



ELSEVIER

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

The time-course of feature interference in agreement comprehension: Multiple mechanisms and asymmetrical attraction

Darren Tanner^{a,*}, Janet Nicol^b, Laurel Brehm^c^a Department of Linguistics, Beckman Institute for Advanced Science and Technology, Neuroscience Program, University of Illinois at Urbana-Champaign, United States^b Departments of Linguistics and Psychology, Program in Cognitive Science, University of Arizona, United States^c Department of Psychology, University of Illinois at Urbana-Champaign, United States

ARTICLE INFO

Article history:

Received 22 March 2014

revision received 13 July 2014

Available online 7 August 2014

Keywords:

Agreement

Morphosyntactic processing

P600

Retrieval interference

Language comprehension

ABSTRACT

Attraction interference in language comprehension and production may be as a result of common or different processes. In the present paper, we investigate attraction interference during language comprehension, focusing on the contexts in which interference arises and the time-course of these effects. Using evidence from event-related brain potentials (ERPs) and sentence judgment times, we show that agreement attraction in comprehension is best explained as morphosyntactic interference during memory retrieval. This stands in contrast to attraction as involving the representation of the subject NP's root-node number feature, which is a strong contributor to attraction in production. We thus argue that the cognitive antecedents of agreement attraction in comprehension are non-identical with those of attraction in production, and moreover, that attraction in comprehension is primarily a consequence of similarity-based interference in cue-based memory retrieval processes. We suggest that mechanisms responsible for attraction during language comprehension are a subset of those involved in language production.

© 2014 Elsevier Inc. All rights reserved.

Introduction

Successful language processing requires the integration of grammatical information across linguistic units within a sentence. One grammatical dependency that must be computed in many languages is subject–verb agreement. Agreement phenomena are of particular interest because they present some interesting paradoxes for language processing. For example, while native English speakers tend to ignore agreement information in determining participant roles in off-line language comprehension tasks

(MacWhinney, Bates, & Kliegl, 1984), violations of subject–verb agreement readily trigger processing difficulty during real-time comprehension (Kreiner, Garrod, & Sturt, 2013; Osterhout & Mobley, 1995; Pearlmutter, Garnsey, & Bock, 1999). Moreover, although English agreement morphology is relatively impoverished and the rules of English agreement are relatively simple, both producers and perceivers are known to make errors in verb number marking. These errors are especially likely to occur in contexts where a sentence's subject noun phrase (NP) contains two nouns that mismatch in number features (e.g., *The key to the cabinets...*). In sentence production tasks, it has been shown that such cases lead to an increased number of errors in number marking on the verb (**The key to the cabinets are...*), where the verb erroneously agrees with the “local”

* Corresponding author. Address: Department of Linguistics, University of Illinois at Urbana-Champaign, 4080 Foreign Languages Building, MC-168, 707 S. Mathews Ave., Urbana, IL 61810, United States.

E-mail address: dstanner@gmail.com (D. Tanner).

(linearly closer) plural noun (e.g., *cabinets*; henceforth ‘attractor’ noun) rather than singular head noun of the subject NP (e.g., *key*).

Importantly, these errors, known as ‘attraction’ errors, can reliably be elicited in laboratory experiments focused on language production (Bock & Cutting, 1992; Bock & Eberhard, 1993; Bock & Miller, 1991; Eberhard, 1999; Gillespie & Pearlmutter, 2011; Hartsuiker, Schriefers, Bock, & Kikstra, 2003; Solomon & Pearlmutter, 2004; Staub, 2009, 2010; Vigliocco, Butterworth, & Garrett, 1996; Vigliocco, Butterworth, & Semenza, 1995; Vigliocco & Nicol, 1998), and later research found that plural attractors can interfere with the normal processing of verb agreement during language comprehension as well (e.g., Kaan, 2002; Nicol, Forster, & Veres, 1997; Pearlmutter, 2000; Pearlmutter et al., 1999; Thornton & MacDonald, 2003; Wagers, Lau, & Phillips, 2009). Since the earlier production-based studies focused on contexts in which errors arise, many comprehension studies also included error conditions (i.e., subject–verb agreement violations), because contrasting the processing of agreement errors in comprehension and production can yield useful information about the mechanisms supporting grammatical dependencies like agreement across the two task types. Studying the conditions under which attraction errors and attraction-based interference arise can inform us not only about the types of linguistic cues that enter into the computation of agreement dependencies, but also the nature of the mechanisms responsible for these computations, grammatical encoding, and the processing of long distance dependencies more generally.

Although agreement in comprehension has received considerably less attention than in production, the findings of attraction effects in both modalities suggests that similar information sources are used when establishing agreement dependencies. An important question is whether the general susceptibility to attraction in comprehension and production also points to shared cognitive, computational, or linguistic mechanisms used to establish agreement dependencies when speaking and reading, or whether attraction interference has different sources depending on the modality. The prevailing assumption in much psycholinguistic work is that the presence of attraction in both modalities indicates that agreement has largely overlapping cognitive or linguistic bases in production and comprehension (Badecker & Kuminiak, 2007; Hartsuiker, 2006; Nicol et al., 1997; Pearlmutter et al., 1999; Severens, Jansma, & Hartsuiker, 2008; Thornton & MacDonald, 2003; Wagers et al., 2009; see also Kreiner et al., 2013), though there remains debate about the exact nature of the processes giving rise to interference effects.

A unified, modality-neutral account is an attractive possibility, not least because it appeals to notions of parsimony: similar patterns of interference in comprehension and production should be explained by assuming a common cognitive mechanism or linguistic representation. This meshes with recent studies that find similarities in syntactic processing in comprehension and production, using behavioral methods (e.g., syntactic priming, Bock, Dell, Chang, & Onishi, 2007), and neuroimaging (e.g., fMRI, Segaert, Menenti, Weber, Petersson, & Hagoort, 2012).

Furthermore, all of this aligns with recent models of language processing which propose an extremely tight link between language comprehension and production mechanisms (e.g., Gennari & MacDonald, 2009; MacDonald, 2013; Pickering & Garrod, 2013).

Despite this, there are fundamental differences in comprehension and production, and the two necessarily use a set of processing systems that are at least partly distinct. Production begins with abstract conceptual messages; comprehension begins with sensory input (auditory, visual, or even tactile). It is therefore possible that a differing set of underlying processes may give rise to only superficially similar profiles of interference. The goal of the present paper is to investigate attraction interference during language comprehension, and in the process, gain a more detailed understanding of the relationship between the processes underlying the formation of agreement dependencies in language production and comprehension. We focus in particular on the contexts in which interference arises, as well as the time course of the interference effects. In doing so, we use a methodology which has not been extensively used to study agreement attraction, namely event-related brain potentials (ERPs).

Mechanisms of agreement and interference symmetries

Despite early findings of similar patterns of interference in production and comprehension, more recent findings on attraction suggest that the two domains may differ with regard to the contexts in which attraction interference arises. This suggests that the assumption of a shared mechanism needs further scrutiny. In particular, these conflicting findings point to a larger debate about the cause of attraction itself.

One family of theories holds that attraction arises because the mental representation of the number feature of the subject NP is faulty or ambiguous (e.g., Eberhard, Cutting, & Bock, 2005; Franck, Vigliocco, & Nicol, 2002; Nicol et al., 1997; Staub, 2009; Staub, 2010; Vigliocco & Nicol, 1998). On this account, attraction arises because the abstract plural feature of an attractor noun can erroneously spread throughout the syntactic representation of the subject NP, either through upwards feature percolation (Franck et al., 2002) or via spreading activation through a hierarchically structured syntactic tree (Eberhard et al., 2005; Hartsuiker, Antón-Méndez, & van Zee, 2001). It should be noted that these theories were initially developed to account for production data; however, these ideas have also been applied to comprehension (e.g., Nicol et al., 1997; Pearlmutter et al., 1999), and we do so here. Others have argued that attraction effects can have a representational locus, but from slightly different perspectives. For example, some argue that attraction arises from stages of a syntactic derivation, as conceptualized within Minimalist syntactic theory (e.g., Franck, Lassi, Frauenfelder, & Rizzi, 2006), or from statistically-based representations formed through language experience (e.g., Haskell, Thornton, & MacDonald, 2010).

A second class of approaches argues that attraction effects arise because the processing mechanisms

responsible for valuing or checking verb agreement may temporarily misidentify the correct controller of agreement (Badecker & Kuminiak, 2007; Dillon, Mishler, Sloggett, & Phillips, 2013; Wagers et al., 2009; see also Martin, Nieuwland, & Carreiras, 2012). This account, couched in a content-addressable working memory architecture, does not place the cause of interference on a globally faulty representation of the subject's number feature. Instead, it holds that retrieval of the agreement controller is subject to similarity-based interference from other items in memory (see also Gordon, Hendrick, & Johnson, 2001; Gordon, Hendrick, & Levine, 2002; Van Dyke & McElree, 2006). On this account, processing verb inflections requires retrieval of possible agreement controllers represented in memory. When the subject NP contains two noun phrases, they may potentially compete for control of verb agreement, even when only one of them is syntactically licensed as an agreement controller. When the two nouns mismatch in number, competition can result in a potential mis-identification of the agreement controller.

Many aspects of agreement production can be accurately modeled within the first approach, and Eberhard and colleagues (2005) have formalized this account within a computational model known as Marking and Morphing (MM).¹ A central tenet of MM is that the number value of the subject NP is continuously valued. The model formalizes this as the $S(r)$ value, which reflects the relative singular and plural (SAP) value of the subject NP and is the "joint product or reconciliation of the initial marking of notional number and the number specifications of morphemes that compose the phrase" (p. 539). This means that, within this model, some root NP nodes can be 'more plural' than others, indicated by a higher $S(r)$ value. For example, when an NP contains a singular head and a plural attractor noun and/or refers to a notionally (semantically) plural or collective entity, the number representation of the overall subject NP will be more plural than when both nouns are singular and/or the referent of the NP is unambiguously singular. This number representation of the subject NP (established through number 'marking' in the model) then triggers number agreement on the verb (established through 'morphing' in the model) via a probabilistic process: the more plural the overall subject NP's value, the greater the probability that the agreement process will select a plural verb. Within this model, the overall $S(r)$ value of the subject NP is the only parameter visible for valuation of verb number, and not the individual nouns within the NP. Thus, it is the relative SAP value of the whole subject NP that specifies verb number and gives rise to attraction effects.

Given a continuously valued number global representation, the MM architecture thus predicts that difficulty in processing should arise regardless of whether or not the correct verb is ultimately chosen in production, or in comprehension, whether or not the verb encountered is properly marked for number. That is, agreement attraction

should be *symmetrical* (Wagers et al., 2009).² In production, symmetry should give rise to processing difficulty for verbs following singular head and plural attractor nouns, even when the correct noun is ultimately chosen. Recent reaction time (RT) evidence squarely supports this prediction. When participants are asked to choose the correct verb (*is* or *are*) to continue a visually-presented sentence preamble either through a forced-choice button pushing task (Staub, 2009, 2010) or through elicited oral production (Brehm & Bock, 2013; Haskell & MacDonald, 2003), they are significantly slower following plural attractor nouns embedded in an otherwise singular NP than when the attractor is singular. This slowdown occurs even when they ultimately choose the correct verb. Staub (2009) argued that these findings supported the notion that the ambiguously-valued global number representation (e.g., a 'slightly plural' $S(r)$ value) gives rise to uncertainty about the correct verb form, which in turn leads to longer RTs. Thus, in addition to accurately predicting effects on attraction rates of notional number (e.g., Bock, Carreiras, & Meseguer, 2012; Brehm & Bock, 2013), morphological richness (e.g., Foote & Bock, 2012; Lorimor, Bock, Zalkind, Sheyman, & Beard, 2008), and other factors (see Bock & Middleton, 2011, for a recent review), the MM model also correctly predicts symmetrical attraction effects during language production.

Shared mechanisms and (a)symmetrical attraction

MM was designed to account solely for production data, and although MM correctly predicts a range of interference effects in language production, including findings of symmetrical attraction, it is less clear to what extent it could account for facts of agreement comprehension. Maintaining a unified account of attraction in comprehension and production would minimally require: (1) that the same linguistic factors that modulate attraction rates in production also modulate attraction in comprehension, and (2) that interference is similarly symmetrical in comprehension. With respect to the first criterion, although years of research have identified numerous linguistic factors modulating attraction in production at play in production only a few studies have investigated whether there may be parallel effects in comprehension. Some findings from comprehension have converged with those from production regarding the role of attractor-predicate plausibility, morphophonological ambiguity in number marking on head noun, and the role of semantics in number marking on verbs versus pronouns (Kreiner et al., 2013; Severens et al., 2008; Thornton & MacDonald, 2003), though findings regarding the role of hierarchical structure in agreement comprehension have been equivocal (cf. Nicol et al., 1997; Pearlmutter, 2000; see Gillespie & Pearlmutter,

¹ We focus here on the MM model. However, the other representational models mentioned above (Franck et al., 2006; Haskell et al., 2010) would seem to make similar predictions for the relationship between language production and comprehension, assuming the same representations are active in both tasks (e.g., MacDonald, 2013).

² Here we follow Wagers et al. (2009) and use the terms "symmetrical" and "asymmetrical" to refer to the contexts in which attraction interference occurs. Symmetrical attraction will occur in both grammatical and ungrammatical sentences, whereas asymmetrical attraction will occur in either grammatical or ungrammatical sentences. Note that this is a different usage from the "markedness asymmetry" sometimes discussed in the agreement attraction, where plural nouns are marked for number and can therefore give rise to interference, whereas singular nouns are unmarked for number and therefore do not give rise to interference (Eberhard, 1997).

2013, for evidence against hierarchical structural constraints in agreement production).

