediction and retrieval in agreement comprehension, which argued that prediction was primarily responsible for the asymmetric effects of agreement attraction with respect to sentence grammaticality (Dillon et al., 2013; Tanner et al., 2014; Wagers et al., 2009). In those studies, prediction was implicated to explain the lack of a difference in reading times or ERP effects at grammatical verbs following singular versus plural

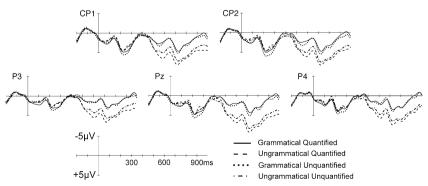


Fig. 3 Grand mean ERP waveforms contrasting brain responses in all four experimental conditions over five centro-parietal electrode sites



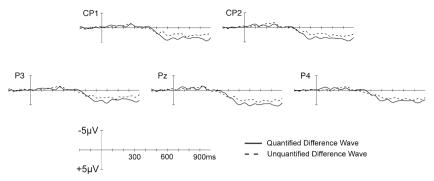


Fig. 4 Grand mean difference waves for the ungrammatical minus grammatical conditions in both the quantified (solid lines) and unquantified (dashed lines) conditions over five centro-parietal electrode sites. Positive deviations from zero indicate a P600 effect

attractors (e.g., when predicted features are encountered, no retrievals are necessary, such that no attraction effects ensue), whereas retrieval was implicated in accounting for differences in the processing of ungrammatical verbs following singular versus plural attractors (e.g., determining the relative degree of processing difficulty for ungrammatical verbs). Our results show not only that prediction in agreement comprehension can be responsible for determining whether or not retrieval processes are initiated (Dillon et al., 2013; Tanner et al., 2014; Wagers et al., 2009), but also that the strength of predictions can determine the processing costs when these predictions are violated. Degree of predictability and prediction accuracy have been shown to be strong determinants of brain responses in prior ERP work on both lexical and syntactic processing, in which earlier and stronger ERP responses were associated in a gradient fashion with increased predictability and prediction accuracy of words in context (e.g., Brothers, Swaab, & Traxler, 2015; DeLong, Urbach, & Kutas, 2005; Wlotko & Federmeier, 2012) and with violations of probabilistic syntactic subcategorization biases (e.g., Osterhout et al., 1994). Here we have shown that this sort of graded sensitivity to prediction strength is also found with simple, obligatory morphosyntactic dependencies such as subject-verb agreement in simple NP contexts, even when there is no possibility for similarity-based retrieval interference (cf. Tanner et al., 2014).

Our results also suggest that cross-linguistic generalizations about the relationship between quantificational number and plural agreement do not hold in the comprehension of English. In reviewing number agreement phenomena in Slavic languages, Corbett (2000) notes that increasing numerosity expressed via numerical quantification is often associated with a decreased likelihood of plural marking on verbs and other agreement-taking elements in a sentence. This suggests that in some languages, the expression of number on the quantifier obviates the need to redundantly mark plurality on the verb via agreement. Because agreement is obligatory in standard English, corpus analyses like those used by Corbett or experimental elicitation tasks like those used in agreement attraction studies would be unlikely to reveal any differences

in the use of agreement in quantified versus unquantified contexts; correct number agreement would likely be at ceiling. However, sensitive processing measures like ERPs could potentially reveal differential sensitivity to agreement on verbs in such contexts. In contrast to the agreement phenomena in Slavic languages, our ERP results from English suggest that plural quantification enhances the salience of agreement on verbs, despite the relative semantic vacuousness of number marking on verbs (see Chomsky, 2001). Although our findings are suggestive of an important role for quantification in the computation of agreement when comprehending English, future cross-linguistic research will be needed to address how such cues influence the production and comprehension of agreement in real time in other languages. For example, research will be needed to understand how factors such as the morphological richness of agreement marking interact with notional cues to number (see Foote & Bock, 2012) or the extent to which morphological richness interacts with cue informativity during real-time comprehension (see Caballero & Kapatsinski, in press).

