Aligning grammatical theories and language processing models

Shevaun Lewis[1] and Colin Phillips[2]

[1] Department of Cognitive Science, Johns Hopkins University
[2] Department of Linguistics, University of Maryland

Corresponding author:

Shevaun Lewis
Department of Cognitive Science
Krieger Hall 237
Johns Hopkins University
3400 N Charles St
Baltimore, MD 21218

Phone: 410-516-6843
Fax: 410-516-8020
Email: shevaun@jhu.edu

Abstract:

We address two important questions about the relationship between theoretical linguistics and psycholinguistics. First, do grammatical theories and language processing models describe separate cognitive systems, or are they accounts of the same system at different levels of abstraction? We argue that most evidence is consistent with the one-system view. Second, how should we relate grammatical theories and language processing models to each other?

Key words:

Parsing; grammatical theories; abstraction; Cognitive architecture of language

Aligning Grammatical Theories and Language Processing Models

Shevaun Lewis[1] and Colin Phillips[2]

1 Introduction

In this paper we address two questions that arise from the long-standing divide between theoretical linguistics and psycholinguistics. We encounter renewed interest in bridging this divide, but in order to do so it is useful to explore the relation between the concerns of the two sub-fields, and what it means for them to be more closely aligned.

The first question is about the object of study. Do grammatical theories and language processing models describe separate cognitive systems, or are they accounts of the same system at different levels of abstraction? We discuss the empirical evidence that is relevant to this question, with particular attention to cases of apparent mismatches between real-time processing phenomena and grammatical phenomena. Our main goal is to clarify the contributions and the limitations of this kind of evidence. We also argue, however, that most evidence is consistent with the view that grammatical theories and language processing models describe a single cognitive system.

The second question is a more practical one: how should we relate grammatical theories and language processing models to each other? The answer depends in part on the answer to the first question. If the grammar and the language processor are independent cognitive systems, then it is difficult for findings about one to inform the other unless there is a well specified account of how they are connected. Such accounts are not widespread. On the other hand, if grammatical theories and language processing models are different abstract characterizations of the same system, then obviously we should expect them to be closely aligned. But under that view one might ask why we need two theories.

---

[1] Department of Cognitive Science, Johns Hopkins University

[2] Department of Linguistics, University of Maryland

2 The sociological divide

Psycholinguists and theoretical linguists differ primarily in whether they collect and explain "offline" or "online" data, as we will discuss in the next section. However, this difference need not have caused a deep divide between the fields. The divide is really a historical artifact. Although the sociological divide between the fields is not of interest to us, it may be helpful to review—and dismiss—the surface differences that have been so influential.

One field is inclined to refer to its accounts as "theories", the other as "models", but this apparent difference is misleading. Both groups seek to provide illuminating accounts of empirically grounded generalizations about human language viewed as a cognitive capacity. So both groups build "theories" in the well-regarded scientific sense. There is no justification for the linguists who regard psycholinguistics as a more menial pursuit that steers clear of theory development. Neither group warrants the more derogatory folk sense of "theory", referring to speculation that is divorced from evidence; that is why "theoretical" linguists are infuriated when psycholinguists suggest that their own work is distinctive in virtue of being "empirical". Both fields are very much accountable to data. The data simply tends to be rather cheaper when it involves untimed judgments.

There are obvious differences in the dominant data collection practices in the two fields: one group can do most of its work with pencil and paper; the other group depends on labs, fancy technologies, and ever more complex statistical analyses. These methodological differences may even be the source of the different concerns of the two fields. In the early days of modern linguistics, online methods were difficult to implement, whereas offline data was reliable and readily available. It was productive for linguists to constrain their domain of inquiry by restricting it to offline data. The lack of recognizable labs aligned linguistics more closely with philosophy than with psychology. By contrast, in many psychology departments it can be hard to be taken seriously if one does not have a busy lab. Psychologists have therefore focused on the online aspects of language competence that lend themselves well to lab-based research.

The artificiality of the division is clear if we compare with research on people whose language abilities are less easy to test, such as children and individuals with neurocognitive disorders or brain damage. With these populations, even untimed judgments call for laboratory skills and ingenious experimental design. And in these sub-fields we encounter no divide between groups concerned with online vs. offline phenomena, and both are considered to be psychologically respectable pursuits.

Other motivations for separation between the research agendas of linguistics and psycholinguistics include skepticism on the part of many psycholinguists about the empirical phenomena that motivate much linguistic research. Some of the more "exotic" phenomena that engage linguists are viewed as either irrelevant to what speakers do in real-time tasks, or simply as bogus effects caused by careless data collection practices (Edelman & Christiansen, 2003; Gibson & Fedorenko, 2013; Wasow & Arnold, 2005). Relatedly, one encounters skepticism on the part of some psycholinguists about the complexity or the abstractness of linguists' representations (Ferreira, 2005; Frank, Bod, & Christiansen, 2012). These concerns are deep-seated and influential, but we regard them as an unfortunate distraction from the more substantive issues facing the two fields. Large-scale testing of acceptability judgments with naïve informants overwhelmingly confirms the results of small-scale testing with expert judges, i.e., linguists (Phillips, 2010; Sprouse & Almeida, 2012; Sprouse, Schütze, & Almeida, in press), and we see little motivation for ignoring robust linguistic phenomena simply because they are not frequently attested in natural speech.

3 The heart of the divide: timing is everything

In accounts of sentence-level phenomena, which is the area that we know best, theoretical linguists and psycholinguists differ primarily in their methods and the types of phenomena they seek to explain. One group focuses on developing accounts of 'offline' data–judgments of acceptability or interpretation that are made under 'ideal' conditions, with no time limit and minimal impact from memory limitations. The other group focuses on developing accounts of ‘online’ data, typically gathered using time-sensitive measures.

Offline and online data have much in common. Grammarians and psycholinguists alike base their inferences about language competence on people’s ability to discriminate word strings in terms of acceptability or interpretation.[3] For example, a competent speaker of English can discriminate (1) and (2) based on acceptability and (1) and (3) based on interpretation. We can observe this discrimination in many different kinds of behavioral or neural responses, ranging from simple acceptability judgments to reading time measures to electrophysiological recordings.