With respect to the second criterion for a shared mechanism, results are less clear. In the RT data provided by Staub (2009), the globally ambiguous number representation of the subject NP symmetrically impacted both correct and incorrect decisions about verb number during a forced choice task (see also Brehm & Bock, 2013; Haskell & MacDonald, 2003). Extending this logic to comprehension, this would predict that plural attractors embedded in PP modifiers should cause symmetrical interference, affecting the processing of both grammatical and ungrammatical verbs. Early studies investigating interference effects in comprehension supported this hypothesis, using maze-reading and sentence judgment tasks (Nicol et al., 1997), and using self-paced reading and eye-tracking (Pearlmutter et al., 1999; Thornton & MacDonald, 2003).

More recently, Wagers and colleagues (Wagers et al., 2009) have replicated the finding that plural attractors decrease ungrammaticality effects in self-paced reading, but questioned the finding that attraction also occurs in grammatical sentences. In a series of self-paced reading studies in English they showed the attraction effect in ungrammatical sentences is quite robust, but also argued that the previously-reported attraction effects in grammatical sentences may result from spillover effects from the increased complexity of the plural noun preceding the verb, and not actually from agreement processes themselves. Wagers and colleagues concluded that plural nouns, which are morphologically and conceptually more complex than singular nouns, incur processing costs which may spill over into subsequent reading regions, and that this spillover had erroneously been interpreted as an attraction effect in grammatical sentences in previous work.³ They concluded that attraction is *asymmetrical*: it impacts ungrammatical sentences, making them easier to process, but it does not make grammatical sentences more difficult to process (see also Dillon et al., 2013, for eye movement evidence of asymmetrical attraction from within relative clauses). Wagers and colleagues argue that their results are consistent with the second family of theories described above, and implicate cue-based working memory retrieval mechanisms in the establishment of agreement dependencies, where similarity-based interference from items represented in memory may give rise to processing difficulty during the processing of linguistic dependencies (Gordon et al., 2001, 2002; Lewis, Vasishth, & Van Dyke, 2006; Lewis & Vasishth, 2005; Martin & McElree, 2008; McElree, Foraker, & Dyer, 2003; McElree, 2000, 2006; Van Dyke, 2007; Van Dyke & McElree, 2011).

Both Wagers et al. and Dillon et al. argue that the asymmetry in attraction effects is predicted by cue-based

retrieval mechanisms in the following way. Upon encountering a verb in a context such as “The author of the speeches *are*...” the processing system deploys a search in working memory for possible agreement controllers. Because the plural attractor noun has a partial feature overlap with an ungrammatical plural verb, an agreement dependency can be established between the two items. In the case of grammatical verbs (e.g. “The author of the speeches *is*...”), a plural attractor noun does not overlap in number features with the verb and therefore is unlikely to cause interference. Moreover, it is additionally possible that retrieval mechanisms operate in tandem with predictive mechanisms, where retrievals are driven by prediction errors. In this case, retrievals would only be initiated when detected agreement features mismatch with predicted features. This would further predict attraction asymmetries, as no retrievals would be necessary in grammatical sentences.

While the evidence from Wagers and colleagues for attraction from within prepositional phrases is compelling, it is difficult to reconcile with Nicol and colleagues' (1997) findings of attraction in grammatical sentences, especially because Nicol et al. showed in a control experiment with uninflected verbs (Experiment 3) that the interference effect disappeared. It is additionally worth noting that the statistical procedure Wagers and colleagues used to control for possible spillover effects from the plural attractor noun may have removed any effects of attraction in grammatical sentences from their reaction time data. That is, by using previous words' reading times as covariates, processing difficulty at a grammatical verb following a plural noun (i.e., a true attraction effect) could be regressed out of the data, since the attraction effect would be of a similar magnitude and direction to the effect of the plural noun in the previous reading regions. Note, however, that in Experiments 2 and 3 Wagers and colleagues did not use statistical covariates and found no differences in reading times for grammatical sentences with singular and plural attractors in relative clauses with a non-intervening attractor (e.g., *The musician(s) who the reviewer praise(s)...*), where spillover from a plural attractor (*musicians*) is not a concern at the target verb (*praise(s)*). However, more recent findings reported by Staub (2010) suggest that interference from non-intervening attractors and interference from intervening (PP-embedded) attractors have a different cognitive source (that is, predication confusion versus true attraction). Thus, more research is needed here. Moreover, Wagers and colleagues maintain that the content-addressable retrieval mechanism they propose to explain asymmetrical attraction effects in comprehension can also be used to account for attraction effects in production (see also Badecker & Kuminiak, 2007). However, the asymmetry they report for comprehension is clearly at odds with more recent findings of symmetrical attraction in verb selection and sentence completion tasks (e.g., Brehm & Bock, 2013; Staub, 2009).

A further possibility is that attraction in grammatical sentences exists, but is exceedingly subtle, such that the self-paced reading task used by Wagers and colleagues was insufficiently sensitive to detect it (compare the eye tracking results of Pearlmutter et al., 1999, Experiment 2).

³ Note that there is a problem with the spillover explanation for Nicol et al.'s (1997) data for plural head cases. First, in their maze-task experiment (Exp. 1), which obtained RTs for verbs, RTs only differed for singular-head conditions, not plural head conditions (i.e. the verb “are” was read equally quickly after “The authors of the speeches...” and “The authors of the speech...”). Second, their whole-sentence judgment task also showed no RT differences for sentences containing plural heads: if a plural head was followed by a plural attractor, the mean RT was no greater than if a plural head was followed by a singular attractor.

The approach we take here, namely using electrophysiological measures, provides an important source of information regarding this possibility, as ERPs have been shown to detect processing effects that go unnoticed in behavioral measures (e.g., McLaughlin, Osterhout, & Kim, 2004; Thierry & Wu, 2007; Tokowicz & MacWhinney, 2005). If it is found that attraction is asymmetrical, in contrast to findings from production-oriented experiments, this would countervail the assumption that attraction interference effects in comprehension and production have isomorphic sources. Finding symmetrical attraction in comprehension would be consistent with, though would not necessarily entail, the hypothesis that attraction in the two tasks has a shared cause. Finally, a third possibility is that attraction arises in both grammatical and ungrammatical sentences but from different sources, as has recently been shown to be the case for attraction from within prepositional phrases and non-intervening relative clauses (Staub, 2010). ERPs' multidimensional outcome measures could therefore potentially distinguish whether attraction in grammatical and ungrammatical sentences might arise from qualitatively different cognitive mechanisms.

Electrophysiology and agreement attraction

ERPs are direct recordings of cortical electrical activity underlying processing and have proven to be extremely useful in studying language comprehension. Importantly for our purposes, ERPs have a measurement scale on the order of milliseconds, which is fast enough to capture comprehension processes as they unfold in time, they provide multiple dimensions of outcome measures (e.g., effect timing, polarity, and scalp distribution), and have shown to be highly sensitive to aspects of morphosyntactic processing.

Most relevant to the present study is the finding that violations of morphosyntactic constraints, including subject–verb agreement violations, typically elicit an enhanced positive-going wave beginning around 500 ms poststimulus that is most prominent over centroparietal scalp regions (the P600 effect: Friederici, Hahne, & Mecklinger, 1996; Kaan, Harris, Gibson, & Holcomb, 2000; Molinaro, Barber, & Carreiras, 2011; Molinaro, Vespignani, Zamparelli, & Job, 2011; Nevins, Dillon, Malhotra, & Phillips, 2007; Osterhout & Holcomb, 1992; Osterhout & Mobley, 1995; Rossi, Gugler, Hahne, & Friederici, 2005). Some studies of morphosyntactic processing have additionally elicited a negativity preceding the P600, which is sometimes strongest over left anterior portions of the scalp (the left anterior negativity, or LAN: e.g., Barber & Carreiras, 2005; Friederici et al., 1996; Hahne & Friederici, 1999; Osterhout & Holcomb, 1992; Osterhout & Mobley, 1995), though the LAN is highly inconsistent across studies (e.g., Allen, Badecker, & Osterhout, 2003; Frenck-Mestre, Osterhout, McLaughlin, & Foucart, 2008; Hagoort, 2003; Hagoort & Brown, 1999, 2000; Kim & Osterhout, 2005; Nevins et al., 2007; Osterhout, Allen, McLaughlin, & Inoue, 2002; Osterhout & Mobley, 1995; Tanner & Van Hell, 2014). We therefore focus on the P600 effect. The onset timing of the P600 has been shown to be sensitive to the relative ease in diagnosing a syntactic anomaly (Friederici, Mecklinger,

Spencer, Steinhauer, & Donchin, 2001; McKinnon & Osterhout, 1996), and the amplitude of the P600 effect can be modulated by the relative perception of ungrammaticality or degree of difficulty in grammatical processing. Larger P600 magnitudes are parametrically associated with less resolvable and more salient grammatical violations, as well as increased processing difficulty in grammatical but difficult-to-process long-distance and verb–argument dependences (e.g., Kaan, 2002; Kaan & Swaab, 2003; Kaan et al., 2000; Nevins et al., 2007; Osterhout, Holcomb, & Swinney, 1994; Xiang, Dillon, & Phillips, 2009). Thus, the P600 effect onset and magnitude can reliably be used as a continuous index of syntactic processing difficulty at the brain level for both grammatical and ungrammatical sentences.

Only a few studies have used ERPs to investigate attraction-like phenomena in comprehension (see Xiang et al., 2009, for additional ERP evidence regarding intrusive licensing of negative polarity items), and as with the behavioral studies reviewed above, the results have been somewhat inconclusive regarding attraction asymmetries. Kaan (2002) investigated the impact of number interference on verb agreement in Dutch subject–object–verb sentences. Kaan's main findings were that P600 amplitudes to disagreeing verbs in clauses where the subject and object NPs' number features were both plural or mismatched were significantly smaller than in cases where both NPs were singular, suggesting that verb agreement violations in the singular–singular condition were more salient than in the other conditions. Hints of asymmetry were also observed: in the singular subject/plural object condition, grammatical verbs elicited a positivity between 250 and 300 ms poststimulus, but ungrammatical verbs showed a more typical positivity in the later time windows. Kaan suggested that participants may have initially perceived the grammatical sentences as ungrammatical given the number mismatch in the local string (i.e., plural attractor noun–singular verb), but that this misanalysis was quickly corrected.

Severens et al. (2008) included an agreement attraction manipulation in a study on the role of morphophonology in Dutch agreement processing. They found that, whereas disagreeing verbs following complex subjects containing PP modifiers with mismatching number features (i.e., singular head/plural attractor) showed a P600, disagreeing verbs following complex subjects with matching number features (i.e., singular head/singular attractor) showed an unexpected N400-like negativity. The authors attributed this unexpected N400 to the differences in the depth of processing needed in the match and mismatch conditions. They speculate that participants may use 'shallow processing' strategies in the relatively simpler match condition, and may use a deeper processing strategy in the more complex mismatch condition (Kuperberg, 2007; see also Ferreira, Bailey, & Ferraro, 2002; Ferreira, 2003). Importantly, neither the Kaan (2002) nor the Severens et al. (2008) studies directly compared brain responses for the singular and plural attractor conditions in grammatical sentences, so they provide no electrophysiological evidence regarding the symmetry of attraction in comprehension.

More recent work has provided tentative support for asymmetrical attraction effects. However, some of the findings have implicated ERP effects other than the canonical P600. Martin and colleagues (Martin et al., 2012) investigated gender attraction effects in Spanish NP ellipsis using ERPs. Sentences contained two NPs, which either matched or mismatched in gender, as well as a gapped NP indicated by the determiner *otro/a*, which either correctly agreed with the first NP, or incorrectly agreed with the second NP (e.g., *Marta se compró la camiseta que estaba al lado de la falda/del vestido y Miren cogió otra/otro [...] para salir de fiesta*, “Marta bought the t-shirt_{Fem} that was next to the skirt_{Fem}/dress_{Masc} and Miren took another_{Fem/Masc} to go to the party”). Results showed that gender mismatches at the gapping site elicited a sustained, widespread negativity, rather than a P600, that largely impacted *grammatical* rather than ungrammatical sentences. Symmetrical attraction emerged only in a late time window.

Shen, Staub, and Sanders (2013) recorded ERPs to agreement violations, including attraction violations, in auditorily-presented spoken English discourses. They found a typical LAN-P600 pattern of effects for simple agreement violations (i.e., those without an intervening attractor noun). Relevant to the current discussion, in complex NPs they found no differences when comparing ERP responses to grammatical verbs following singular and plural attractors, suggesting no increased processing difficulty in these contexts, in line with the asymmetrical attraction findings reported by Wagers and colleagues. However, other effects of number and grammaticality they reported were atypical of most ERP studies of morphosyntactic processing (either auditory or visual), as they largely involved posterior negativities to ungrammaticalities. Given the differences in materials and procedure relative to other studies of agreement attraction, it is unclear how well these results might generalize to the broader agreement attraction literature.

Here we present data from three experiments (two ERP, one behavioral) on agreement processing in the comprehension of English sentences. Given the conflicting findings regarding the contexts in which attraction interference exerts itself in comprehension, our primary goals were to identify the contexts in which attraction interference occurs in English agreement as well as the time course of the interference. Establishing asymmetric effects of attraction during comprehension would indicate that, counter to some current assumptions, attraction has differing sources in comprehension and production.

Experiment 1

In Experiment 1 we investigated classic agreement attraction effects in English using ERPs. We tested interference effects from plural attractor nouns in NPs with singular head nouns, as singular attractor nouns with plural head nouns have been shown to exert very little interference in both production and comprehension (Bock & Miller, 1991; Eberhard, 1997; Nicol et al., 1997; Pearlmutter et al., 1999; Thornton & MacDonald, 2003).