Furthermore, our results contrast with prior work on language production showing that quantification of plural nouns had little effect on number marking on verbs (Eberhard, 1997). Eberhard showed that embedding plural nouns within an otherwise singular NP (e.g., The key to the cabinets . . .) caused an increase in (erroneous) plural marking on verbs when participants completed sentence fragments, whereas quantifying the plural-embedded noun phrase (e.g., The key to a few cabinets . . .) caused a trend toward a decrease in the proportion of erroneous plural marking on verbs. This trend was driven by a significant increase in the likelihood of producing a correctly marked singular verb in the pluralquantifier versus the unquantified condition. This suggests that providing an additional cue to plurality via quantification does not cause speakers' representations of the subject NP's number feature to become "more plural" than when the number feature is marked only on the noun itself via plural morphology, as conceptualized within the MM model. Taken in the context of MM as a model for agreement in production, our results suggest that there are interesting differences in how



quantificational cues are used when formulating agreement in production and when processing it in comprehension.

It is important to ask whether the differences between the results observed here for comprehension and those observed by Eberhard (1997) for production might not have resulted from comprehension/production differences, but instead from differences in the syntactic constructions used in the two studies. We used simple NPs, in which the quantified expression was the head of the NP; Eberhard used complex NPs, in which the quantified expression was embedded. Perhaps our positive result for plural quantification effects was a consequence of the relative simplicity of the grammatical constructions in our sentences, whereas the null results for plural quantification obtained by Eberhard may have resulted from the embeddedness of the quantified phrase, even if the basic representations or processes underlying agreement in comprehension and production are isomorphic. However, there are sufficient reasons to suspect that this difference in syntactic complexity was not responsible for the comprehension/production differences that we observed.

Notably, the MM model does not predict qualitatively different effects of quantification in simple and complex NP contexts. That is, although much of the data for marking and morphing comes from attraction experiments, MM is not a model of attraction—it is a model of agreement more generally. Attraction effects serve as a window into subtle information sources that impact agreement representation, since performance on agreement in non-attraction contexts would otherwise be near ceiling. Within the model, information from all nouns within the sentence can potentially contribute to the representation of the subject NP's number feature in similar ways. The crucial difference in the relative quantitative contributions to that representation from embedded versus nonembedded nouns is that nouns syntactically more distant from the root NP node should show a correspondingly smaller effect than those closer to the root NP node. This is modeled within MM as a weighting parameter for each noun, with weights decreasing as a function of increasing syntactic distance in order to simulate spreading activation decay over structural hierarchy. The model does not, however, propose any qualitative difference in how quantification or other forms of markedness contribute to the number representations for head versus embedded nouns. That is, any effects of quantification and markedness on number representation would be qualitatively similar for embedded nouns and head nouns (in either simple or complex NPs), though the quantitative influence from embedded nouns might be correspondingly smaller than the influence from head nouns.

Evidence in support of this qualitative similarity comes from a comparison of Experiments 1 and 2 reported by Eberhard (1997). In Experiment 1, Eberhard showed that singular quantification of a singular head noun reduced the influence of embedded plural attractors (e.g., *The/One key to the*

cabinets . . .); Experiment 2 showed that singular quantification of an embedded noun increased attraction rates with plural head nouns (e.g., *The keys to the/one cabinet* . . .). That is, adding a marked singular feature to an otherwise unmarked (default) singular NP via quantification caused qualitatively similar effects on both head and attractor nouns. Thus, assuming the plural-markedness account, MM predicts that plural quantification should not affect the number representation of either head or attractor nouns in production, since they are already fully marked for plurality.

This raises the additional question of whether the null results for embedded plural quantification in Eberhard's study might have resulted from a lack of sensitivity to quantificational effects due to structural decay. Two empirical results argue against this. First, the lack of an effect of plural quantification in Eberhard's (1997) results cannot be reduced to the fact that the quantified NP was embedded, because her results showed robust effects of singularly quantified NPs in embedded positions. Because the structural distances were identical in the singularly and plurally quantified conditions, the amounts of decay in the two conditions should have been identical in terms of the MM model. Second, if plural quantification were to similarly affect comprehension and production, and Eberhard's null result was simply due to decay (and our positive result due to lack of decay), one would expect to see a trend toward greater attraction in the plurally quantified than in the unmarked condition in the production data, even if it did not reach statistical significance. However, this is the opposite of what was found. Eberhard reported a numerical decrease in the proportion of agreement errors produced and a significant increase in the number of correct singular verbs produced following quantified versus unquantified plural attractor nouns.