(1) The dog follows the man.

(2) *The dog follow the man.

(3) The man follows the dog.

What distinguishes offline from online data is the time when the response is elicited. Offline responses are elicited with no time restrictions, after the presentation of a complete unit of linguistic information, such as a sentence. Online responses are elicited during limited time windows, often after the presentation of an incomplete unit of linguistic information, such as in the middle of a sentence.

The difference in timing between offline and online responses leads to another difference. Offline responses are always conscious: acceptability judgments, truth value judgments, “makes-sense” judgments, sentence completions, sentence rephrasing, or sentence repetition all require explicit attention to a task, calling for varying degrees of metalinguistic awareness. Online responses can be conscious as well, such as when they are elicited in a speeded version of any offline task. But they are often unconscious, when measured through behaviors that people do not consciously control—reaction times or

[3] Speakers demonstrate their language competence in many other ways, of course, such as in corpora of naturally occurring speech or writing, and in elicited responses that cannot be classified as “discrimination.” We focus on discrimination of word strings due to its prominence in both linguistics and psycholinguistics.

eye movements, for example—or when measured through brain activity. In cases where offline and online responses diverge, it is important to consider whether the difference could be due to conscious processes, rather than the timing of unconscious processes. Generally, however, there is little evidence that conscious control has much of an effect on offline judgments, perhaps because there are very few linguistic constraints that native speakers are consciously aware of.

Two additional methodological differences between the fields are fairly prominent, but we regard them as orthogonal to the fundamental timing difference. First is a difference in information availability. Grammarians are typically interested in speakers' ability to endorse sound-meaning pairings under ideal conditions in which both sound and meaning are known to the speaker. In contrast, psycholinguists are more commonly concerned with comprehension or production, in which only the sentence form or the sentence meaning is provided and the speaker must infer an appropriate meaning or form to pair with it. These tasks introduce questions about ambiguity resolution that are generally absent from grammarians' discussions. Sometimes it can appear that those questions are the main questions in the field: accounts of ambiguity resolution alone are often referred to as theories of sentence processing, implying that sentence processing involves little other than ambiguity resolution.

Second, there is a difference in the treatment of variability in the two fields. Grammarians tend to focus on speakers' maximal sensitivity to contrasts, whereas psycholinguists are more likely to highlight the full range of sensitivity. Consider the examples in (4), in which the auxiliary either agrees with the head noun of the subject (4a) or with a noun embedded inside the subject (4b). It is a fact that in online and offline tasks alike comprehenders sometimes discriminate and sometimes fail to discriminate the contrast between (a) and (b) (Pearlmutter, Garnsey, & Bock, 1999; Wagers, Lau, & Phillips, 2009). But whereas grammarians tend to emphasize the maximal differences that speakers detect, and therefore regard (4a) as acceptable and (4b) as unacceptable, psycholinguists tend to be more interested in the variable sensitivity that speakers show to the contrast.

(4) a. The solution to the problems is likely out of reach.

b. The solution to the problems are likely out of reach.

But questions of ambiguity and variability are logically independent of issues of timing. It is quite possible to investigate ambiguity resolution and variable sensitivity in offline judgments. It is equally feasible to abstract away from ambiguity and variability in online measures—as we do in our discussion of empirical evidence in §4.3.

Now that the two fields have collected substantial bodies of reliable offline and online data, we must confront the theoretical and practical problem of reconciling the claims that are made based on each kind of data. The theoretical problem is our first question—whether grammatical theories and language processing models describe separate cognitive systems, or the same system at different levels of abstraction. The practical problem is our second question—how should we connect grammatical theories and language processing models, and how should they inform one another (if at all)?

4 One language system, or two?

Most researchers would agree that slow, conscious offline responses reflect implicit "knowledge" of grammar. What people disagree on is the role that the grammar plays in normal language use, in the mechanisms for speaking and understanding. We next outline two views on this question and the motivations for each before discussing what we regard as the most relevant empirical evidence.

4.1 The two-system hypothesis

Under the *two-system* hypothesis, the cognitive system responsible for the grammar is separate from the system(s) responsible for language processing. Under this view, the grammar system is usually thought of as a static body of knowledge—rules for combining lexical items—whereas the language processing system is a set of procedures for comprehension and production. The properties of the grammar system are assumed to be more clearly revealed in offline data, and the language processing system in online data.

There are a number of different motivations for the two-system hypothesis, and they vary in terms of how amenable they are to empirical scrutiny. It is certainly not the case that all those who endorse the two-system view would agree with all of the motivations; here we simply attempt to marshal a range of different arguments.

One motivation derives from claims about the function of human language and its evolutionary origins. The suggestion is that the core of the human language capacity is an ability to productively and recursively combine concepts. This ability to store and manipulate complex representations could have conferred a substantial evolutionary advantage, independent of the ability to use it for communication. Proponents of this view emphasize that language is often used for non-communicative purposes. Some have argued that this function of language would have had more immediate adaptive benefits, since it does not depend on the availability of conversation partners (Berwick, Friederici, Chomsky, & Bolhuis, 2013; Jacob, 1977). There also is interesting evidence that language confers cognitive benefits independent of communication. For example, Spelke has argued that language allows humans to create representations that integrate information from different cognitive domains. For example, the thought "to the left of the blue wall" integrates representations of geometry and color. Humans can represent both of these concepts with lexical items, and the grammar allows us to combine them. Spelke has argued that the ability to encode such notions appears to depend on the ability to encode it using language, based on evidence from cognitive development in children, from dual task paradigms in adults, and from cases of extreme language deprivation (Hermer & Spelke, 1996; Hermer-Vazquez, Spelke, & Katsnelson, 1999; Hyde et al., 2011; Spelke, 2003). Similarly, language has been argued to confer advantages in domains such as reasoning about false beliefs (de Villiers, 2007; de Villiers & de Villiers, 2009; de Villiers & Pyers, 2002) and representing exact quantities (Condry & Spelke, 2008; Pica, Lemer, Izard, & Dehaene, 2004).