In line with previous research, we expected that subject-verb agreement violations would elicit robust P600 (and possibly LAN) effects. Of particular importance to the theoretical questions posed here is whether interference would be apparent primarily in grammatical sentences (Martin et al., 2012), primarily in ungrammatical sentences (Wagers et al., 2009), or both (Pearlmutter et al., 1999). As P600 effects can reflect processing difficulty in both grammatical and ungrammatical sentences, attraction interference should manifest itself in ungrammatical sentences as a reduction in P600 amplitude, whereas attraction in grammatical sentences should manifest itself as an increase in P600 amplitude, relative to grammatical sentences with singular attractors.

Method

Participants

Our participants included 24 native English speakers. Participants were all strongly right-handed, as assessed by an abridged version of the Edinburgh Handedness Inventory (Oldfield, 1971), and had normal or corrected-to-normal vision. Data from one participant were excluded because a language history questionnaire revealed significant early exposure to a non-English language with agreement patterns substantially different from English (Mandarin Chinese); all other participants reported no significant knowledge of or exposure to languages other than English and were functionally monolingual. Two additional participants were excluded due to an excessive amount of blink artifact in the raw EEG. This left 21 participants in the final analyses (13 male; mean age: 21.2 years, range: 18–32). Participants either volunteered or received course credit for taking part.

Materials

120 sentence quadruplets were constructed to serve as critical materials for this experiment in a 2×2 factorial design. The factor grammaticality (grammatical, ungrammatical) was crossed with attractor number (singular, plural), resulting in four versions of each sentence frame (see Table 1). Critical words were short auxiliary verbs (*is/are, was/were, has/have*) which either agreed with the singular head noun of the subject NP (grammatical verbs) or disagreed (ungrammatical verbs); the number of the attractor noun embedded within a PP subject modifier was also manipulated as either singular or plural. Some preambles were adapted from published stimulus materials investigating agreement attraction (Bock & Miller, 1991; Staub, 2009; Wagers et al., 2009); others were newly constructed. The four versions of each sentence were distributed across four experimental lists, such that each participant saw only one version of each sentence and there were 30 critical trials in each condition per list.

Each list also contained 120 filler sentences. Half of the filler sentences were grammatical and half contained a violation of determiner-noun agreement or numerical quantifier-noun agreement (e.g., *I read one/two good book last summer*). Each list thus contained 240 sentences, half of which contained an ungrammaticality. Each list was pseudo-randomized such that no more than three trials

Table 1
Example experimental sentences from Experiment 1.

| Grammaticality | Attractor number | Sentence |
|----------------|------------------|---|
| Grammatical | Singular | The chemist with the test tube <u>is</u> conducting an experiment |
| Ungrammatical | Singular | The chemist with the test tube <u>are</u> conducting an experiment |
| Grammatical | Plural | The chemist with the test tubes <u>is</u> conducting an experiment |
| Ungrammatical | Plural | The chemist with the test tubes <u>are</u> conducting an experiment |

Note: The critical word for ERP averaging is underlined.

from any single condition occurred in succession and no more than three grammatical or ungrammatical trials occurred in succession. Lists were divided into 3 blocks of 80 sentences each.

Procedure

Participants were tested in a single session lasting approximately eighty-five minutes (including about thirty minutes of experimental preparation). Each participant was randomly assigned to one of the stimulus lists and was seated in a comfortable recliner in front of a CRT monitor. Participants were instructed to relax and minimize movements while reading and to read each sentence as normally as possible. Each trial consisted of the following events: each sentence was preceded by a blank screen for 1000 ms, followed by a fixation cross, followed by a stimulus sentence, presented one word at a time. The fixation cross and each word appeared on the screen for 450 ms followed by a 200 ms blank screen between words. Sentence-ending words were followed by a full stop, followed by a “yes” or “no” prompt, instructing participants to give an acceptability judgment for the sentence. Participants were instructed to respond “yes” if they felt that it was a well-formed English sentence or “no” if they felt something was wrong with the sentence. The “yes” response hand was counterbalanced across participants. Participants were put under no time pressure to make their judgments.

Data acquisition and analysis

Continuous EEG was recorded from 19 tin electrodes attached to an elastic cap (Electro-cap International) in accordance with the 10–20 system (Jasper, 1958). Eye movements and blinks were monitored by two electrodes, one placed beneath the left eye and one placed to the right of the right eye. All electrodes were referenced to an electrode placed over the left mastoid and were amplified with a bandpass of 0.01–40 Hz (–3dB cutoff) by an SA Instruments bioamplifier system. EEG was recorded from an additional electrode placed on the right mastoid to identify if there were any experimental effects detectable over the mastoids; no such effects were found. Impedances at scalp and mastoid electrodes were held below 5 k Ω and below 15 k Ω at eye electrodes.

Continuous analog-to-digital conversion of the EEG and stimulus trigger codes was performed at a sampling

frequency of 200 Hz. ERPs, time-locked to the onset of the critical word, were averaged off-line for each participant at each electrode site in each condition. Grand average waveforms were created by averaging over participants. Trials characterized by eye blinks, excessive muscle artifact, or amplifier blocking were not included in the averages; 4.6% of trials overall were removed due to artifacts and the number of rejected trials did not differ across conditions.

Behavioral data was quantified both as proportion judged correctly in each of the four conditions in Table 1 and as d-prime scores for each of the singular and plural attractor conditions to test decreases in sensitivity to the grammaticality manipulation from attractor interference. Proportions were arcsine transformed and analyzed using a 2×2 repeated-measures ANOVA with factors grammaticality (grammatical, ungrammatical) and match (match, mismatch); d-prime data were analyzed using a one-factor repeated measures ANOVA to show whether overall sensitivity to the agreement violations decreased in the match versus mismatch conditions. ERP components of interest were quantified by computer as mean voltage within a window of activity. In accordance with previous literature and visual inspection of the data, the following time windows were chosen: 150–300 ms (P2), 300–500 ms (LAN/N400), and 500–800 ms (P600), relative to a 200 ms prestimulus baseline. Differences between conditions were analyzed using a repeated-measure ANOVA with two levels of grammaticality (grammatical, ungrammatical) and two levels of attractor number (singular, plural). Additionally, to investigate the scalp distribution of the relevant effects, data from midline (Fz, Cz, Pz), medial (right hemisphere: Fp2, F4, C4, P4, O2; left hemisphere: Fp1, F3, C3, P3, O1), and lateral (right hemisphere: F8, T8, P8; left hemisphere: F7, T7, P7) were treated separately. ANOVAs on midline electrodes included electrode as an additional within-subjects factor (3 levels), ANOVAs on medial-lateral electrodes included hemisphere (2 levels) and electrode pair (5 levels) as additional within-subjects factors, and ANOVAs over lateral-lateral electrodes included hemisphere (2 levels) and electrode pair (3 levels) as additional within-subjects factors. The Greenhouse–Geisser correction for inhomogeneity of variance was applied to all repeated measures with greater than one degree of freedom in the numerator. In such cases, the corrected *p*-value is reported.

Results

Behavioral results

Mean proportion of sentences judged correctly as acceptable or unacceptable and d-prime scores are presented in Table 2. A repeated measures ANOVA on arcsine-transformed proportions judged correctly showed no main effect of grammaticality ($p > .27$). There was a main effect of attractor number, $F(1,20) = 17.899$, $MSE = 0.016$, $p < .001$, and an interaction which did not reach significance, $F(1,20) = 2.454$, $p = .133$. Because the theoretical questions of interest crucially hinge on whether there is an asymmetry in attraction effects between grammatical and ungrammatical sentences, we computed planned

Table 2

Means and standard errors for behavioral judgments in Experiment 1. Proportion judged correctly and d-prime scores. Standard errors are reported in parentheses.

| | Grammatical | Ungrammatical | d-Prime |
|---------------|-------------|---------------|------------|
| Sg. Attractor | .92 (.02) | .90 (.04) | 3.23 (.23) |
| Pl. Attractor | .90 (.02) | .79 (.05) | 2.48 (.26) |

pair-wise tests comparing the effects of attractor number within both grammatical and ungrammatical sentences. These effects showed no reliable effect of attractor number within grammatical sentences, $p > .05$, but a reliable effect in ungrammatical sentences, $t(20) = 3.719$, $p = .001$. Analysis of d-prime scores showed that sensitivity to grammatical violations decreased significantly in the presence of a plural attractor, as there was a reliable effect of attractor number, $F(1,20) = 18.143$, $MSE = .326$, $p < .001$. In sum, behavioral accuracy showed a reliable attraction effect, as sensitivity to subject–verb agreement grammaticality was significantly poorer in the plural attractor condition. However, this effect was asymmetrical and driven primarily by a reduction in sensitivity to ungrammatical verbs following plural attractor nouns and not by a perception of ungrammaticality for grammatical verbs following plural attractors.

ERP results

Grand mean ERP waveforms comparing brain responses to grammatical and ungrammatical verbs in the singular attractor condition are presented in Fig. 1, and for the plural attractor condition in Fig. 2. Waveforms comparing all four conditions at central midline electrode Cz are presented in Fig. 3. In these and all subsequent waveforms, the general shapes of the waveforms were consistent with previous data using visually presented language stimuli (e.g., Osterhout & Holcomb, 1992). Visual inspection of the waveform showed that, relative to grammatical verbs, brain responses to ungrammatical verbs were more positive-going beginning around 450 ms poststimulus, though the magnitude of the positivity was much larger in the singular attractor condition than the plural attractor condition. In both the singular and plural attractor conditions there appeared to be a small left-central negativity elicited by ungrammatical verbs beginning around 300 ms, and in the singular attractor condition, ungrammatical verbs showed an additional early positivity between 150 and 300 ms poststimulus. Statistical analysis confirmed all of these observations.

In the 150–300 ms time window, there were no main effects of grammaticality or attractor number over any scalp region; however, there was a significant interaction between grammaticality and attractor number over midline electrodes, $F(1,20) = 6.614$, $MSE = 4.447$, $p = .018$, and which neared significance over lateral electrodes, $F(1,20) = 3.843$, $MSE = 3.350$, $p = .064$. Follow-up tests over midline sites showed that this interaction was driven by an increased positivity to ungrammatical verbs in the singular attractor condition, $F(1,20) = 11.776$, $MSE = 3.349$, $p = .003$, which was not present in the plural attractor condition ($F < 1$). In the 300–500 ms time window, the only reliable

effect of interest was a significant grammaticality by electrode by hemisphere interaction over lateral electrodes, $F(2,40) = 4.165$, $MSE = .538$, $p = .041$. This effect indicates a small negativity over left hemisphere sites and a small positivity over right hemisphere sites, particularly over electrodes T7/T8 (cf. Tanner & Van Hell, 2014).

Between 500 and 800 ms there was a large, widespread main effect of grammaticality, indicating a P600 effect to ungrammatical verbs [midline: $F(1,20) = 23.747$, $MSE = 14.339$, $p < .001$; medial: $F(1,20) = 16.017$, $MSE = 28.607$, $p < .001$; lateral: $F(1,20) = 12.454$, $MSE = 10.361$, $p = .002$], which was largest over posterior sites [grammaticality \times electrode interaction, midline: $F(2,40) = 9.095$, $MSE = 3.027$, $p = .004$; medial: $F(4, 80) = 12.127$, $MSE = 11.691$, $p < .001$; lateral: $F(2,40) = 14.634$, $MSE = 2.336$, $p < .001$], and which also was dominant over right posterior lateral sites [grammaticality \times electrode \times hemisphere interaction: $F(2,40) = 6.339$, $MSE = .758$, $p = .040$].

Importantly, there was also a grammaticality by attractor number interaction [midline: $F(1,20) = 14.081$, $MSE = 5.099$, $p = .001$; medial: $F(1,20) = 7.794$, $MSE = 13.576$, $p = .011$; lateral: $F(1,20) = 6.193$, $MSE = 7.141$, $p = .022$]. We investigated this interaction by computing pair-wise ANOVAs between all four conditions over midline electrodes, since this is where the interaction was strongest. This analysis showed that the effect of grammaticality was highly reliable in the singular attractor condition, $F(1,20) = 41.263$, $MSE = 8.785$, $p < .001$, but much less so in the plural attractor condition, $F(1,20) = 4.675$, $MSE = 10.652$, $p = .043$. A comparison of brain responses in the two ungrammatical conditions revealed a reliable main effect of attractor number, $F(1,20) = 10.216$, $MSE = 10.514$, $p = .005$, indicating that the positivity to ungrammatical verbs was significantly smaller following plural attractor nouns. However, when comparing brain responses in the two grammatical conditions, there were no detectable differences, $F(1,20) = .517$.