Taking all of these points together leads us to the conclusion that the ways quantification is used as a cue to agreement in comprehension and production are nonidentical. We hypothesize that the difference between the way cues to agreement are used during production and comprehension results from the basic nature of the two tasks. During production, the planning process allows for numerous information sources to be simultaneously active (e.g., the notional number of the reference set and the grammatical number(s) of the noun(s) in the subject NP being formulated; cf. Gillespie & Pearlmutter, 2011), leading to a strong influence from the global representation of subject NP number, in line with the MM model. In comprehension, conceptual and morphosyntactic number information about a to-beconveyed message is not always available to the comprehender during first-pass comprehension, leading to an increased role for predictive processes and retrievals when prediction fails (see Tanner et al., 2014, for more detailed discussion), at least during initial processing. However, some evidence now suggests that global number representations (as conceived by the MM model) can be important in very late



measures of comprehension (Patson & Husband, 2015; see also Kreiner, Garrod, & Sturt, 2013).

Finally, given the divergence between the effects of plural quantification observed here and the lack of effects observed in production research, our results provide an important caveat to recent proposals highlighting the overlap between comprehension and production processes at representational, cognitive, and neural levels (e.g., Dell & Chang, 2014; Gennari & MacDonald, 2009; MacDonald, 2013; Pickering & Garrod, 2013; Segaert, Menenti, Weber, Petersson, & Hagoort, 2012; Silbert, Honey, Simony, Poeppel, & Hasson, 2014). Although it is clear that language comprehension and production share many underlying mechanisms, and that the systems supporting the two tasks interact very strongly, our present results support Tanner et al.'s (2014) proposal that there are some nontrivial divergences in the cues used during the two tasks. Taken with the findings of divergent attraction interference effects between comprehension and production reported by Tanner et al., and the divergent effects of quantification reported here, we argue that agreement in production is influenced by numerous information sources, with a large contribution from representational information about subject numerosity. On the other hand, the present data suggest a strong role for cue-based effects in comprehension, in the form of morphosyntactic predictions, which can operate in tandem with memory retrieval operations.

Acknowledgments This research was supported in part by Grant Number BCS-1431324 from the National Science Foundation, and by University of Illinois Campus Research Board Grant Number RB14158. We thank Kailen Shantz, Chase Krebs, Yisi Liu, Jack Dempsey, and Fiona Weingartner for assistance in data collection. We also thank two anonymous reviewers for helpful comments on an earlier version of the manuscript, and Laurel Brehm for important discussions about the issues addressed in this article.

References

- Badecker, W., & Kuminiak, F. (2007). Morphology, agreement and working memory retrieval in sentence production: Evidence from gender and case in Slovak. *Journal of Memory and Language*, 56, 65–85. doi:10.1016/j.jml.2006.08.004
- Bock, K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology*, 23, 45–97.
- Bock, K., Carreiras, M., & Meseguer, E. (2012). Number meaning and number grammar in English and Spanish. *Journal of Memory and Language*, 66, 17–37. doi:10.1016/j.jml.2011.07.008
- Brehm, L., & Bock, K. (2013). What counts in grammatical number agreement? *Cognition*, 128, 149–169. doi:10.1016/j.cognition. 2013.03.009
- Brothers, T., Swaab, T. Y., & Traxler, M. J. (2015). Effects of prediction and contextual support on lexical processing: Prediction takes precedence. *Cognition*, 136, 135–149. doi:10.1016/j.cognition.2014. 10.017
- Caballero, G., & Kapatsinski, V. (in press). Perceptual functionality of morphological redundancy in Choguita Rarámuri (Tarahumara).