Under this view, the grammar system could have been shaped by the need for a flexible, productive means to represent complex thoughts. Mechanisms for translating these internal representations into a linear format for communication might have been separately shaped by the need for efficiency in the face of the time and resource constraints on real-time language use.

The idea that the language processing system (if not the grammar system) is optimized for communication has led some researchers to suggest that comprehension and production mechanisms should build the simplest representations possible that allow for effective communication. Proponents of this view emphasize the fact that most utterances have simple structures that do not require reference to complex grammatical constraints. They argue that people rarely need to consult their grammatical knowledge, because simple heuristics are sufficient for most tasks (Ferreira, Ferraro, & Bailey, 2002; Ferreira & Patson, 2007).

The "evolutionary" motivation for the two-system hypothesis can be summarized as follows. Representing complex thoughts and achieving effective communication are two very different kinds of goals. A complex hierarchical grammar system is a good solution for representing complex thoughts, but less good for efficient linear externalization for communication. By contrast, a simpler, more linear system is efficient for communication, but less efficient for representing complex thoughts. Given the different pressures associated with these two different functions, it may be that they are best implemented with different kinds of systems.

A second kind of motivation for the two-system view is more closely tied to observable empirical phenomena. It is a fact that people often produce and easily interpret sentences that are deemed unacceptable in traditional offline acceptability measures. For example, expressions with resumptive pronouns like (5) are generally judged to be unacceptable by native speakers of English. Nevertheless, native speakers often produce such sequences, and listeners have little trouble understanding the intended meaning (Zukowski & Larsen, 2004). The two-system hypothesis provides a natural explanation for this contrast: if comprehension and production operate independently of the grammar, there is no reason to suppose that they should be derailed by ungrammaticality.

(5) *The man who the caterpillar is falling on his head.

In section 4.3 we discuss in more detail the status of the various misalignments between online and offline measures of language abilities.

A third motivation for the two-system hypothesis appeals to arguments from the 1960s and 1970s about the irrelevance of transformational grammar to the concerns of psycholinguists, typically under the heading of the demise of the Derivational Theory of Complexity (DTC). The DTC was a briefly popular idea that the complexity of a sentence's transformational derivation should predict its perceptual complexity (Miller & Chomsky, 1963). After some initial enthusiasm and promising results (e.g., Miller & McKean, 1964), a consensus emerged that DTC was not fruitful (Fodor, Bever, & Garrett, 1974; Levelt, 1974). Strangely, the finding that transformations are not unique or dominant contributors to perceptual complexity led to the widespread conclusion that transformations are not cognitively "real" (Bresnan, 1978; Sag & Wasow, 2011), and contributed to arguments for the two-system hypothesis (Townsend & Bever, 2001). We are skeptical of this conclusion, and do not regard it as a strong argument for the two-system hypothesis (Phillips, 1996; Phillips, 2013a). A modern instantiation of this kind of argument could point to the tension between the bottom-to-top derivation featured in current syntactic theories and the necessarily left-to-right—and thus largely top-to-bottom—direction of real-time sentence processing. We also find this argument to be unpersuasive (Phillips & Lewis, 2013).

A fourth motivation for the two-system view is that comprehension and production are plausibly served by distinct task-specific systems, and yet they reliably exploit the same words, rules, and constraints. This observation motivates a third system that is the source of the commonalities between comprehension and production systems, i.e., the grammar. There is little evidence on the distinctness of systems for comprehension and production, as the two tasks have been studied largely independently. But to the extent that the mechanisms proposed in the comprehension and production literatures differ from one another, this provides a serious argument for the two-system hypothesis.

The most important challenge for the two-system hypothesis is explaining how the language processing system and the grammar system interact. Some interaction between the two systems is necessary to explain how offline judgments come to reflect the properties of the grammar, even though they are mediated by comprehension or production mechanisms. Close interaction between the two systems is also necessary to explain why the outputs of comprehension and production look so similar to the representations licensed by the grammar. If language use were really independent of the grammar, there would be no need for heuristic rules to approximate grammatical constraints. We know of no linguistic or psycholinguistic instantiation of the two-system view that addresses these problems by providing an explicit theory of how the two systems interact with each other. Some initial psycholinguistic suggestions are provided by Townsend and Bever (2001), although they do not begin to address the grammatical richness of online processes, and hence they fall far short of an account of how the human parsing system might exploit detailed grammatical constraints. Meanwhile, in computer science there are well-specified methods for translating a grammar into a corresponding parsing device (Aho, Lam, Sethi, & Ullman, 2006; Grune & Jacobs, 2008; for an application to minimalist grammars see Stabler, 2013). However, these methods presuppose a highly transparent mapping from grammar to parser, and they may be understood as relating different levels of analysis, as in a one-system approach, rather than relating independent cognitive systems.

4.2 The one-system hypothesis

Under the *one-system* hypothesis, there is only one cognitive system for language, and it is suitable for real-time comprehension and production. Under this view, the grammar is an abstract description of the representations that this cognitive system builds.

Just as the two-system view is motivated in part by an intuition about the function and evolution of language, so the one-system view is sometimes motivated by a different intuition about "what language is for". One might imagine that the capacity for language developed under simultaneous pressures to represent complex thoughts and to externalize them for communication. Under this view, the grammar would be shaped by the need for representations that can both encode complex thoughts and be transmitted through a serial medium. Although many of the complex properties of grammars do not make sense when considering information capacity or externalization alone, they may represent one solution to the problem of satisfying both needs simultaneously. This view seems to be accepted either explicitly or implicitly by most people who study language evolution and take syntax seriously (e.g. Bickerton, 2003; Pinker & Bloom, 1990). Perhaps relevant to this claim, even non-communicative uses of language seem to make use of externalization mechanisms. For example, the cross-domain representations that Spelke claims are parasitic on linguistic abilities appear to be available only when the person knows the external (phonological) properties of the relevant lexical items. Children who understand the difference between right and left, but do not know which word applies to which direction, are unable to use a concept like "to the left of the blue wall" to plan actions (Hermer-Vazquez, Moffet, & Munkholm, 2001).