Discussion

The main results of Experiment 1 are as follows. Plural attractors showed a reliable interference effect in both participants' judgments of the sentences and in the pattern of brain responses elicited by verbs during reading. Importantly, the attraction effect was asymmetrical: plural attractors ameliorated effects of ungrammaticality, whereas they did not cause grammatical sentences to be perceived as correspondingly less grammatical. For behavioral judgments, this manifested itself as a reduction in accuracy for ungrammatical verbs following plural attractors, with no such reduction in accuracy for grammatical singular verbs following plural attractors. For participants' brain responses, effects of ungrammaticality emerged as early as the 150–300 ms time window, but only following singular attractors; effects of ungrammaticality following plural attractors were less reliable and did not appear until the 500–800 ms time window, suggesting these agreement violations were detected more slowly and processed less robustly in that condition. Indeed, the P600 for ungrammatical verbs was reliably smaller in the plural attractor than singular attractor condition. However, the plural

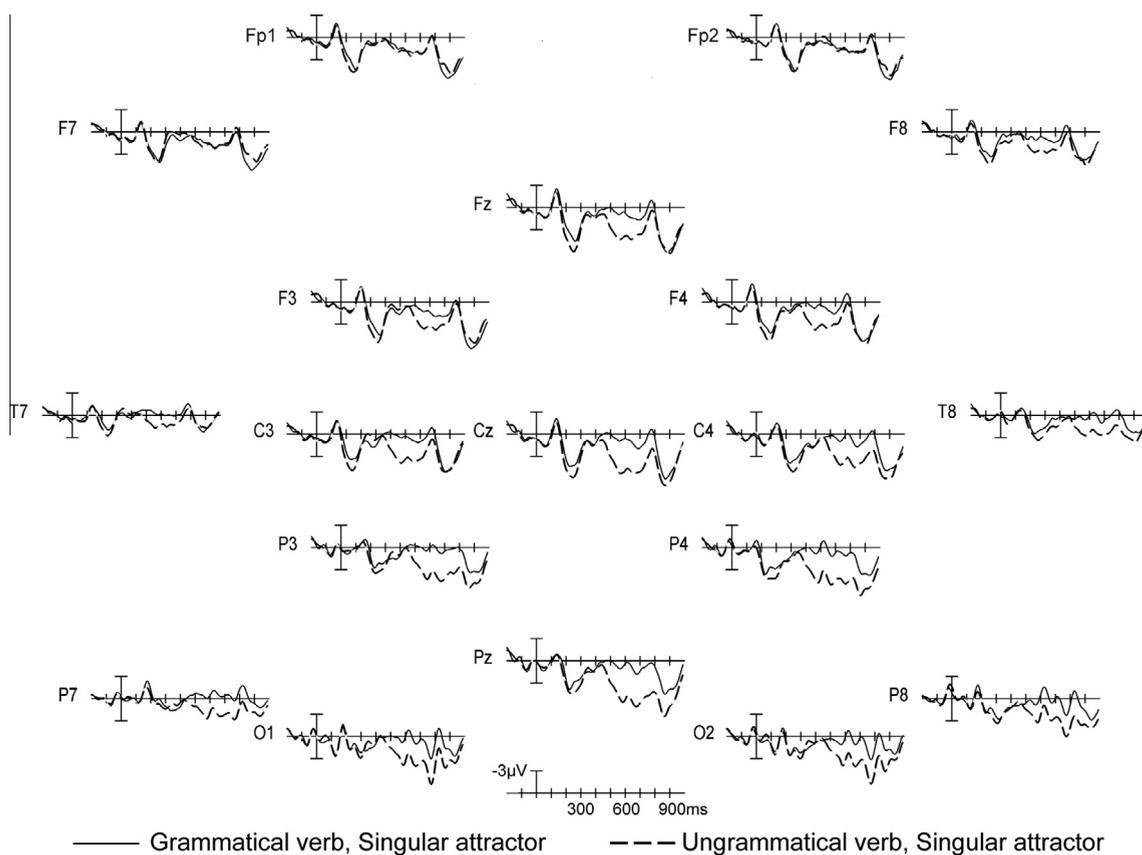


Fig. 1. Grand mean ERP waveforms from Experiment 1 for all 19 scalp electrodes in the grammatical (solid line) and ungrammatical (dashed line) singular attractor conditions. Onset of the verb is indicated by the vertical calibration bar; each tick mark represents 100 ms of time. Negative voltage is plotted up. ERPs were filtered with a 15 Hz low-pass filter for plotting purposes only in these and all subsequent waveforms.

attractor had no effect on grammatical verbs: no brain-wave effect of processing difficulty emerged as a P600 or any other ERP component.

Overall these results are consistent with the notion that attraction interference in comprehension is asymmetrical with respect to grammaticality. Moreover, our results also indicate that this asymmetrical interference manifested itself very quickly after onset of the verb: the early effect of ungrammaticality on the P2 peak between 150 and 300 ms was absent following plural attractors. Note that we did not replicate the early positivity for grammatical verbs following plural attractors reported by [Kaan \(2002\)](#), nor did we replicate the pattern of agreement attraction reported by [Martin et al. \(2012\)](#). Kaan suggested that grammatical singular verbs following plural attractor nouns may initially be perceived as ungrammatical, but that participants quickly overcome this initial misparse. In the current experiment both visual inspection of the waveforms as well as statistical analyses showed no evidence of an early misparse in grammatical sentences, though there was a clear early effect of ungrammaticality, modulated by agreement attraction. Martin and colleagues showed that gender agreement violations involving NP ellipsis elicited a broadly distributed, sustained negativity, but that attraction impacted primarily grammatical sentences. We return to this point in the general discussion.

The pattern of results reported here is inconsistent with the MM model of agreement processing. If features of attractor nouns can erroneously value the subject NP as 'slightly more plural' (in the terms of the MM model), it is predicted that ungrammatical verbs will be perceived as 'slightly grammatical' and that grammatical verbs will be perceived as 'slightly ungrammatical' to the same extent. However, neither the judgment data nor ERP results from this experiment support that view for comprehension, suggesting that – assuming the MM is the right characterization for production – the processing mechanisms which give rise to attraction interference in comprehension and production are at least partly distinct.

One possible objection to this claim is that the time course of activation of the attractor noun's plural feature may differ in production and comprehension. For example, work by [Nicol \(1995\)](#), and more recently by [Pearlmutter and colleagues \(Gillespie & Pearlmutter, 2011, 2013; Solomon & Pearlmutter, 2004\)](#), argues that the timing of activation of the head and attractor nouns can impact rates of agreement attraction, where plural attractor nouns that are activated simultaneously with a singular head noun will cause more interference than those that are activated later. During speech planning, conceptual representations of the subject NP's referent(s) and associated numerosities are activated prior to morphosyntactic marking of number

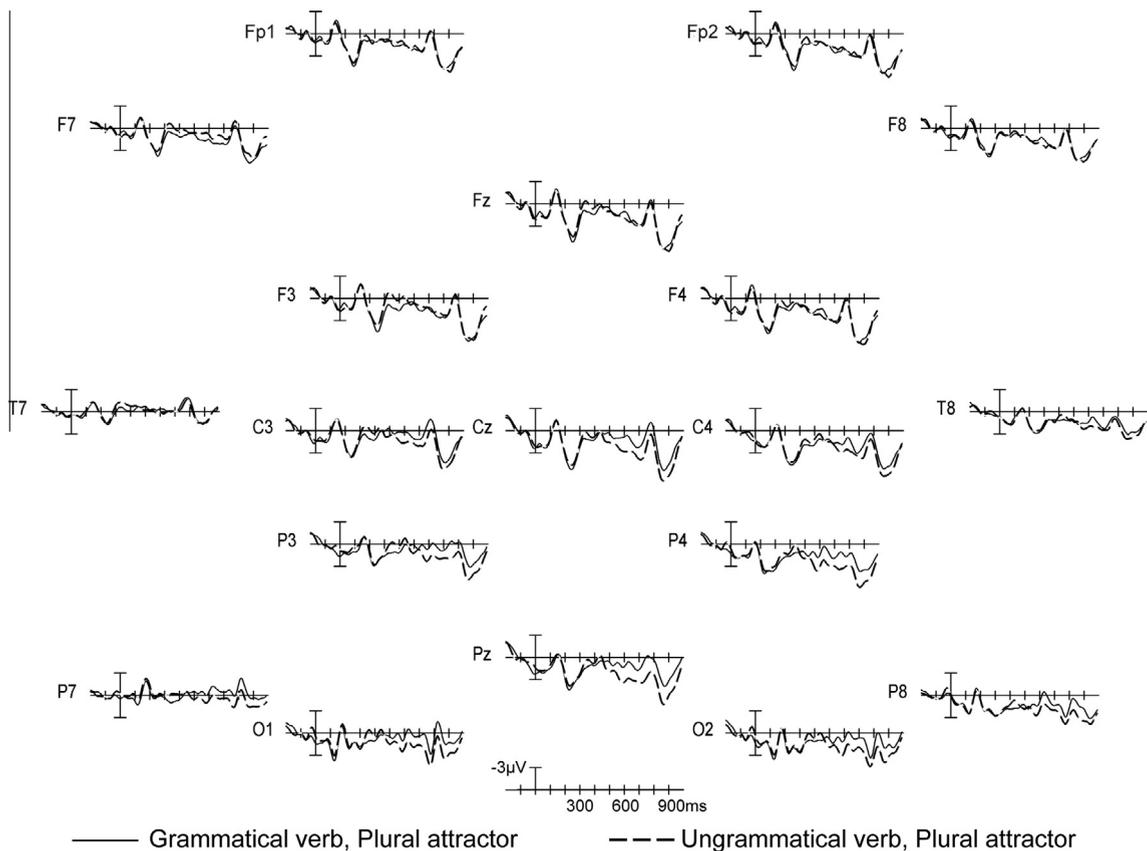


Fig. 2. Grand mean ERP waveforms from Experiment 1 in the grammatical (solid line) and ungrammatical (dashed line) plural attractor conditions. Onset of the verb is indicated by the vertical calibration bar; each tick mark represents 100 ms of time. Negative voltage is plotted up.

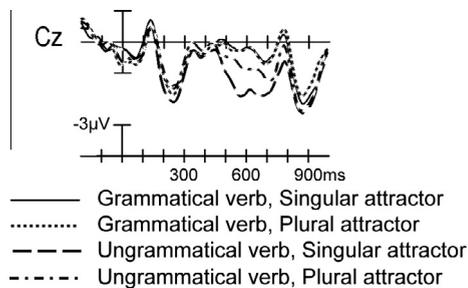


Fig. 3. Grand mean ERP waveforms for all four experimental conditions in Experiment 1 at midline vertex electrode Cz. Onset of the verb is indicated by the vertical calibration bar; each tick mark represents 100 ms of time. Negative voltage is plotted up.

on the verb according to the MM model. On this approach, a plural attractor's number feature will have adequate time to infiltrate the representation of the subject NP number and lead to interference in verb marking during production. However, in comprehension, information is added incrementally to a conceptual representation of the sentence as the language input unfolds over time. A reader or hearer will have no knowledge of an upcoming noun's number feature until that noun is eventually encountered, and incorporating a plural number feature into the syntactic or conceptual representation of the sentence will

necessarily take time. Thus, it may have been the case that, given the linear adjacency of the attractor and verb, participants did not have time to update the representation of the subject NP's number before encountering the verb, such that there was no increase in P600 amplitude for grammatical verbs following plural attractors.

Some evidence from this experiment argues against this objection. Principally, if attraction is driven primarily by the overall representation of the subject NP number (e.g., the $S(r)$ value, within the MM model), and a plural attractor noun had not had adequate time to alter participants' representations of subject NP number, there should have been no amelioration effect from plural attractors on ungrammatical verbs in the ERP record. The fact that plural attractors reduced positivities to ungrammatical verbs as early as 150 ms after onset of the verb suggests that the processes responsible for checking number agreement are sensitive to attraction interference from the earliest stages of anomaly detection. Nonetheless, there may be multiple mechanisms that can give rise to attraction in comprehension, each operating on a different time course. For example, retrieval interference, which would trigger an illusion of grammaticality like that reported here, may operate extremely rapidly. Reevaluation of the subject number feature, which would trigger an illusion of ungrammaticality, may require more time (see also Brehm & Bock, 2013) and

show up only in late measures of processing. Our judgment data do not support this possibility, however. While there was an overall decrease in accuracy following a plural attractor noun in participants' acceptability judgments, this was driven primarily by a decrease in accuracy in the ungrammatical condition. That is, attraction effects were similarly asymmetrical in both the ERP and off-line judgments. These off-line judgments represent a late, global assessment of sentence-level acceptability and should include representational information about subject NP numerosity. Nonetheless, allowing more time between the attractor noun and verb might facilitate reevaluation of subject NP number as more plural, allowing interference to manifest in grammatical sentences. We investigate this possibility in Experiment 2.

Experiment 2

In this experiment we tested the hypothesis that reevaluation of the subject NP's number feature following a plural attractor noun requires time, meaning that creating greater linear (and therefore temporal) separation between the attractor noun and verb would give rise to a more symmetrical pattern of attraction. Specifically, we linearly separated the attractor noun and verb with an adverb (cf. Wagers et al., 2009), thereby increasing the amount of time available for a plural attractor noun's number feature to impact the representation of the subject NP's number. If number representation is a crucial contributor to attraction interference in comprehension, as predicted by the MM model for production, we should see attraction effects for both grammatical and ungrammatical verbs in this experiment. Alternately, if agreement dependencies during comprehension are established primarily via cue-based search in working memory (e.g., McElree, 2006; Van Dyke, 2007; Van Dyke & McElree, 2011), we predict a pattern similar to that seen in Experiment 1.

Method

Participants

Participants in Experiment 2 were 22 native English speakers. Participants were all strongly right-handed, as assessed by an abridged version of the Edinburgh Handedness Inventory (Oldfield, 1971), and had normal or corrected-to-normal vision. Data from one participant were excluded after a language history questionnaire revealed significant early exposure to a language with agreement patterns very different from English (Tagalog). This resulted in 21 participants being included in the final analyses (14 male; mean age: 19.3 years, range: 18–23). These participants reported no significant knowledge of or exposure to languages other than English and were functionally monolingual. Participants either volunteered or received course credit for taking part.

Materials

Sentences from Experiment 1 were adapted by including a sentential adverb between the attractor noun

and auxiliary verb (e.g., *The chemist with the test tube(s) probably is/are conducting an experiment*). Sentence versions were distributed across four lists with 30 items per condition per list, as in Experiment 1. Filler sentences were identical to those used in Experiment 1.

Procedure and analysis

Procedures, EEG recording, and analyses were identical to those used in Experiment 1. An average of 6.6% of trials were rejected due to artifact, equally distributed across conditions.