- Language, Cognition and Neuroscience. doi:10.1080/23273798. 2014.940983
- Chomsky, N. (2001). Derivation by phase. In M. Kenstowicz (Ed.), *Ken Hale: A life in language* (pp. 1–52). Cambridge, MA: MIT Press.
- Corbett, G. G. (2000). Number. Cambridge, UK: Cambridge University Press
- Dell, G. S., & Chang, F. (2014). The P-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society*, 369, 20120394. doi:10.1098/rstb.2012.0394
- 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
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134,* 9–21. doi:10.1016/j.jneumeth.2003.10.009
- Dillon, B., Mishler, A., Sloggett, S., & Phillips, C. (2013). Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language*, 69, 85– 103. doi:10.1016/j.jml.2013.04.003
- Eberhard, K. M. (1997). The marked effect of number on subject–verb agreement. *Journal of Memory and Language, 36,* 147–164. doi:10. 1006/jmla.1996.2484
- Eberhard, K. M. (1999). The accessibility of conceptual number to the processes of subject–verb agreement in English. *Journal of Memory and Language*, 41, 560–578.
- Eberhard, K. M., Cutting, J. C., & Bock, K. (2005). Making syntax of sense: Number agreement in sentence production. *Psychological Review*, 112, 531–559. doi:10.1037/0033-295X.112.3.531
- Fedorenko, E., Gibson, E., & Rohde, D. (2006). The nature of working memory capacity in sentence comprehension: Evidence against domain-specific working memory resources. *Journal of Memory* and *Language*, 54, 541–553. doi:10.1016/j.jml.2005.12.006
- Foote, R., & Bock, K. (2012). The role of morphology in subject–verb number agreement: A comparison of Mexican and Dominican Spanish. *Language and Cognitive Processes*, 27, 429–461. doi:10. 1080/01690965.2010.550166
- Gennari, S. P., & MacDonald, M. C. (2009). Linking production and comprehension processes: The case of relative clauses. *Cognition*, 111, 1–23. doi:10.1016/j.cognition.2008.12.006
- Gillespie, M., & Pearlmutter, N. J. (2011). Hierarchy and scope of planning in subject–verb agreement production. *Cognition*, 118, 377–397. doi:10.1016/j.cognition.2010.10.008
- Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 27, 1411–1423. doi:10.1037/ 0278-7393.27.6.1411
- Gordon, P. C., Hendrick, R., & Levine, W. H. (2002). Memory-load interference in syntactic processing. *Psychological Science*, 13, 425–430.
- Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, 8, 439–483. doi:10.1080/01690969308407585
- Harris, A. C., & Samuel, A. G. (2011). Perception of exuberant exponence in Batsbi: Functional or incidental? *Language*, 87, 447–469. doi:10.1353/lan.2011.0068
- Hawkins, J. A. (2011). Processing efficiency and complexity in typological patterns. In J. J. Song (Ed.), *The Oxford hand-book of language typology* (pp. 206–226). Oxford, UK: Oxford University Press.
- Humphreys, K. R., & Bock, K. (2005). Notional number agreement in English. *Psychonomic Bulletin & Review*, 12, 689–695. doi:10. 3758/BF03196759



- Jasper, H. H. (1958). Report of the Committee on Methods of Clinical Examination in Electroencephalography: 1957. Electroencephalography and Clinical Neurophysiology, 10, 370-375. doi:10.1016/0013-4694(58)90053-1
- Kaan, E., & Swaab, T. Y. (2003). Repair, revision and complexity in syntactic analysis: An electrophysiological differentiation. *Journal* of Cognitive Neuroscience, 15, 98–110.
- Kail, M. (1989). Cue validity, cue cost, and processing types in sentence comprehension in French and Spanish. In B. MacWhinney & E. A. Bates (Eds.), *The crosslinguistic study of sentence processing* (pp. 77–117). Cambridge, UK: Cambridge University Press.
- Kreiner, H., Garrod, S., & Sturt, P. (2013). Number agreement in sentence comprehension: The relationship between grammatical and conceptual factors. *Language and Cognitive Processes*, 28, 829–874. doi: 10.1080/01690965.2012.667567
- Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, 29, 375–419. doi:10.1207/s15516709cog0000 25
- Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, 10, 447–454. doi:10.1016/j.tics.2006.08.007
- Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. Frontiers in Human Neuroscience, 8, 213. doi:10.3389/fnhum.2014.00213
- MacDonald, M. C. (2013). How language production shapes language form and comprehension. Frontiers in Psychology, 4(226). doi: 10.3389/fpsyg.2013.00226
- McElree, B. (2006). Accessing recent events. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 46, pp. 155–200). San Diego, CA: Academic Press.
- McElree, B., Foraker, S., & Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language*, 48, 67–91. doi:10.1016/S0749-596X(02)00515-6
- McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: Minimal instruction produces rapid change. *Nature Neuroscience*, 7, 703–704.
- Molinaro, N., Barber, H., & Carreiras, M. (2011). Grammatical agreement processing in reading: ERP findings and future directions. Cortex, 47, 908–930. doi:10.1016/j.cortex.2011.02.019
- Montgomery, J. W. (2000). Verbal working memory and sentence comprehension in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 43, 293–308.
- Nevins, A., Dillon, B., Malhotra, S., & Phillips, C. (2007). The role of feature-number and feature-type in processing Hindi verb agreement violations. *Brain Research*, 1164, 81–94.
- Nicol, J. L., Forster, K. I., & Veres, C. (1997). Subject–verb agreement processes in comprehension. *Journal of Memory and Language*, 36, 569–587. doi:10.1006/jmla.1996.2497
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97–113. doi:10.1016/0028-3932(71)90067-4
- Osterhout, L., & Mobley, L. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, 34, 739–773.
- Osterhout, L., Holcomb, P. J., & Swinney, D. A. (1994). Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 28,* 786–803. doi: 10.1037/0278-7393.20.4.786