The one-system hypothesis of course provides a straightforward explanation for the similarity between online and offline language responses: they both come from the same system. Under this view, online and offline responses represent different "snap-shots" of processes that take some amount of time to complete, and grammatical theories and language processing models are characterizations of different outputs of those processes, stated at different levels of description.

The one-system hypothesis certainly allows for divergence between online and offline responses. The real-time mechanisms that implement the grammar may not be perfectly suited to the task, especially when they recruit domain-general resources like working memory and cognitive control. Under time and resource limitations, these mechanisms may produce unintended outputs. Since these unintended outputs can be regarded as errors rather than features of the system, they could be overlooked in the higher-level descriptions of the system that grammarians provide.

However, the fact that the language processing system might be error-prone does not give free license to maintain a one-system hypothesis in the face of arbitrary mismatches between what the language processor constructs and what the grammar licenses. The most important challenge for the one-system hypothesis is to provide an explanation of how and why real-time language processes sometimes give rise to representations that are not licensed by the grammar. It is easy to provide post-hoc accounts of differences between online and offline responses in particular cases. A more convincing one-system theory should be able to systematically predict where these mismatches occur. In light of the importance of this concern, we turn next to specific cases of alignment and misalignment between online and offline responses to assess whether they are systematic and predictable within a one-system approach.

4.3 Empirical arguments: alignment and misalignment

The empirical evidence that we focus on speaks to whether the linguistic representations that are built in the earlier stages of real time processing match those that are motivated by offline measures such as untimed acceptability judgments. In other words, do parsing and production systems build representations that are licensed by the grammar?

4.3.1 Alignment

The closer the alignment between the representations tracked by online and offline measures, the more feasible it is to maintain a one-system view. Although close alignment is also compatible with a two-system view, it cannot be explained or predicted without an explicit theory of how the two systems interact.

A growing body of evidence suggests that the representations built during online language processing are usually constrained in the same way as those licensed by the grammar. Online measures often show rapid detection of grammatical anomalies and avoidance of ungrammatical parses or interpretations.

In anomaly-detection paradigms, especially the sizeable literature based on event-related brain potentials (ERPs), it is routine for grammatical anomalies to be detected within a few hundred milliseconds (Friederici, Pfeifer, & Hahne, 1993; Neville, Nicol, Barss, Forster, & Garrett, 1991; Osterhout & Holcomb, 1992; for reviews see Kaan, 2007; Sprouse & Lau, 2013). ERP responses typically track the same fine-grained degrees of anomaly measured in offline tasks (e.g., Nevins, Dillon, Malhotra, & Phillips, 2007). In fact, rapid detection of grammatical anomalies is so routine that it is newsworthy when anomalies are not immediately registered (negative polarity: Vasishth, Brüssow, Lewis, & Drenhaus, 2008; Xiang, Dillon, & Phillips, 2009; subcategorization: Wagers & Phillips, in press; agreement: Wang, Bastiaansen, Yang, & Hagoort, 2012).

Studies of long distance dependencies demonstrate that online processes generally avoid constructing dependencies that would be unacceptable in offline judgments. For example, the interpretation of a reflexive pronoun requires a dependency between the reflexive and an antecedent from earlier in the sentence. Studies on the online interpretation of reflexives (e.g., 'himself', 'herself', 'themselves') in comprehension have tested whether the parser considers only antecedents that would be acceptable in offline judgments: c-commanding clausemates (Binding Principle A: Chomsky, 1981). That is, when interpreting (6), does the parser ever consider 'Jonathan' as a potential antecedent for 'himself'?

(6) The surgeon who treated Jonathan had pricked himself with a used syringe needle.

Most studies suggest that grammatically illicit antecedents are not considered, based on evidence from cross-modal priming (Nicol & Swinney, 1989), eye-tracking during reading (Dillon, Mishler, Sloggett, & Phillips, 2013; Sturt, 2003), self-paced reading (Badecker & Straub, 2002, experiments 4-5; Clifton, Frazier, & Deevey, 1999), visual world eye-tracking (Clackson, Felser, & Clahsen, 2011), and

ERPs (Xiang et al., 2009). Thus, online and offline responses both indicate the same set of candidate antecedents.[4]

The interpretation of *filler-gap dependencies* in *wh*-questions and relative clauses requires associating a displaced element like a *wh*-word with a "gap" later in the sentence. Many studies have tested whether comprehenders ever attempt to associate fillers with gap positions that would not be acceptable in offline judgments, i.e., those inside syntactic "islands". That is, when interpreting (7), does the parser ever consider the illicit gap site marked with an asterisk, taking 'the book' to be the object of 'wrote'?

(7) We like the book that the author who wrote * unceasingly and with great dedication saw __ while waiting for a contract.

Most studies have found that the parser respects island constraints that are observed offline (Bourdages, 1992; Neville et al., 1991; Omaki & Schulz, 2011; Phillips, 2006; Stowe, 1986; Traxler & Pickering, 1996; Wagers & Phillips, 2009; Yoshida, Aoshima, & Phillips, 2004). Other studies have found that comprehenders readily detect the boundaries of islands while parsing filler-gap dependencies (Kluender & Kutas, 1993; McElree & Griffith, 1998; Neville et al., 1991).[5]

---

[4] Some studies have challenged the generality of these conclusions about reflexives (Badecker & Straub, 2002, experiment 3; King, Andrews, & Wagers, 2012; Patil, Vasishth, & Lewis, 2011; Runner, Sussman, & Tanenhaus, 2006), but it is clear that the parser is able to ignore at least some grammatically irrelevant material in memory access.

[5] A couple of studies have reached more equivocal conclusions (Clifton & Frazier, 1989; Pickering, Barton, & Shillcock, 1994), but have stopped short of concluding that the parser is insensitive to island constraints. Some further studies have argued that the parser is able to construct island-violating filler-gap dependencies when other parses are not available (Freedman & Forster, 1985; Hofmeister & Sag, 2010), but these findings do not conflict with the findings about island effects in active dependency formation. We discuss these apparent misalignments between online and offline responses in the next section.

