Original Articles

Negative polarity illusions and the format of hierarchical encodings in memory

Dan Parker [a,*], Colin Phillips [b,c]

[a] Linguistics Program, Department of English, College of William & Mary, United States
[b] Department of Linguistics, University of Maryland, United States
[c] Language Science Center, University of Maryland, United States

ABSTRACT

Linguistic illusions have provided valuable insights into how we mentally navigate complex representations in memory during language comprehension. Two notable cases involve illusory licensing of agreement and negative polarity items (NPIs), where comprehenders fleetingly accept sentences with unlicensed agreement or an unlicensed NPI, but judge those same sentences as unacceptable after more reflection. Existing accounts have argued that illusions are a consequence of faulty memory access processes, and make the additional assumption that the encoding of the sentence remains fixed over time. This paper challenges the predictions made by these accounts, which assume that illusions should generalize to a broader set of structural environments and a wider range of syntactic and semantic phenomena. We show across seven reading-time and acceptability judgment experiments that NPI illusions can be reliably switched "on" and "off", depending on the amount of time from when the potential licensor is processed until the NPI is encountered. But we also find that the same profile does not extend to agreement illusions. This contrast suggests that the mechanisms responsible for switching the NPI illusion on and off are not shared across all illusions. We argue that the contrast reflects changes over time in the encoding of the semantic/pragmatic representations that can license NPIs. Just as optical illusions have been informative about the visual system, selective linguistic illusions are informative not only about the nature of the access mechanisms, but also about the nature of the encoding mechanisms.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Successful language comprehension routinely requires establishing dependencies between non-adjacent words and phrases. These dependencies are subject to diverse syntactic, semantic, and discourse constraints, and often rely on memory access mechanisms to recover the appropriate information from the encoding of the previous context. For instance, to relate the verb *were* in (1) to its subject for subject-verb number agreement, memory access mechanisms must recover the encoding of the plural subject noun *paintings*, and avoid interference from similar information in structurally irrelevant locations, such as the non-subject plural noun *curators*.

(1) The paintings that impressed the curators were recently sold at auction.

Speakers are typically highly accurate in retrieving the appropriate information from memory for dependency formation, but a growing number of studies have reported that memory retrieval in language comprehension is sometimes susceptible to interference from structurally irrelevant items (Drenhaus, Saddy, & Frisch, 2005; Gordon, Hendrick, & Johnson, 2001; Lewis, Vasishth, & Van Dyke, 2006; Van Dyke & McElree, 2006, 2011; Wagers, Lau, & Phillips, 2009). In this paper, we focus on interference effects that have been argued to trigger 'linguistic illusions'. Linguistic illusions are cases where speakers appear to accept incoherent or ungrammatical sentences during the early stages of comprehension, but judge those same sentences as unacceptable after more reflection. These effects can arise during the comprehension of linguistic dependencies, and have been presented as evidence that comprehenders temporarily consider ungrammatical linguistic dependencies (Pearlmutter, Garnsey, & Bock, 1999; Vasishth, Brüssow, Lewis, & Drenhaus, 2008; Wagers et al., 2009; Xiang, Dillon, & Phillips, 2009).

The current study examines the cause of linguistic illusions in order to diagnose the cognitive mechanisms for encoding and accessing linguistic information in memory. Linguistic illusions

* Corresponding author at: Linguistics Program, College of William & Mary, P.O. Box 8975, Williamsburg, VA 23187, United States.
E-mail address: dparker@wm.edu (D. Parker).

could, in principle, reflect either an error in how we mentally encode structured linguistic representations in memory, or an error in how we access information in those representations later. Many recent accounts have argued that linguistic illusions are the product of faulty memory access processes, with the additional assumption that the encoding of the sentence remains fixed over time (e.g., Vasishth et al., 2008; Wagers et al., 2009; see also Lewis & Phillips, 2015, and Phillips, Wagers, & Lau, 2011). Importantly, these accounts predict that linguistic illusions should generalize to a broad range of syntactic and semantic environments.

In contrast, we report the results from seven reading-time and acceptability judgment experiments that challenge our current understanding of how linguistic illusions arise. We show that one type of illusion, which depends on semantic/pragmatic licensing mechanisms, shows a fleeting time profile, such that it is present or absent depending on the amount of time from when the potential licensor is processed until the licensee is encountered. But we find that the same profile does not arise for another type of illusion, which depends on a morphosyntactic licensing mechanism. These results are unexpected under existing accounts, which predict that illusions should be rather general. We take these results to suggest that the encoding of emerging semantic representations is not fixed, as previously assumed, but rather, changes over time. Up to now, linguistic illusions have been taken to be especially informative about the access mechanisms used in language comprehension. Here, we argue that they are also informative about the nature of the encoding mechanisms.