Results

Behavioral results

Mean proportion of sentences judged correctly as acceptable or unacceptable and d-prime scores are presented in Table 3. A repeated measures ANOVA on arcsine-transformed proportions showed no main effect of grammaticality ($p > .1$), a significant effect of attractor number, $F(1,20) = 41.546$, $MSE = .012$, $p < .001$, and a trend toward an interaction between the two, $F(1,20) = 3.202$, $MSE = .012$, $p = .089$. Critical planned pairwise comparisons showed that a plural attractor decreased accuracy both for grammatical, $t(20) = 3.689$, $p = .002$, and ungrammatical sentences, $t(20) = 5.356$, $p < .001$. Analysis of d-prime scores showed a strong effect of attractor number, $F(1,20) = 43.944$, $MSE = .217$, $p < .001$. Overall this shows that plural attractors caused reliable interference, decreasing accuracy and sensitivity to subject–verb grammaticality. However, in contrast to Experiment 1, plural attractors decreased accuracy across both levels of grammaticality.

ERP results

Waveforms comparing participants' brain responses to grammatical and ungrammatical verbs are presented for the singular and plural attractor conditions in Figs. 4 and 5, respectively; waveforms comparing all four conditions at electrode Cz are presented in Fig. 6. Analysis of participants' brain responses revealed no significant effects in either the 150–300 ms or 300–500 ms time windows over any scalp region. Statistical analysis in the 500–800 ms showed a large main effect of grammaticality [midline: $F(1,20) = 33.765$, $MSE = 17.159$, $p < .001$; medial: $F(1,20) = 31.418$, $MSE = 34.974$, $p < .001$; lateral: $F(1,20) = 28.359$, $MSE = 10.200$, $p < .001$], indicating a large P600 effect for ungrammatical verbs. This effect was largest over posterior electrodes [grammaticality \times electrode interaction, midline: $F(2,40) = 27.703$, $MSE = 2.224$, $p < .001$; medial: $F(4, 80) = 20.108$, $MSE = 11.771$, $p < .001$; lateral: $F(2,40) = 26.329$, $MSE = 2.079$, $p < .001$].

Table 3

Means and standard errors for behavioral judgments in Experiment 2. Proportion judged correctly and d-prime scores. Standard errors are reported in parentheses.

| | Grammatical | Ungrammatical | d-Prime |
|---------------|-------------|---------------|------------|
| Sg. Attractor | .81 (.05) | .92 (.02) | 2.82 (.23) |
| Pl. Attractor | .75 (.05) | .80 (.04) | 1.87 (.24) |

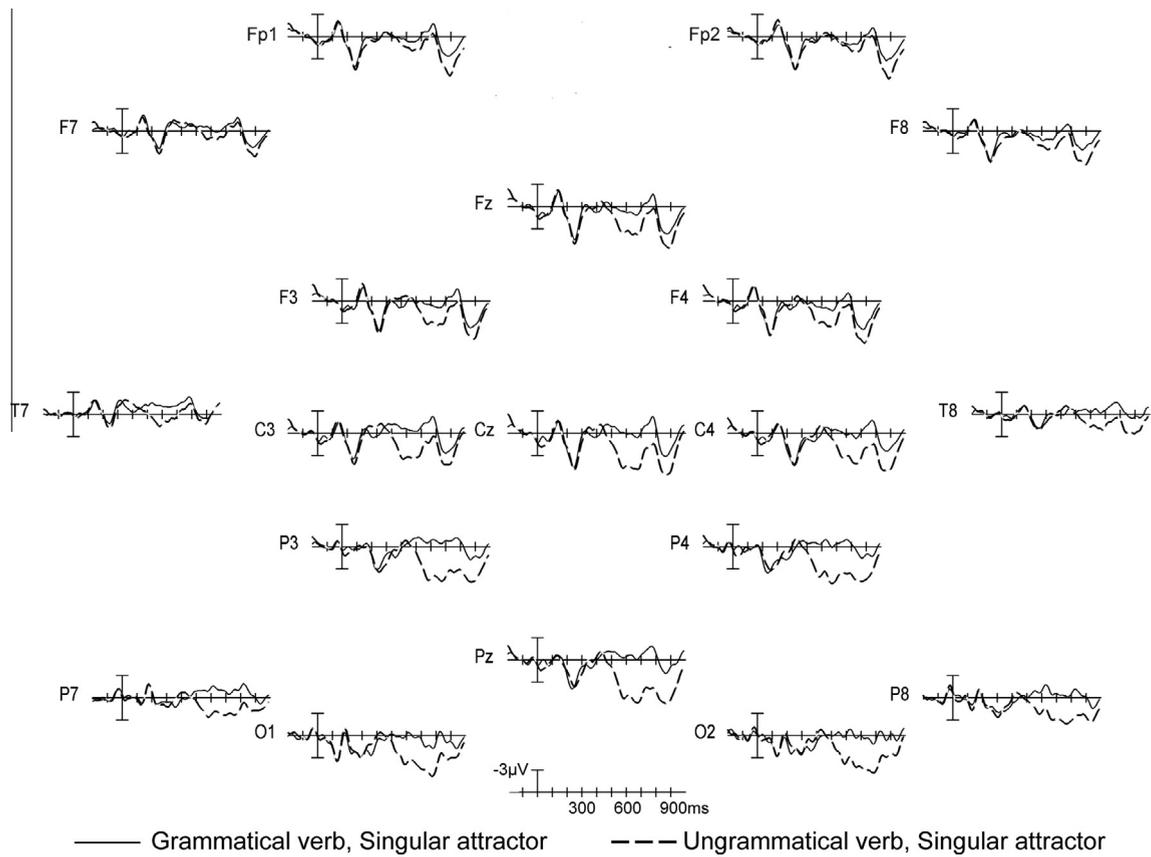


Fig. 4. Grand mean ERP waveforms from Experiment 2 for all 19 scalp electrodes in the grammatical (solid line) and ungrammatical (dashed line) singular attractor conditions. Onset of the verb is indicated by the vertical calibration bar; each tick mark represents 100 ms of time.

As in Experiment 1, there was an additional grammaticality by attractor number interaction over midline electrodes, $F(1,20) = 6.495$, $MSE = 10.564$, $p = .019$, and which neared significance over lateral sites, $F(1,20) = 3.338$, $MSE = 12.377$, $p = .083$. Pair-wise follow-up tests over midline electrodes showed that the effect of grammaticality was significant in both the singular attractor, $F(1,20) = 33.487$, $MSE = 15.630$, $p < .001$, and plural attractor conditions, $F(1,20) = 10.303$, $MSE = 12.094$, $p = .004$. Comparison of brain responses to grammatical verbs showed no reliable difference in the singular versus plural attractor conditions, $F(1,20) = 1.370$, $p > .25$, whereas there was a reliable difference between the two attractor conditions for ungrammatical verbs, $F(1,20) = 5.477$, $MSE = 10.905$, $p = .030$. Again, this indicates that the positivity elicited by ungrammatical verbs in the plural attractor condition was smaller than in the singular attractor condition.

Discussion

The results of Experiment 2 provide a somewhat mixed picture. The ERP results largely replicated the findings from Experiment 1, showing an asymmetrical pattern of attraction interference. This is despite an increased linear (and therefore temporal) separation between the attractor

noun and verb. If verb number is checked against a continuously valued global representation of subject NP number, as conceptualized within the MM model, comprehension measures should show that grammatical singular verbs following plural attractors should be perceived as slightly ungrammatical to the same extent as ungrammatical plural verbs are perceived as slightly grammatical. Allowing time for the plural attractor feature to spread through the subject NP representation should maximize the opportunity for this effect to manifest and become observable. However, as in Experiment 1, plural attractors reduced the P600 effect to ungrammatical verbs, but did not increase the P600 to grammatical verbs or trigger any other type of brainwave response. In contrast to these findings, participants' acceptability judgments of sentences did show a degree of attraction symmetry. Judgment accuracies were reliably lower following plural attractors, both for grammatical and ungrammatical sentences. Thus, although plural attractors caused no interference in grammatical sentences during initial processing as measured by ERPs, there was a degree of interference in late global measures of sentence acceptability.

One interpretation of these findings is that interference based on global subject number representation manifests itself only very late in processing. That is, initial parsing of verb number relies on a combination of predictive

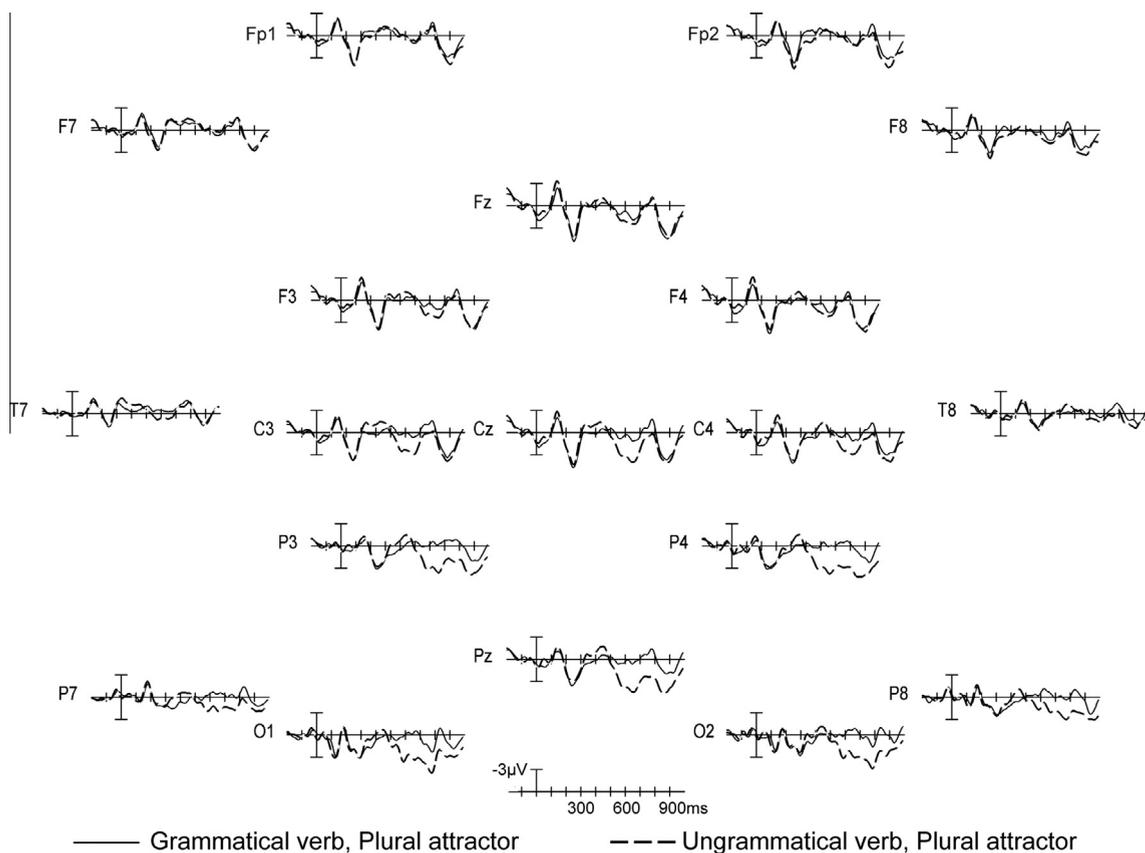


Fig. 5. Grand mean ERP waveforms from Experiment 2 in the grammatical (solid line) and ungrammatical (dashed line) plural attractor conditions. Onset of the verb is indicated by the vertical calibration bar; each tick mark represents 100 ms of time. Negative voltage is plotted up.

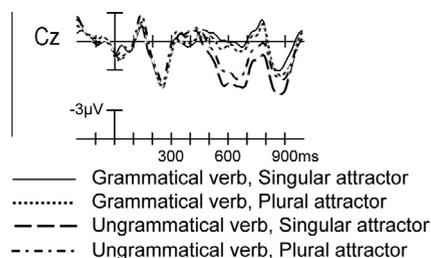


Fig. 6. Grand mean ERP waveforms for all four experimental conditions in Experiment 2 at midline vertex electrode Cz. Onset of the verb is indicated by the vertical calibration bar; each tick mark represents 100 ms of time. Negative voltage is plotted up.

mechanisms in tandem with cue-based retrieval mechanisms, which are susceptible to similarity-based interference, while later processes are more sensitive to the global representation of subject NP number, consistent with the MM model. Indeed, such a model could reconcile conflicting findings which have shown no interference in grammatical sentences using on-line measures at the verb (Wagers et al., 2009), but reliable interference in off-line, global measures of sentence comprehension (Nicol et al., 1997).

However, there are sufficient grounds to entertain caution in making this conclusion. First, interference effects

were asymmetrical in the judgment task in Experiment 1. In neither experiment were participants put under any time timing pressure to make responses, making a timing-based explanation of attraction interference unlikely to account for the data. Additionally, participants' accuracy in judging grammatical sentences was markedly lower in Experiment 2 than Experiment 1. This is likely a consequence of relative infrequency of adverb–auxiliary verb strings, which may lead to a lowered rate of acceptability in these instances. A search of the BYU Corpus of Contemporary American English (Davies, 2008–) showed that there were 2,653 instances of adverbs followed by a form of the verb *be* (5.72 per million corpus tokens), whereas there were 4,024,516 instances of *be* followed by an adverb (8673.15 per million corpus tokens). Although adverb–verb word orders are in principle grammatical in English, they are relatively restricted in their distribution. The decrease in accuracy in the judgment task for grammatical sentences in the plural attractor condition may therefore not have been a direct result of attraction interference *per se*, but rather an interaction between plural attractor number and lowered acceptability of adverb–auxiliary strings in general. A clear test of whether attraction in grammatical sentences manifests itself as a late, global effect therefore requires a test similar to that used in previous off-line studies of agreement processing.