Similar to the findings on island constraints, studies on the processing of *backwards anaphora* or *cataphora,* interpretive dependencies in which a pronoun precedes its antecedent, have tested whether comprehenders consider potential antecedents that are grammatically illicit due to Binding Principle C, a cross-linguistically robust constraint that excludes pronouns that c-command their antecedent (Chomsky, 1981; Büring, 2005). Results from multiple languages indicate that the parser respects Principle C (English: Cowart & Cairns, 1987; Kazanina, Lau, Lieberman, Yoshida, & Phillips, 2007; Japanese: Aoshima, Yoshida, & Phillips, 2009; Russian: Kazanina & Phillips, 2010; Dutch: Pablos, Ruijgrok, Doetjes, & Cheng, 2012). For example, comprehenders never take 'Kathryn' to be a potential antecedent for 'she' in sentences like (8).

(8) Because last semester she was taking classes full-time while Kathryn was working two jobs to pay the bills, Erica felt guilty.

Taken together, these findings and many others indicate that online responses exhibit fine-grained sensitivity to many of the constraints identified by grammarians using offline measures. They do not lend support to the notion of a comprehension system that deploys rough-and-ready mechanisms that sacrifice grammatical detail for efficiency, and as such they are encouraging for a one-system view. But we also find many cases where online and offline responses appear to diverge, which we turn to next.

4.3.2 Misalignment

There are a number of interesting cases of misalignment in the literature. The more we see arbitrary mismatches between online and offline representations, the more motivation we have for a two-system view, in which comprehension and production mechanisms may frequently make use of task-specific rules that differ substantially from the grammar. Misalignments may be consistent with a one-system view, but only if the explanation for those mismatches is based on general properties of the language processing system.

We argue that the observed misalignments plausibly arise from limitations of general-purpose mechanisms—particularly memory access and control mechanisms—that are used to implement language-specific processes. We do not observe the diverse and arbitrary misalignments that would be expected under a two-system view. We discuss four categories of misalignment: garden paths and revision failures, resource overload, consequences of memory access mechanisms, and internal stages of computation. We first discuss each type of mismatch individually, and then consider how they may be related to one another.

Garden paths and revision failures

Misalignment between online and offline responses can arise in comprehension in cases where the incrementality of the input to the system ends up misleading the parser. A notorious and uncontroversial example of this is garden path sentences like (9) (Bever, 1970). Readers or listeners initially perceive the sentence to be ungrammatical, but with enough time they can eventually recognize that it does have an acceptable parse. This misalignment between online and offline responses to the sentence does not suggest that parsing ignores grammatical constraints. Quite the contrary: it is the parser's zeal in pursuing a grammatical and highly likely syntactic structure (with 'horse' as the subject of 'raced') that increases the difficulty of considering an alternative structure.

(9) The horse raced past the barn fell.

Comprehenders not only misjudge the acceptability of garden-path sentences, but also sometimes maintain the interpretation associated with their initial parse. For example, in (10) the noun phrase 'the baby that was small and cute' is likely to be initially parsed as the direct object of the verb 'dressed', but it must later be reanalyzed as the subject of the verb 'spit up'. After reading a sentence like (10), speakers answer "yes" about 60% of the time to the question, 'Did Anna dress the baby?', compared to only 12% of the time when the sentence was disambiguated with a comma or a different clause order (Christianson, Hollingworth, Halliwell, & Ferreira, 2001).

(10) While Anna dressed the baby that was small and cute spit up on the bed.

Christianson and colleagues interpret their finding as evidence that in cases of high processing load, real-time comprehension processes can give rise to "good enough" representations that are not consistent with grammatical constraints. For example, in (10) comprehenders might fail to fully reanalyze the embedded clause object as a main clause subject and end up with a "good enough" parse in which 'the baby that was small and cute' is simultaneously an argument of both 'dressed' and 'spit up'. A subsequent study, however, shows that interpretations associated with initial (mis-)analyses persist even in cases where syntactic reanalysis is relatively easy and there is no reason to suppose that the parser resorts to good-enough representations (Sturt, 2007).

The persistence of incorrect interpretations in cases like (10) is less surprising if we consider that conceptual representations are not the same as syntactic-semantic representations. It is relatively uncontroversial to assume that comprehenders incrementally update their beliefs as they parse incoming sentences. But once a parse of the sentence has been used to update the comprehender's non-linguistic representation of the event described by the sentence, the link between the parse and the updated beliefs need not be maintained. If the parse is subsequently revised, there is no straightforward way to automatically update the corresponding beliefs. The difficulty of retracting beliefs based on syntactic reanalysis can easily be seen if we expand the scale from one sentence to several hundred sentences. In the course of reading this paper you may have updated your beliefs about the nature of the language system—or at least your beliefs about our views—but you would not be able to recall how those new beliefs arose from the specific sentence structures you read. If we decided to change a sentence here and there, you might recognize the changes if they led to different beliefs, but you would be unlikely to recognize changes that merely substituted syntactic structures. Likewise, the persistence of interpretations following syntactic reanalysis need not reflect a parser-grammar misalignment, but may simply reflect the memory limitations for tracking links between linguistic representations and non-linguistic beliefs.

Notorious cases of illusory comparative sentences like (11) present an apparent mismatch between online and offline judgments, which we argue arises from a "garden path" at the semantic level. Sentences like (11) sound natural at first, but further reflection reveals that they are incoherent. The first clause establishes a comparison involving a number of entities, but there is no corresponding countable noun in the comparative clause to complete the comparison. If the initial percept of acceptability and the subsequent judgment of incoherence are the product of separate cognitive systems, then the mismatch is unsurprising, and therefore these cases provide potential evidence for a two-system view (Townsend & Bever, 2001).

(11) More people have been to Russia than I have.