1.1. Linguistic illusions

One type of linguistic dependency that is highly susceptible to illusory licensing involves subject-verb agreement (Clifton, Frazier, & Deevy, 1999; Dillon, Mishler, Slogget, & Phillips, 2013; Pearlmutter et al., 1999; Staub, 2009; Tanner, Nicol, & Brehm, 2014; Wagers et al., 2009). Subject-verb number agreement in English and many other languages is subject to specific structural and morphological requirements: the number feature of the verb or auxiliary verb must agree with the number feature of the subject. Agreement attraction arises when comprehenders are temporarily misled during agreement resolution by a feature-matched item that is not the subject of the verb. For instance, Wagers et al. (2009) examined the comprehension of subject-verb agreement in sentences like (2) using self-paced reading and speeded acceptability judgments.[1] The sentences in (2) are ungrammatical due to the number mismatch between the plural verb and its subject. Comprehenders are typically highly sensitive to such errors. However, Wagers and colleagues found that the presence of a plural distractor (e.g., *cabinets* and *runners*) reduced the reading time disruption associated with the number mismatch and increased acceptability judgments, relative to the singular distractor condition. The eased processing and increased acceptability suggests that comprehenders were misled by the structurally irrelevant plural distractor, giving rise to an illusion of acceptability.

(2) a. *The key to the *cabinet(s)* unsurprisingly <u>were</u> rusty after many years of disuse.
　　b. *The *runner(s)* who the driver <u>see</u> each morning always wave.

Agreement illusions are not simply cases of proximity concord (e.g., Quirk, Greenbaum, Leech, & Svartvik, 1985) or local coherence (e.g., Tabor, Galantucci, & Richardson, 2004), as illusions are observed when the distractor does not intervene between the verb and its subject, as in (2b). Nor is the effect simply a consequence of dialectal variation, as speakers agree on the unacceptability of sentences like (2) when given ample time, and all speakers are prone to the illusion.

Furthermore, the illusion cannot reflect misrepresentation or faulty encoding of the subject, as has been previously claimed (e.g., Eberhard, Cutting, & Bock, 2005). If comprehenders simply misrepresented the number feature of the subject, we might expect them to experience "illusions of ungrammaticality", where sentences with grammatical agreement would be misperceived as ungrammatical. However, comprehenders generally do not experience illusions of ungrammaticality (Wagers et al., 2009). Wagers et al. (2009) argued that this grammatical asymmetry is expected if agreement illusions are due to properties of faulty memory access mechanisms, rather than misrepresentation or faulty encoding of the subject phrase. Under this account, encountering a plural-marked verb triggers a retrieval process that probes all items in memory at once, in parallel, for a match to the required structural and morphological cues, e.g., [+subject] and [+plural]. In sentences with ungrammatical agreement like (2), the competition between the true subject and the distractor is relatively even, since both items only partially match the retrieval cues. On some portion of trials, the distractor may be incorrectly retrieved due to a partial-match to [+plural]. Misretrieval of the plural distractor can give the comprehender the false impression that agreement is licensed, resulting in an illusion of acceptability. In sentences with grammatical agreement, by contrast, the distractor is less likely to interfere because the fully matched subject should outcompete partial matches. Crucially, this contrast would be unexpected if comprehenders simply misrepresented or incorrectly encoded the subject phrase.

Another linguistic illusion that has sometimes been argued to reflect misretrieval involves negative polarity items (NPIs). NPIs are expressions like *ever, any, yet, lift a finger*, and *a rat's ass* that are generally only acceptable in sentences that contain a downward entailing operator in a structurally higher position (Ladusaw, 1979). Negation is the canonical example of a downward entailing operator (see Giannakidou, 2011, for a review of the contexts that license NPIs). For instance, the NPI *ever* in (3a) is licensed because it appears in the scope of the negative phrase *no diplomats*. The scope of negation for current purposes corresponds roughly to the 'c-command' domain of negation, i.e., the structural sister of the negation in a syntactic tree and any element contained within the structural sister.[2] When negation is absent, as in (3b), or is not structurally higher than the NPI, as in (3c), the NPI is not licensed.

(3) a. *No diplomats* have <u>ever</u> supported a drone strike.
　　b. *The diplomats have <u>ever</u> supported a drone strike.
　　c. *The diplomats that *no congressman* could trust have <u>ever</u> supported a drone strike.