In Experiment 3 we attempt to replicate two crucial findings reported by Nicol et al. (1997) using the same timed sentence judgment methodology.

Experiment 3

Nicol et al. (1997), in a timed sentence acceptability judgment task, showed that participants were reliably slower and less accurate when judging grammatical sentences with singular verbs and plural attractors versus sentences with singular verbs and singular attractors (Experiments 2, 4 and 5). A control experiment showed that, when then sentences included modal verbs, which are not marked for agreement in English, there was no difference in RT or accuracy for grammatical sentences in the singular and plural attractor conditions (Experiment 3). This is a critical finding, because it showed that the slow-down in their Exp. 2 was not a result of increased complexity from plural attractors, but depended on the need to process agreement. Here we attempt to replicate the results from Nicol et al. (1997) Experiments 2 and 3, using a within-subjects design. The sentence judgment task should be sensitive to attraction interference no matter when it occurs, even if it is a very late effect. If attraction in grammatical sentences manifests as a late effect, RT slow-downs caused by plural attractors should be modulated by the agreement status of the verb: the slow-downs should only be present in sentences with agreement-taking auxiliary verbs and not with modal verbs. Finding a slow-down in sentences containing plural attractors that is not modulated by the agreement-taking status of the auxiliary verb would suggest that it results from a more general complexity, and not a process crucial for agreement.

Method

Participants

Our participants were 36 monolingual native English speakers (13 male, mean age: 18.8 years, range: 18–20) and had normal or corrected-to-normal vision. Participants received course credit for taking part.

Materials

Sixty sentence frames were adapted from Experiments 1 and 2. Four versions of each sentence frame were created in a 2×2 design by crossing the factors attractor number (singular, plural) and verb type (agreeing auxiliary, modal), see Table 4. The four versions of each sentence frame were counterbalanced across four experimental lists in a Latin square design, such that each list contained 15 items per condition and no participant saw more than one version of a given item. Following Nicol et al. (1997), we also included a set of filler sentences for which we expected to see an RT difference: we constructed 20 pairs of sentences containing either subject or object relative clauses (RCs), all of which had animate NPs as the head of the RC and within the embedded clause (e.g., *The salesman that liked the fireman drank some coffee*/*The salesman that the fireman liked drank some coffee*). These sentences were

counterbalanced across lists 1 and 2, and again across lists 3 and 4. These grammatical sentences were randomized among 80 ungrammatical sentences that contained blatant violations of English word order restrictions (e.g., *The men's faces had for photographed been the police report*), such that each list contained an equal number of grammatical and ungrammatical sentences (160 sentences total). Experimental lists were randomized separately for each participant. 10 practice items preceded experimental trials.

Procedure and analysis

Participants were tested individually in a sound-attenuating booth. Each trial began with a fixation cross which was presented for 500 ms, followed by a 500 ms blank screen. Then the entire sentence was presented at once, and participants were asked to indicate as quickly and accurately as possible whether the string of words on the screen formed a grammatical sentence in English or not by pressing a button on a button box. Sentences were presented for a maximum of 4000 ms, after which the trial timed out and the next trial automatically began. 3.1% of trials were excluded due to timing out. Trials with RTs 2.5 SD larger or smaller than the participant's mean RT were replaced with the cutoff value. This impacted 1.1% of remaining trials. Only correctly-answered trials were included in the RT analyses.

We conducted ANOVAs on RTs and arcsine-transformed accuracies with both subjects (F1) and items (F2) as random factors. We additionally computed linear mixed effects models, with maximal random effects structures (crossed random intercepts for subjects and items, with random slopes for each experimental factor and their interaction by both items and subjects: Baayen, Davidson, & Bates, 2008; Barr, Levy, Scheepers, & Tily, 2013) for RTs and weighted empirical logit models by subjects and items for accuracies (Barr, 2008). For brevity, we report only the ANOVAs unless there were differences between the results. Differences only occurred for the agreement experiment error data.

Results

For the RC filler sentences, participants responded faster to subject RCs (mean RT = 2640.65 ms, $SE = 51.91$) than to object RCs (mean RT = 2767.89 ms, $SE = 51.0$), $F(1, 35) = 12.672$, $MSE = 22998.469$, $p = .001$, $F(1, 19) = 16.917$, $MSE = 9451.494$, $p < .001$. A similar pattern was obtained for accuracy, with subject RC sentences judged more accurately (mean proportion correct = .963, $SE = .016$) than object RC sentences (mean proportion correct = .916, $SE = .021$), $F(1, 35) = 4.928$, $MSE = .039$, $p = .033$, $F(1, 19) = 6.966$, $MSE = .019$, $p = .016$. As in Nicol et al. (1997), this indicates that the task is sufficiently sensitive to detect processing difficulty in grammatical sentences.

Mean RTs and proportion correct for the agreement experimental items are presented in Table 5. Analyses of arcsine-transformed proportion judged correctly showed a non-significant main effect of attractor number by subjects, $F(1, 35) = 2.862$, $p = .10$, and a significant main effect by items, $F(2, 59) = 4.117$, $MSE = .017$, $p = .047$. However, there were no effects of verb type, $F_s < .1$, nor was there

Table 4
Example experimental sentences from Experiment 3.

| Attractor number | Verb type | Sentence |
|------------------|-----------|--|
| Singular | Agreeing | <i>The chemist with the test tube is conducting an experiment</i> |
| Singular | Modal | <i>The chemist with the test tube might be conducting an experiment</i> |
| Plural | Agreeing | <i>The chemist with the test tubes is conducting an experiment</i> |
| Plural | Modal | <i>The chemist with the test tubes might be conducting an experiment</i> |

an interaction between attractor number and verb type, $F_s < .1$. Results obtained using a weighted empirical logit analysis failed to reveal any significant differences between conditions in the accuracy data (all $p_s > .27$), suggesting that the ~1% advantage in accuracy for sentences with singular attractors may not be statistically meaningful.

Analyses of RT data showed a reliable main effect of attractor number, $F(1,35) = 8.167$, $MSE = 19183.89$, $p = .007$, $F(2,1,59) = 7.673$, $MSE = 31141.39$, $p = .007$, and a reliable main effect of verb type, $F(1,35) = 52.311$, $MSE = 19099.811$, $p < .001$, $F(2,1,59) = 34.126$, $MSE = 47599.808$, $p < .001$, but no interaction between the two, $F(1,35) = .070$, $p = .794$, $F(2,1,59) = .0003$, $p = .986$. This indicates that participants were slower responding to sentences containing modal verbs, possibly because of increased length of these sentences, and that participants were slower responding to sentences with plural attractors. However, there was no sign of an interaction between the factors, indicating that the impacts of verb type and attractor number were independent.

Discussion

This experiment investigated whether interference effects manifested themselves in grammatical sentences using a measure that is sensitive to late-occurring processes. A previous study using this measure showed in a between-subject design that attraction in grammatical sentences led to slower RTs and a higher error rate in sentence judgments, but only when the auxiliary verb was overtly marked for agreement (Nicol et al., 1997). Using a within-subjects design, we were unable to replicate this finding. RTs were slower and judgments were slightly less accurate for sentences containing plural attractor nouns, though this difficulty was not modulated by auxiliary verb type: it occurred both in sentences containing overtly agreement-marked verbs and sentences containing modal verbs with no overt agreement marking. The current data therefore suggest that the slowdown and decrease in accuracy are not directly linked with the agreement processes

themselves. Instead there seems to be a more general cost associated with processing complex NPs with plural attractors, which is independent of agreement computation.

Based on discussion by Wagers et al. (2009), one possibility is that plural nouns incur processing costs over and above the cost associated with per-letter reading time. That is, the effect may reflect the processing of plurality itself. Alternatively, it is possible the cost is associated with integrating a plural noun into an otherwise singular NP. This latter possibility is supported by Nicol et al. (1997). They included sentences in which a plural head noun was modified by a PP containing either a plural or singular attractor noun. Across experiments, a comparison of RTs to sentences containing a singular head and plural attractor with RTs to sentences containing a plural head and plural attractor shows longer RTs for singular-head/plural attractor sentences. Therefore, it seems likely that the configuration of singular-head/plural attractor is what is difficult. Our results do clearly show that there was no added cost associated with processing agreement in sentences containing a plural attractor noun, even with a late, global measure of sentence processing. This is consistent with the view that number attraction interference does not manifest itself in grammatical sentences as a direct result of processing agreement dependencies.

There remains a discrepancy between the current findings and those of Nicol et al. (1997), who used a between-subjects design. A possibility which we would like to pursue here is that perhaps in Nicol et al.'s original study, the lack of agreeing verbs in Experiment 3 created a comprehension strategy that downplayed number specification within the complex NP subject: participants simply stop paying much attention to noun number if it never interacts with a verb. This would give rise to the observed effect of attractor number in their Experiment 2 (resulting from difficulty integrating a plural noun into a singular NP) and lack of effect in their Experiment 3, and would simultaneously explain why attractor number and verb type did not interact in the present study. Similar effects of stimulus list composition are well-established in lexical processing tasks (e.g., Hutchison, Neely, & Johnson, 2001; Lau, Holcomb, & Kuperberg, 2013; Neely, 1991). The results of the present study, using the within-subject design, may therefore be more representative of more naturalistic language processing contexts, where comprehenders must be prepared to process a mix of verb types.

It is also useful to ask whether effects of list composition could be partly responsible for the pattern of effects seen across the three experiments here, since half of the sentences in Experiments 1 and 2 contained some sort of agreement ungrammaticality (either subject-verb or quantifier-noun agreement violation). Since agreement-

Table 5
Mean RTs and proportion judged correctly in Experiment 3. Standard errors are reported in parentheses.

| Attractor number | Agreeing aux | | Modal aux | |
|------------------|-----------------|--------------------|-----------------|--------------------|
| | RT (ms) | Proportion correct | RT (ms) | Proportion correct |
| Singular | 2393.74 (42.39) | .984 (.005) | 2565.85 (48.55) | .986 (.005) |
| Plural | 2465.23 (53.51) | .973 (.007) | 2626.30 (55.23) | .973 (.008) |

violating sentences like these are relatively rare in normal language experience (though the production experiments demonstrate that errors do arise with some regularity), the inclusion of ungrammatical sentences may have altered participants' expectations about upcoming information and more general processing strategies. We see two ways which the inclusion of ungrammatical sentence might have altered participants' processing behaviors. On the one hand, knowing that in half of the sentences expectations about upcoming agreement will be wrong, participants might 'give up' trying to predict upcoming elements and show decreased sensitivity to violations (cf. Coulson, King, & Kutas, 1998; Hahne & Friederici, 1999). However, our results showed clear sensitivity to violations, indexed by P600 effects, and we found clear evidence of prediction both in the time course of effects (Experiment 1, see below) and in the observed attraction asymmetries in Experiments 1 and 2. Alternately, the inclusion of ungrammatical sentences could heighten the salience of subject–verb agreement within the experimental context, given that it was relevant to the acceptability judgment task employed in Experiments 1 and 2. Raising the salience of agreement would increase the likelihood of detecting some sort of processing difficulties in grammatical plural attractor sentences in Experiments 1 and 2. However, none was found. Importantly, this lack of agreement-based attraction was similar to the finding of Experiment 3, where there were no agreement-violating sentences. We therefore feel confident in concluding that the pattern of effects across our experiments is not a result of experimental artifacts caused by including ungrammaticalities.

General discussion

We studied effects of subject–verb agreement attraction during comprehension in three experiments in order to understand the contexts in which attraction interference arises, the time course of attraction effects, and most importantly, to assess whether the mechanisms purported to influence production of agreement are operative during comprehension. This latter point is important in that although prior theoretical accounts of attraction interference differ in their interpretation of attraction's causes, the prevailing assumption across the literature is that attraction interference in comprehension and production have isomorphic sources (Badecker & Kuminiak, 2007; Hartsuiker, 2006; Nicol et al., 1997; Pearlmutter et al., 1999; Severens et al., 2008; Thornton & MacDonald, 2003; Wagers et al., 2009; see also Kreiner et al., 2013). However, in clear contrast to this assumption, our data support the conclusion that the causes of attraction interference in comprehension and production are non-identical, and more specifically, that the mechanisms responsible for attraction in comprehension are a subset of those operative in production. This conclusion is also relevant to recent general models of language processing that posit a tight link between the two tasks (MacDonald, 2013; Pickering & Garrod, 2013). Our findings highlight one interesting area where there is divergence in some of the subprocesses involved in the two tasks.

Asymmetric attraction during comprehension

The results of our experiments squarely support the notion that attraction in comprehension is *asymmetric*, in that it targets primarily ungrammatical contexts (Dillon et al., 2013; Shen et al., 2013; Wagers et al., 2009). This contrasts with mounting findings from production-oriented tasks showing processing difficulty in singular head/plural attractor configurations, even when the correct verb is ultimately chosen (i.e., *symmetric* attraction: Brehm & Bock, 2013; Haskell & MacDonald, 2003; Staub, 2009, 2010). In these production studies, attraction interference is frequently attributed to faulty global representation of the subject NPs number feature, in line with the MM model (Brehm & Bock, 2013; Eberhard et al., 2005; Staub, 2009). However, across three experiments, we saw no clear evidence for global number representation impacting agreement processes in either early or late measures. Experiments 1 and 2 used ERPs' high temporal resolution to investigate the time course of subject–agreement processing and attraction interference at the verb itself and found that agreement attraction ameliorated the effect of ungrammaticality at disagreeing verbs, but did not induce a perception of ungrammaticality, or evidence of any other processing difficulty, at agreeing verbs. These findings held both when the attractor and verb were linearly adjacent (Experiment 1), as well as linearly and temporally separated by an adverb (Experiment 2). Moreover, Experiment 1 showed an effect of ungrammaticality modulated by attraction as early as 150–300 ms after onset of the verb. This suggests that the mechanisms supporting subject–verb computation and giving rise to attraction interference operate very rapidly and that attraction interference was detected from the earliest stages of agreement integration.