Closer investigation suggests that illusory comparatives might be more akin to garden path phenomena, reflecting detailed use of grammatically licit semantic options, rather than reflecting the operations of a grammar-independent heuristic analyzer (Wellwood, Pancheva, Hacquard, & Phillips, submitted). English allows sentences that have the form of assertions about quantities of individuals to be understood as assertions about quantities of events. For example (12a) is most naturally understood as a claim about the number of events in which a car crossed the George Washington Bridge, and not about the number of distinct cars that crossed the bridge. Real-world knowledge tells us that the total probably includes many cars that crossed the same bridge 10 times per week, accounting for as many as 500 bridge crossings per year, but speakers do not consider this to falsify the sentence. (12b) shows that this use of the form of individual quantification to express event quantification extends to comparatives. But (12c) shows that the use of this strategy appears to be contextually constrained: in situations where the tracking of distinct individuals is likely to be relevant to the assertion, the event quantification interpretation is less available. (12b) would be regarded as an accurate description of a situation where the same set of vehicles made more bridge crossings, but (12c) would be regarded as a misleading description of a scenario where more hamburgers were eaten by the same number of individuals.

(12) a. 106 million cars crossed the George Washington Bridge in 2007.

b. More cars crossed the George Washington Bridge in 2007 than in any other year.

c. More Americans ate at McDonalds last year than in any other year.

Wellwood and colleagues argue that illusory comparatives like (11) induce semantic garden path effects, in which speakers initially interpret the first clause as an instance of event quantification, an option that is independently allowed in English. In support of this, they show in a series of judgment studies that speakers are less susceptible to comparative illusions when the predicate in the initial clause is non-repeatable, i.e., it cannot be carried out multiple times by the same person, and hence disfavors an event quantification reading. What is left unexplained under this account is why the semantic garden path is hard to detect. In the case of syntactic garden paths like (9), speakers immediately recognize when their initial parse runs into difficulty, but the corresponding failure is less apparent in sentences like (11). The difficulty of detecting the failure could be due to the fact that an overt sortal (i.e., countable noun) is not required in the comparative clause, as examples like (11b) show. Or it could be that recognition of the unsuitability of the uncountable subject noun ('I' in (11)) for the comparison is masked by the task of identifying an ellipsis antecedent for the comparative clause, a process that is successful. Or there could be other reasons for the failure. Summarizing, it is becoming clearer why comparative illusions are triggered: they are initiated due to an entirely legitimate semantic option in English. What remains unclear is why the illusions are not readily detected.

Taken together, the large literature on syntactic garden paths and reanalysis processes in comprehension, and the far smaller literature on semantic garden paths, suggest that parsing mechanisms are highly sophisticated, and that they take advantage of detailed grammatical constraints, together with other sources of information, in order to deduce structure from linear sequences of words. Misalignments between online and offline responses arise when partial information is misleading. Even with full knowledge of the space of grammatical possibilities, reanalysis is often difficult because of the limitations of memory and control mechanisms. These phenomena do not motivate a two-system view.

Processing overload

A second type of misalignment between online and offline linguistic representations arises in situations where the comprehension system's resources are overloaded to the point that it is difficult to arrive at any parse for the incoming sentence, as in the well-known examples of unparsable center embedded sentences like (13). While these cases look like misalignments—the parser fails to construct a representation licensed by the grammar—they are not misalignments at the level of representational capacity, and thus not of interest to us. The parser can in principle build structures with multiple center embedding, and succeeds in doing so in easier cases like (14).

(13) The student who the professor that the counselor recommended disappointed appealed the grade.

(14) Every student that the professor you work with wrote a recommendation letter for ended up getting a job.

More relevant to our current concerns are cases where there is a conflict between the representations constructed online and in time-unlimited tasks—for example, when online comprehension processes seem to allow sentences that are recognized as ill-formed in offline tasks. An example of this can be found by comparing (13) with (15). Whereas (13) is grammatical but hard-to-parse, (15) is simply ungrammatical: it contains three clauses but only two verbs. Yet a number of studies have found that the ungrammatical (15) is judged as more acceptable than the grammatical (13) (Frazier, 1985; Gibson & Thomas, 1999; Gimenes, Rigalleau, & Gaonach, 2009).

(15) * The student who the professor that the counselor recommended appealed the grade.

The relative acceptability of (15) presents an interesting case of misalignment, and such effects have been argued to motivate a two-system architecture for language (Trotzke, Bader, & Frazier, 2013). Like the comparative illusion in (11), the contrast between initial acceptance of (15) and its status as uncontroversially ungrammatical fits with the view that online comprehension and associated percepts of acceptability are implemented by a system that is distinct from—and imperfectly related to—the grammar. However, it may be premature to regard such examples as evidence for a two-system view. First, the fact that (15) is judged as more acceptable than (13) is a relative judgment, which does not entail that speakers reliably judge it to be a well-formed sentence of English. That is reassuring, since speakers are surely unable to report a well-formed interpretation for (15), as it does not have one. A plausible account of how speakers overlook the missing verb in (15) is that the second and third subject NPs ('the professor', 'the counselor') are successfully associated with the most deeply embedded verb ('recommended'), at which point speakers shift their attention to the needs of the one remaining disconnected NP ('the student'), while failing to notice that the second NP ('the professor') needs to be the subject of a further verb (Whitney, 2004). The unsatisfied dependency may fail to generate an error signal because it has simply been "forgotten", leading to a percept of acceptability. Thus, this misalignment could arise because of limitations on memory and control mechanisms. There is no need to assume two distinct structure building systems.

Properties of memory access mechanisms

We attributed the first two types of apparent misalignments to simple limitations of the capacity of memory access and control mechanisms. Even when capacity is not a problem, the properties of domain-general mechanisms can lead to other kinds of misalignment. In comprehension and production, speakers frequently fail to notice unacceptable number marking on the verb when some NP other than the subject has features that match with the verb (Bock & Miller, 1991; Clifton, Frazier, & Deevy, 1999; Pearlmutter, Garnsey, & Bock, 1999), as in (16).

(16) *The key to the cabinets are missing.