- Patson, N. D., & Husband, E. M. (2015). Misinterpretations in agreement and agreement attraction. *Quarterly Journal of Experimental Psychology*, 1–22. doi:10.1080/17470218.2014.992445
- Pearlmutter, N. J., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language*, 41, 427–456. doi:10.1006/jmla.1999.2653
- Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, 36, 329–347. doi:10.1017/S0140525X12001495
- Segaert, K., Menenti, L., Weber, K., Petersson, K. M., & Hagoort, P. (2012). Shared syntax in language production and language comprehension—An fMRI study. *Cerebral Cortex*, 22, 1662–1670. doi: 10.1093/cercor/bhr249
- Severens, E., Jansma, B. M., & Hartsuiker, R. J. (2008). Morphophonological influences on the comprehension of subject–verb agreement: An ERP study. *Brain Research*, 1228, 135–144. doi:10.1016/j.brainres. 2008.05.092
- Silbert, L. J., Honey, C. J., Simony, E., Poeppel, D., & Hasson, U. (2014). Coupled neural systems underlie the production and comprehension of naturalistic narrative speech. *Proceedings of the National Academy* of Sciences, 111, E4687–E4696. doi:10.1073/pnas.1323812111
- Staub, A. (2009). On the interpretation of the number attraction effect: Response time evidence. *Journal of Memory and Language*, 60, 308–327. doi:10.1016/j.jml.2008.11.002
- Tanner, D., Nicol, J., & Brehm, L. (2014). The time-course of feature interference in agreement comprehension: Multiple mechanisms and asymmetrical attraction. *Journal of Memory and Language*, 76, 195–215. doi:10.1016/j.jml.2014.07.003
- Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings* of the National Academy of Sciences, 104, 12530–1255. doi:10. 1073/pnas.0609927104
- Thornton, R., & MacDonald, M. C. (2003). Plausibility and grammatical agreement. *Journal of Memory and Language*, 48, 740–759.
- Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar—An eventrelated potential investigation. Studies in Second Language Acquisition, 27, 173–204.
- Van Dyke, J. A. (2007). Interference effects from grammatically unavailable constituents during sentence processing. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 33*, 407–430. doi:10. 1037/0278-7393.33.2.407
- Van Dyke, J. A., & Johns, C. L. (2012). Memory interference as a determinant of language comprehension. *Language and Linguistics Compass*, 6, 193–211. doi:10.1002/lnc3.330
- Van Dyke, J. A., & McElree, B. (2011). Cue-dependent interference in comprehension. *Journal of Memory and Language*, 65, 247–263.
- VanPatten, B. (2004). Processing instruction: Theory, research, and commentary. Mahwah, NJ: Erlbaum.
- Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory* and Language, 61, 206–237. doi:10.1016/j.jml.2009.04.002
- Wlotko, E. W., & Federmeier, K. D. (2012). So that's what you meant! Event-related potentials reveal multiple aspects of context use during construction of message-level meaning. *NeuroImage*, 62, 356–366. doi:10.1016/j.neuroimage.2012.04.054
- Zhang, J., & Mueller, S. T. (2005). A note on ROC analysis and non-parametric estimate of sensitivity. *Psychometrika*, 70, 203–212. doi: 10.1007/s11336-003-1119-8