Experiment 3 used whole-sentence judgments to test whether a late process that might compute global number representation contributed to the comprehension of subject–verb agreement. If re-valuation of the subject NP's global number feature occurs slowly, it may only be evident in late measures as the whole sentence is conceptualized. Results showed that while sentences containing a singular head and singular attractor were judged to be grammatical reliably faster than sentences with a singular head and plural attractor, this effect was not modulated by the need to compute agreement. It occurred both in sentences with agreement-taking auxiliaries (e.g., *is* or *was*) and in sentences with modal verbs, which are not overtly marked for agreement in English. The present data therefore suggest that sentences with a singular head and plural attractor may indeed be more difficult to process than sentences with matching singular attractors, but that this difficulty likely results from difficulty incorporating a plural noun into an otherwise singular NP, and not from the processing of the subject–verb agreement dependency itself. In contrast to previous reports attributing attraction interference to faulty global number representation, we found no evidence of this across three experiments, regardless of the dependent measure used (ERPs, sentence judgment RT and accuracy).

Using ERPs to study attraction adjudicates between some possible reasons for the contrast between our results

and the previous literature. First, ERPs revealed attraction's asymmetry without the use of statistical covariates, which could conceivably have inadvertently removed the effects of attraction in grammatical sentences (cf. Wagers et al., 2009). Second, ERPs allowed us to establish that a plural attractor following a singular head noun does not introduce any detectable processing difficulty at the verb, which might have manifested as either a P600 or any other qualitatively different ERP component. Instead, interference served merely to reduce P600 magnitude, suggesting it ameliorated processing difficulty associated with the ungrammatical verb. Third, we established a time course for the appearance of both ungrammaticality and attraction effects, as both effects appeared in a time window beginning as early as 150 ms after presentation of the verb. This rapid time course of syntactic processing is consistent with other research showing early effects of agreement ungrammaticality within the P2 time window in agreement tasks (Osterhout & Mobley, 1995), as well as other syntactic manipulations where diagnosis of an anomaly can occur rapidly (Friederici et al., 2001; McKinnon & Osterhout, 1996).

The pattern of results across all three experiments reported here indicates that agreement attraction effects in comprehension is (1) asymmetrical, targeting primarily ungrammatical contexts, (2) extremely fast, and (3) fails to show evidence of interference from global NP number representation in agreement processing, even in late measures. Overall, our results are consistent with the hypothesis that attraction interference in comprehension is driven primarily by similarity-based memory retrieval interference. However, in contrast to previous interpretations of attraction effects across the two modalities (e.g., Hartsuiker, 2006; Pearlmutter et al., 1999; Thornton & MacDonald, 2003; Wagers et al., 2009), we argue that there are non-identical loci for interference phenomena in agreement production and comprehension.

Comprehension and production

The existing body of work on agreement production supports the idea that several factors affect the formation of agreement dependencies during speaking, including structural factors (e.g., conceptual number and NP morphology as formalized within the MM model), timing and scope-of-planning effects (Gillespie & Pearlmutter, 2011, 2013; Nicol, 1995; Solomon & Pearlmutter, 2004), and similarity-based memory retrieval interference (Badecker & Kuminiak, 2007; Hartsuiker et al., 2003). Each of these sources, and possibly others, potentially conspire in giving rise to the interference effects established across experiments.

In comprehension, however, recent findings suggest that similarity-based retrieval interference at the level of morphological features is the primary source of interference, at least for agreement features on verbs (cf. Dillon et al., 2013; Parker, Du, & Phillips, 2014). Some have also argued that this mechanism accounts equally well for both comprehension and production data (Wagers et al., 2009). Our results similarly support retrieval interference as the primary driver of attraction in comprehension. However,

unlike Wagers and colleagues, who suggest that similarity-based retrieval interference can account for the broad set of findings across the two tasks, we argue that attraction effects in comprehension are primarily driven by a subset of the mechanisms active in production, and this difference arises from basic facts about the two tasks.

One major difference between comprehension and production is that in language production, the intended message to be conveyed is known to the speaker prior to the execution of the speech plan. This intended message includes information about the individual NPs to be uttered, their individual number specifications, as well as the notional number of the referent set indicated by the subject NP, and the semantic relations among the nouns. This suggests that there are more avenues for interference in subject-verb agreement production than in subject-verb agreement comprehension. These mechanisms give rise to several well-attested effects in production, including effects of notional number on verb agreement and effects of scope-of-planning. In later stages of speech planning, morphophonological information about number and overt case marking become available, such that attraction in these cases receives a straightforward explanation as resulting from similarity-based interference within a content-addressable working memory architecture.

In contrast, in language comprehension, top-down conceptual information about notional number of the subject NP is unknown to the comprehender, at least under the conditions tested in typical psycholinguistic experiments (that is, sentences presented in isolation). In this type of task, linguistic input is accessed serially, restricting the window of temporal co-activation of nouns: linguistic input is added incrementally to the comprehender's representation of the sentence over time. This predicts that interference from semantic integration and notional plurality should be restricted in comprehension (e.g. late; contingent upon discourse context). Previous research is consistent with this idea: the few studies investigating the role of notional number in comprehension suggest that it has a limited role. Nicol et al. (1997, Experiment 4) investigated the role of notional distributivity by contrasting sentences with notionally singular NPs (e.g., *The bridge across the canyon(s) is safe*) to sentences with distributed referent NPs (e.g., *The ad on the billboard(s) was very offensive*). They showed effects of grammatical number, with slowed acceptability judgments for sentences with singular head/plural attractor NPs, but no effects of distributivity even in late, whole-sentence processing measures. Work by Kreiner and colleagues (Kreiner et al., 2013) used eye-tracking to demonstrate that noun collectivity had no effect on the processing of subject-verb agreement violations in early processing measures, with semantic effects present only in second pass reading times. This suggests that initial processing of subject-verb agreement is relatively immune to semantic influences, though notional number may play a role late in processing. Future research may wish to investigate whether strengthening the availability of notional number cues might enhance semantic effects in agreement comprehension (e.g., using discourse context or a using images to strengthen the distributivity of the NP references; cf. Eberhard, 1999).

Differing roles for prediction

Both comprehension and production require tracking of what goes together; however, differences in what is tracked implies that prediction has different impacts on agreement across domains. A large body of research points to a prominent role for predictive mechanisms during language comprehension at multiple levels, including word-form, semantics, discourse, and syntax (DeLong, Urbach, & Kutas, 2005; Dikker, Rabagliati, Farmer, & Pykkänen, 2010; Dikker, Rabagliati, & Pykkänen, 2009; Farmer, Christiansen, & Monaghan, 2006; Federmeier & Kutas, 1999; Fine, Jaeger, Farmer, & Qian, 2013; Levy, 2008; Lew-Williams & Fernald, 2007; Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005; Wicha, Moreno, & Kutas, 2004; Wlotko & Federmeier, 2012; see Pickering & Garrod, 2013; Van Petten & Luka, 2012, for recent reviews). These predictive mechanisms in comprehension are rooted in the input that the perceiver has been exposed to and what speakers have previously produced, and as such, are yet another link between production and comprehension.

In language production, one way that predictive information has been shown to impact agreement is through lexical co-occurrence probabilities obtained through language experience (Haskell et al., 2010). In particular, Haskell and colleagues argue that speakers' past experience with agreement shapes production and can account for rates of plural agreement with collective nouns. These statistics are similar to the statistically-based distributional properties that have also been shown to have widespread effects in language comprehension (e.g., Fine et al., 2013; Garnsey, Pearlmutter, Myers, & Lotocky, 1997; Gennari & MacDonald, 2009; Reali & Christiansen, 2007; Trueswell, Tanenhaus, & Kello, 1993), including studies using ERPs (Coulson et al., 1998; Hahne & Friederici, 1999; Osterhout et al., 1994).

However, there are critical differences between the effects demonstrated by Haskell and colleagues (2010) and the effects demonstrated in the current paper. In a corpus analysis, the Haskell group showed that attraction in spontaneous production data is primarily restricted to constructions involving particular lexical items, in particular collective NPs (e.g., "a number of considerations"). The stimuli in the current experiment were designed to be construed as neither collective nor distributive NPs. Given this, the lexical co-occurrence account fails to adequately predict attraction interference with our stimuli for two reasons. First, comprehenders should have little experience with plural agreement in non-distributive singular head/plural attractor configurations (at least based on Haskell et al.'s corpus analyses). Second, and more importantly, this account would predict that attraction interference should be symmetrical with respect to grammaticality (i.e., impacting both grammatical and ungrammatical sentences). That is, co-occurrence statistics should influence agreement processing in general, not just in ungrammatical sentences. Thus, an experienced-based account of agreement focused on co-occurrence statistics cannot explain the pattern of results obtained here as effectively as an account relying on a combination of prediction and

similarity-based interference, suggesting again, another potential divide between production and comprehension.

The current account does support a role for prediction in agreement comprehension, though prediction of a different sort. We suggest that predictive mechanisms, in addition to retrieval mechanisms, are active in the processing of subject–verb agreement dependencies during comprehension, but that what is predicted is syntactic structure and feature specification (see also Dillon et al., 2013; Wagers et al., 2009). Prediction of upcoming syntactic structure and lexical categories is a prominent feature of Lewis and Vasishth's parsing model, which incorporates the same cue-based retrieval mechanisms advocated here (Lewis & Vasishth, 2005; Lewis et al., 2006). We suggest that these syntactic predictions include specifications of upcoming verbal agreement features.

In the case of the present data, appealing to this sort of predictive mechanism can potentially account for the rapid time course of the early positivity seen in Experiment 1, where ungrammaticalities were detected within 150 ms of stimulus onset. Moreover, the prediction account can explain the apparent differences in attraction profiles observed here for subject–verb agreement relations and by Martin and colleagues (2012) for gender agreement violations in Spanish NP ellipsis constructions. Recall that in the Martin study, attraction was found primarily to impact grammatical sentences, though attraction was also weakly detected in ungrammatical sentences. At first glance it unclear why cue-based retrieval mechanisms would operate differently in subject–verb agreement dependences and ellipsis constructions involving gender agreement, where attraction is asymmetrical in the first case and shows a degree of symmetry in the latter. However, predictability of a particular dependency type can play an important role in how retrieval operations are deployed (Dillon et al., 2013), and provide a potential explanation for the observed differences between the dependency types.

Within the parsing model outlined by Lewis and Vasishth (2005), syntactic predictions for obligatory upcoming constituents are automatically generated. Thus, given the detection of an initial subject NP, a prediction for an upcoming verb will be entered into the predicted structure stored in memory. Following Wagers et al. (2009), retrieval operations for predicted features are error-driven, only being executed when there is a mismatch between the top-down predicted and bottom-up detected input. One aspect of retrieval as an error-driven process is that it can account for the asymmetrical bias for agreement attraction to occur only in ungrammatical sentences (see Wagers et al., 2009), though this would require morphological features (e.g. a singular VP) to be specified within the syntactic prediction. However, the need to compute agreement in the NP ellipsis constructions investigated by Martin and colleagues cannot reliably be predicted, and so syntactic predictions would contain no morphological feature specification for anaphoric gender. Given no top-down feature prediction, a search for an agreement controller would have to be executed in order to establish a dependency, and this would occur for both grammatical and ungrammatical sentences. In such a scenario, a degree of attraction symmetry would be

expected. Further research in this domain could be fruitful in investigating the role of error-based learning in language comprehension.

Conclusion

The results of the three experiments reported here support the view that attraction interference in subject–verb agreement comprehension arises due to a combination of morphosyntactic prediction coupled with cue-based memory retrieval operations. The results are not consistent with the view that the source of interference is the global number representation of the subject NP. In contrast to previous proposals, our results suggest that the mechanisms giving rise to attraction interference in comprehension are non-isomorphic with those in production, where global number representation plays a prominent role.

Acknowledgments

We would like to thank Lee Osterhout and Judith McLaughlin. Portions of this research were conducted while the first author was supported by NSF Grant OISE-0968369 (Experiment 3) and the William Orr Dingwall Neurolinguistics Fellowship (Experiments 1 and 2). Experiments 1 and 2 were also partially funded by NIDCD Grant R01DC01947 and NSF Grant BCS-0951595. We would like to thank Martin Pickering and two anonymous reviewers for helpful comments on a previous version of this manuscript.