Three types of evidence suggest that these illusions of acceptability should not be regarded as instances of "proximity concord", attributable to memory capacity limitations or to "good enough" representations more concerned with linear proximity than hierarchical relations (e.g., Francis, 1986; Quirk, Greenbaum, Leech, & Svartvik, 1985). First, production evidence from more complex NPs shows that nouns that are closer to the agreeing verb are less disruptive than nouns that are closer to the true subject noun (Bock & Cutting, 1992; Franck, Vigliocco, & Nicol, 2002). Second, nouns that are more distant from the verb than the true subject noun, as in (17), can induce agreement illusions in both production (Bock & Miller, 1991) and comprehension (Staub, 2009; Staub, 2010; Wagers, Lau, & Phillips, 2009). Third, people rarely experience the opposite phenomenon, i.e., illusions of ungrammaticality, in sentences like those in (18), although we might expect such illusions to be equally frequent if speakers are simply ignoring structure in computing agreement[6].

(17) *The musicians who the reviewer praise so highly will probably win a Grammy.

(18) a. The keys to the cabinet are on the table.

b. The musicians who the reviewer praises so highly will probably win a Grammy.

The full pattern of results can be explained by appealing to independently motivated properties of memory retrieval mechanisms, under the hypothesis that agreement relations are implemented in real-time processing by retrieving the subject at the point of the verb using a parallel, cue-based memory access mechanism (Wagers et al., 2009). Parallel, cue-based memory access works by simultaneously probing all objects in memory for their match to particular featural cues—[+subject] and [+plural], for example. This kind of mechanism is less affected by distance between the subject and the verb, but would be subject to interference from 'partial matches' (Lewis & Vasishth, 2005; McElree, Foraker, & Dyer, 2003). In the sentences that give rise to illusions of grammaticality, each NP preceding the verb satisfies only one of the two search criteria: the true subject is [+subject] and [-plural], while the 'attractor' is [-subject] and [+plural]. In such configurations, the retrieval process launched by the verb may frequently retrieve the wrong NP rather than the right one, leading to an illusion of grammaticality in some cases. Illusions of ungrammaticality are predicted to not occur, since in those cases the attractor noun is a poor match to the retrieval cues. Under this account, the parser's actions are fully compatible with the grammar, in the respect that the retrieval instructions are entirely consistent with offline grammatical generalizations. The errors arise simply because the grammar's constraints are implemented within a noisy general memory architecture. If this characterization of the errors is accurate, then it is consistent with a single system hypothesis.

[6] For small-but-reliable illusions of ungrammaticality see Lago & Phillips (in prep.) and Wagers (2008).

Other examples of grammatical illusions can be captured in a similar fashion under the view that grammatically licensed operations are implemented in a noisy memory architecture. German speakers overlook certain classes of case mismatches (Bader, Meng, & Bayer, 2000) and are susceptible to case attraction effects similar to agreement attraction (Sloggett, 2013). In the domain of anaphora processing, a number of studies have reported evidence of fleeting misretrieval of grammatically inappropriate antecedents for pronouns (Badecker & Straub, 2002; Kennison, 2003).[7]

Misretrieval of partially matching items in memory has also been invoked to explain the robust illusory licensing effects that have been found for the negative polarity item (NPI) 'ever'. The contrast between (17a) and (17b) shows that 'ever' must be licensed by a negative element. The quantifier 'no' is just one among many potential licensors (Giannakidou, 2011; Ladusaw, 1996; Linebarger, 1987). The unacceptability of (17c), in which the negative element is embedded inside a relative clause, illustrates the fact that the NPI must be licensed by a c-commanding element.

(19) a. No bills [that the senators voted for] will ever become law.

b. *The bills [that the senators voted for] will ever become law.

c. *The bills [that no senators voted for] will ever become law.

Although offline judgments show that the NPI 'ever' in (17c) is unacceptable, online studies in German (Drenhaus, Saddy, & Frisch, 2005; Vasishth et al., 2008) and English (Xiang et al., 2009) consistently find that it is fleetingly treated as if it is appropriately licensed: initial responses to (17c) typically fall between responses to acceptable (17a) and unacceptable (17b). There are competing accounts of this illusion, treating it either as reflecting a partial match to memory retrieval cues, as has

[7] The empirical record on this point is mixed. Other studies of pronoun processing have argued that ungrammatical antecedents for pronouns are not considered online (Nicol & Swinney, 1989; Clifton, Kennison, & Albrecht, 1997; Chow, Lewis, & Phillips, in prep). If ungrammatical antecedents are not considered, this would of course be a case of alignment, not misalignment.

been proposed for agreement illusions (Vasishth et al., 2008), or as reflecting the over-application of a pragmatic licensing mechanism (Xiang et al., 2009). But this disagreement reflects a corresponding debate in the grammatical literature on NPI licensing, and whether it should be treated as an instance of an item-to-item dependency like agreement and anaphora, or as a case of licensing by the compositional meaning of the entire sentence (cf. Giannakidou, 2011). Therefore, the licensing mechanisms that are invoked to explain the illusions are transparently related to grammatical accounts of NPI licensing. A case for true misalignment of mechanisms arises only if we find that one type of licensing constraint is invoked online and a different type of licensing constraint is most appropriate for capturing offline judgments.

Therefore, in all of these cases the mismatch between online and offline judgments arises not because there are performance mechanisms that implement an alternative set of grammatical constraints, but rather because online mechanisms use grammatical constraints in a cognitive architecture that creates opportunities for error. The difference between immediate responses and slower responses may simply reflect the improvement with time of the signal-to-noise ratio in the responses, rather than the deployment of distinct mechanisms: if a slow judgment involves repeated attempts at retrieval in a noisy architecture, then increased time for judgment should improve grammatical accuracy. An outcome that has a 25% probability of occurrence on a single retrieval trial has a much smaller probability of being the dominant outcome over the course of multiple retrieval trials.

Internal stages of computation

A final way in which misalignments between online and offline percepts can arise is via access to internal stages of linguistic computation. Real-time linguistic computation takes some amount of time, and it is relatively uncontroversial that the computation might sometimes involve multiple steps. It is therefore possible that some sensitive experimental measures might tap into the results of intermediate steps of that computation. This can create the impression of mismatches between the representations revealed by online and offline processes, but such situations clearly should not be taken as evidence for a two-system view, as the following two examples illustrate.