References

- Allen, M., Badecker, W., & Osterhout, L. (2003). Morphological analysis in sentence processing: An ERP study. *Language and Cognitive Processes, 18*, 405–430.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.
- 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.
- Barber, H., & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience, 17*, 137–153.
- Barr, D. J. (2008). Analyzing “visual world” eyetracking data using multilevel logistic regression. *Journal of Memory and Language, 59*, 457–474. <http://dx.doi.org/10.1016/j.jml.2007.09.002>.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255–278. <http://dx.doi.org/10.1016/j.jml.2012.11.001>.
- Bock, K., Carreiras, M., & Meseguer, E. (2012). Number meaning and number grammar in English and Spanish. *Journal of Memory and Language, 66*, 17–37. <http://dx.doi.org/10.1016/j.jml.2011.07.008>.
- Bock, K., & Cutting, J. C. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language, 31*, 99–127.
- Bock, K., Dell, G. S., Chang, F., & Onishi, K. H. (2007). Persistent structural priming from language comprehension to language production. *Cognition, 104*, 437–458. <http://dx.doi.org/10.1016/j.cognition.2006.07.003>.
- Bock, K., & Eberhard, K. (1993). Meaning, sound, and syntax in English number agreement. *Language and Cognitive Processes, 8*, 57–99.
- Bock, K., & Middleton, E. L. (2011). Reaching agreement. *Natural Language & Linguistic Theory, 29*, 1033–1069. <http://dx.doi.org/10.1007/s11049-011-9148-y>.
- Bock, K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology, 23*, 45–97.
- Brehm, L., & Bock, K. (2013). What counts in grammatical number agreement? *Cognition, 128*, 149–169. <http://dx.doi.org/10.1016/j.cognition.2013.03.009>.
- Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes, 13*, 21–58.
- Davies, M. (2008-). *The Corpus of Contemporary American English: 450 million words, 1990-present*. <<http://corpus.byu.edu/coca/>>.
- 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. <http://dx.doi.org/10.1038/nn1504>.
- Dikker, S., Rabagliati, H., Farmer, T. A., & Pykkänen, L. (2010). Early occipital sensitivity to syntactic category is based on form typicality. *Psychological Science, 21*, 629–634. <http://dx.doi.org/10.1177/0956797610367751>.
- Dikker, S., Rabagliati, H., & Pykkänen, L. (2009). Sensitivity to syntax in visual cortex. *Cognition, 110*, 293–321. <http://dx.doi.org/10.1016/j.cognition.2008.09.008>.
- 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. <http://dx.doi.org/10.1016/j.jml.2013.04.003>.
- Eberhard, K. (1997). The marked effect of number on subject-verb agreement. *Journal of Memory and Language, 36*, 147–164.
- Eberhard, K. (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., Cutting, J. C., & Bock, K. (2005). Making syntax of sense: Number agreement in sentence production. *Psychological Review, 112*, 531–558.
- Farmer, T. A., Christiansen, M. H., & Monaghan, P. (2006). Phonological typicality influences on-line sentence comprehension. *Proceedings of the National Academy of Sciences of the United States of America, 103*, 12203–12208. <http://dx.doi.org/10.1073/pnas.0602173103>.
- Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language, 41*, 469–495. <http://dx.doi.org/10.1006/jmla.1999.2660>.
- Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology, 47*, 164–203. [http://dx.doi.org/10.1016/S0010-0285\(03\)00005-7](http://dx.doi.org/10.1016/S0010-0285(03)00005-7).
- Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science, 11*, 11–15.
- Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. *PLoS ONE, 8*(10), e77661. <http://dx.doi.org/10.1371/journal.pone.0077661>.
- 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. <http://dx.doi.org/10.1080/01690965.2010.550166>.
- Franck, J., Lassi, G., Frauenfelder, U. H., & Rizzi, L. (2006). Agreement and movement: A syntactic analysis of attraction. *Cognition, 101*, 173–216.
- Franck, J., Vigliocco, G., & Nicol, J. (2002). Subject-verb agreement errors in French and English: The role of syntactic hierarchy. *Language and Cognitive Processes, 17*, 371–404.
- Frencq-Mestre, C., Osterhout, L., McLaughlin, J., & Foucart, A. (2008). The effect of phonological realization of inflectional morphology on verbal agreement in French: Evidence from ERPs. *Acta Psychologica, 128*, 528–536.
- Friederici, A. D., Hahne, A., & Mecklinger, A. (1996). Temporal structure of syntactic processing: Early and late event-related potential effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1219–1248.
- Friederici, A. D., Mecklinger, A., Spencer, K. M., Steinhauer, K., & Donchin, E. (2001). Syntactic parsing preferences and their on-line revisions: A spatio-temporal analysis of event-related brain potentials. *Cognitive Brain Research, 11*, 305–323.
- Garnsey, S. M., Pearlmutter, N. J., Myers, E., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language, 37*, 58–93.
- Gennari, S. P., & MacDonald, M. C. (2009). Linking production and comprehension processes: The case of relative clauses. *Cognition, 111*, 1–23. <http://dx.doi.org/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. <http://dx.doi.org/10.1016/j.cognition.2010.10.008>.

- Gillespie, M., & Pearlmutter, N. J. (2013). Against structural constraints in subject-verb agreement production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39, 515–528. <http://dx.doi.org/10.1037/a0029005>.
- Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 1411–1423.
- Gordon, P. C., Hendrick, R., & Levine, W. H. (2002). Memory-load interference in syntactic processing. *Psychological Science*, 13, 425–430.
- Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience*, 15, 883–899. doi:10.1162/089892903322370807.
- Hagoort, P., & Brown, C. M. (1999). Gender electrified: ERP evidence on the syntactic nature of gender processing. *Journal of Psycholinguistic Research*, 28, 715–728.
- Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech compared to reading: The P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia*, 38, 1531–1549.
- Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience*, 11, 194–205.
- Hartsuiker, R. J. (2006). Are speech error patterns affected by a monitoring bias? *Language and Cognitive Processes*, 21, 856–891.
- Hartsuiker, R. J., Antón-Méndez, I., & van Zee, M. (2001). Object attraction in subject-verb agreement construction. *Journal of Memory and Language*, 45, 546–572.
- Hartsuiker, R. J., Schriefers, H. J., Bock, K., & Kikstra, G. M. (2003). Morphophonological influences on the construction of subject-verb agreement. *Memory & Cognition*, 31, 1316–1326.
- Haskell, T. R., & MacDonald, M. C. (2003). Conflicting cues and competition in subject-verb agreement. *Journal of Memory and Language*, 48, 760–778. [http://dx.doi.org/10.1016/S0749-596X\(03\)00010-X](http://dx.doi.org/10.1016/S0749-596X(03)00010-X).
- Haskell, T. R., Thornton, R., & MacDonald, M. C. (2010). Experience and grammatical agreement: Statistical learning shapes number agreement production. *Cognition*, 114, 151–164. <http://dx.doi.org/10.1016/j.cognition.2009.08.017>.
- Hutchison, K. A., Neely, J. H., & Johnson, J. D. (2001). With great expectations, can two “wrongs” prime a “right”? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 1451–1463. <http://dx.doi.org/10.1037/0278-7393.27.6.1451>.
- Jasper, H. H. (1958). The ten–twenty system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, 10, 371–375.
- Kaan, E. (2002). Investigating the effects of distance and number interference in processing subject-verb dependencies: An ERP study. *Journal of Psycholinguistic Research*, 31, 165–193.
- Kaan, E., Harris, A., Gibson, E., & Holcomb, P. J. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, 15, 159–201.
- Kaan, E., & Swaab, T. Y. (2003). Repair, revision and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience*, 15, 98–110.
- Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, 52, 205–225. <http://dx.doi.org/10.1016/j.jml.2004.10.002>.
- 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. <http://dx.doi.org/10.1080/01690965.2012.667567>.
- Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, 1146, 23–49. <http://dx.doi.org/10.1016/j.brainres.2006.12.063>.
- Lau, E. F., Holcomb, P. J., & Kuperberg, G. R. (2013). Dissociating N400 effect of prediction from association in single word contexts. *Journal of Cognitive Neuroscience*, 25, 484–502. http://dx.doi.org/10.1162/jocn_a.00328.
- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106, 1126–1177. <http://dx.doi.org/10.1016/j.cognition.2007.05.006>.
- Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science*, 18, 193–198. <http://dx.doi.org/10.1111/j.1467-9280.2007.01871.x>.
- Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, 29, 375–419.
- 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.
- Lorimor, H., Bock, K., Zalkind, E., Sheyman, A., & Beard, R. (2008). Agreement and attraction in Russian. *Language and Cognitive Processes*, 23, 769–799. <http://dx.doi.org/10.1080/01690960701774182>.
- MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, 4(226). <http://dx.doi.org/10.3389/fpsyg.2013.00226>.
- MacWhinney, B., Bates, E., & Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior*, 23, 127–150.
- Martin, A. E., & McElree, B. (2008). A content-addressable pointer underlies comprehension of verb–phrase ellipsis. *Journal of Memory and Language*, 58, 879–906. <http://dx.doi.org/10.1016/j.jml.2007.06.010>.
- Martin, A. E., Nieuwland, M. S., & Carreiras, M. (2012). Event-related brain potentials index cue-based retrieval interference during sentence comprehension. *NeuroImage*, 59, 1859–1869. <http://dx.doi.org/10.1016/j.neuroimage.2011.08.057>.
- McElree, B. (2000). Sentence comprehension is mediated by content-addressable memory structures. *Journal of Psycholinguistic Research*, 29, 111–123.
- McElree, B. (2006). Accessing recent events. In B. H. Ross (Ed.), *Psychology of learning and motivation* (Vol. 46, pp. 155–200). San Diego: Academic Press.
- McElree, B., Foraker, S., & Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language*, 48, 67–91.
- McKinnon, R., & Osterhout, L. (1996). Constraints on movement phenomena in sentence processing: Evidence from event-related brain potentials. *Language and Cognitive Processes*, 11, 495–523.
- 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. <http://dx.doi.org/10.1016/j.cortex.2011.02.019>.
- Molinaro, N., Vespignani, F., Zamparelli, R., & Job, R. (2011b). Why brother and sister are not just siblings: Repair processes in agreement computation. *Journal of Memory and Language*, 64, 211–232. <http://dx.doi.org/10.1016/j.jml.2010.12.002>.
- Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Hillsdale, NJ: Erlbaum.
- 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. (1995). Effects of clausal structure on subject-verb agreement errors. *Journal of Psycholinguistic Research*, 24, 507–516.
- Nicol, J., Forster, K., & Veres, C. (1997). Subject-verb agreement processes in comprehension. *Journal of Memory and Language*, 36, 569–587.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia*, 9, 97–113.
- Osterhout, L., Allen, M. D., McLaughlin, J., & Inoue, K. (2002). Brain potentials elicited by prose-embedded linguistic anomalies. *Memory & Cognition*, 30, 1304–1312.
- Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, 31, 785–806.
- Osterhout, L., Holcomb, P. J., & Swinney, D. A. (1994). Brain potentials elicited by garden-path sentences: Evidence for the application of verb information during parsing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 786–803.
- Osterhout, L., & Mobley, L. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, 34, 739–773.
- Parker, D., Du, A., & Phillips, C. (2014). Time heals semantic illusions, but not syntactic illusions. *Paper presented at the 2014 CUNY Conference on Human Sentence Processing*. Columbus, OH: The Ohio State University.
- Pearlmutter, N. J. (2000). Linear versus hierarchical agreement feature processing in comprehension. *Journal of Psycholinguistic Research*, 29, 89–98.
- Pearlmutter, N. J., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language*, 41, 427–456.
- Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, 36, 329–347. <http://dx.doi.org/10.1017/S0140525X12001495>.

- Reali, F., & Christiansen, M. H. (2007). Processing of relative clauses is made easier by frequency of occurrence. *Journal of Memory and Language*, 57, 1–23. <http://dx.doi.org/10.1016/j.jml.2006.08.014>.
- Rossi, S., Gugler, M. F., Hahne, A., & Friederici, A. D. (2005). When word category information encounters morphosyntax: An ERP study. *Neuroscience Letters*, 384, 228–233.
- 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. <http://dx.doi.org/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. <http://dx.doi.org/10.1016/j.brainres.2008.05.092>.
- Shen, E. Y., Staub, A., & Sanders, L. D. (2013). Event-related brain potential evidence that local nouns affect subject–verb agreement processing. *Language and Cognitive Processes*, 28, 498–524.
- Solomon, E. S., & Pearlmuter, N. J. (2004). Semantic integration and syntactic planning in language production. *Cognitive Psychology*, 49, 1–46.
- Staub, A. (2009). On the interpretation of the number attraction effect: Response time evidence. *Journal of Memory and Language*, 60, 308–327.
- Staub, A. (2010). Response time distributional evidence for distinct varieties of number attraction. *Cognition*, 114, 447–454.
- Tanner, D., & Van Hell, J. G. (2014). ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia*, 56, 289–301. <http://dx.doi.org/10.1016/j.neuropsychologia.2014.02.002>.
- Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 12530–12535. <http://dx.doi.org/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 event-related potential investigation. *Studies in Second Language Acquisition*, 27, 173–204.
- Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preferences from garden-paths. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 19, 528–553.
- Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 443–467. <http://dx.doi.org/10.1037/0278-7393.31.3.443>.
- 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.
- Van Dyke, J. A., & McElree, B. (2006). Retrieval interference in sentence comprehension. *Journal of Memory and Language*, 55, 157–166.
- Van Dyke, J. A., & McElree, B. (2011). Cue-dependent interference in comprehension. *Journal of Memory and Language*, 65, 247–263.
- Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology*, 83, 176–190. <http://dx.doi.org/10.1016/j.ijpsycho.2011.09.015>.
- Vigliocco, G., Butterworth, B., & Garrett, M. (1996). Subject-verb agreement in Spanish and English: Differences in the role of conceptual constraints. *Cognition*, 61, 261–298.
- Vigliocco, G., Butterworth, B., & Semenza, C. (1995). Constructing subject-verb agreement in speech: The role of semantic and morphological factors. *Journal of Memory and Language*, 34, 186–215.
- Vigliocco, G., & Nicol, J. (1998). Separating hierarchical relations and word order in language production: Is proximity concord syntactic or linear? *Cognition*, 68, 13–29.
- Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, 61, 206–237.
- Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*, 16, 1272–1288. <http://dx.doi.org/10.1162/0898929041920487>.
- 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. <http://dx.doi.org/10.1016/j.neuroimage.2012.04.054>.
- Xiang, M., Dillon, B., & Phillips, C. (2009). Illusory licensing effects across dependency types: ERP evidence. *Brain and Language*, 108, 40–55.