The first example comes from studies on island constraints on unbounded dependencies. There is a difference between what we are able to represent in our native language and what we judge to be acceptable. For example, we can readily understand (20) and (21), despite the presence of an agreement violation and an argument structure violation, respectively.

(20) John are happy.

(21) She explained them the story.

The difference between what is representable and what is well-formed plays a larger role in some grammatical theories that distinguish a powerful generative component from a set of filters or constraints that apply to the output of the generative component. This distinction is prominent in Government-Binding (GB) theory (Chomsky, 1981), and it is a core property of Optimality Theory (Prince & Smolensky, 2004). In the past this has led to some interesting psycholinguistic arguments, in which evidence for "overgenerated" representations has been offered as evidence for specific grammatical models. Freedman and Forster (1985) used evidence from a sentence-matching task to argue that certain island constraint violations, which in GB theory are claimed to be representable but ungrammatical, patterned with well-formed sentences rather than with flat-out ungrammatical sentences. Freedman and Forster argued that these results show that their experimental task is sensitive to the class of representable sentences rather than the class of acceptable sentences. The results do not challenge the view that island constraints have rapid online effects. These experimental findings have been disputed (Crain & Fodor, 1987), but the argument remains interesting.

A different example of access to internal stages of linguistic computation comes from research by Peter Gordon and colleagues on the *repeated name penalty*, a discourse constraint that makes it infelicitous to repeat a name that is already prominent in a discourse (22a), such that use of a pronoun would be more natural. Gordon and colleagues found that ERP N400 responses are larger to the repeated name than to a non-repeated counterpart (22b), transparently reflecting the difficulty caused by violation of the discourse constraint (Ledoux, Gordon, Camblin, & Swaab, 2007; Swaab, Camblin, & Gordon, 2004). In contrast, early eye-tracking measures show the opposite pattern: repeated names are read more quickly than non-repeated names, presumably due to repetition priming (Ledoux et al., 2007). This suggests that the status of the repeated name is different at the stages of lexical access and discourse integration. It certainly reflects a mismatch between a representation that is favored at one moment and then disfavored just a couple of hundred milliseconds later. But this clearly reflects different steps in the workings of a single parsing mechanism, rather than the results of mutually incompatible systems.

(22) a. # Tom moved the desk because Tom needed room.

b. Dave moved the desk because Tom needed room.

Misalignments such as these are what we should hope to find if we have sufficiently sensitive tools for investigating language processing, as we should want to be able to look inside the stages of linguistic computation.

Summary: the argument for the one-system hypothesis

We claim that all the different cases of potential parser-grammar misalignment can be accounted for without recourse to a two-system view. This conclusion invites the objection that perhaps anything could be explained under a one-system view by invoking an ad hoc series of noise factors, garden paths, or multi-step computations in order to account for any kind of misalignment that might arise.

We acknowledge this concern, but we think that our account of the misalignments is far from ad hoc. In fact, the four different sub-types of misalignment described here fit naturally into an account of real-time linguistic computation, as outlined in (23).

(23) *Types of misalignment between online and offline responses*

i. Computations that are not yet complete ("Internal stages of computation")

ii. Computations that fail to complete, due to resource limitations ("Processing overload")

iii. Computations that complete, but inaccurately, due to noisy architecture ("Properties of memory access mechanisms")

iv. Computations that complete successfully, but that are later challenged by subsequent input ("Garden paths and revision failures")

According to this approach, online and offline representations are the product of a single structure-building system (the grammar) that is embedded in a general cognitive architecture, and misalignments between online ("fast") and offline ("slow") responses reflect the ways in which linguistic computations can fail to reflect the ideal performance of that system. The computations and their failures are all independently motivated.

First, the grammatical computations that we assume operate in real time reflect independently motivated grammatical constraints. For example, the linguistic features that are used to retrieve agreement controllers or antecedents of anaphors are the same features that govern offline acceptability. Similarly, the constraints that are used to license NPIs online should be the same constraints that govern the offline acceptability of NPIs, even if the online implementation of those constraints is noisy. Additionally, in appealing to the internal stages of linguistic computation to account for misalignments we rely on plausible claims about what those internal stages are.

Second, the properties of the general cognitive architecture are independently motivated based on non-linguistic evidence. For example, the account of 'illusions of grammaticality' as mis-retrieval of items from memory is based on independently motivated assumptions about parallel access in content-addressable memory. This is attractive, as it provides constraints on accounts of retrieval errors.

Third, it would be attractive if an account of misalignments based on resource limitations were based on independently motivated measures of individual memory resources and cognitive control abilities. This should make it possible to predict individual variation in language processing abilities. However, our theories are not yet as advanced as we would like in that regard. There is evidence that individuals with greater working memory resources or better cognitive control abilities can parse more complex sentences or more readily handle garden-path sentences (working memory: Just & Carpenter, 1992; MacDonald, Just, & Carpenter, 1992; cognitive control: Hussey & Novick, 2012; Novick et al. 2013), but more precise predictions are not yet available.

Although the sketch given here provides a schematization of the general types of online-offline misalignments that we should expect to encounter, we should ultimately be able to predict which linguistic phenomena should yield illusions and misalignments, and which phenomena should not. It should be possible to make predictions about the profile of as-yet unstudied phenomena, in English or in other languages. This is something that we have begun to do (e.g., Phillips, Wagers, & Lau, 2011), but many specifics remain poorly understood. For example, Dillon et al. (2013) show that subject-verb agreement and reflexive licensing in English exhibit sharply different susceptibility to interference from irrelevant NPs, despite being subject to very similar grammatical constraints. They capture the contrast by proposing that person/number/gender features are used as retrieval cues for subject-verb agreement, giving rise to mis-retrievals in cases of partial matches, but that those same features are only used as post-retrieval well-formedness checks in reflexive licensing, thereby avoiding mis-retrieval. This contrast does not follow straightforwardly from our proposal here (for suggestions see Dillon, 2011; Kush, 2013).

Meanwhile, under the alternative two-system view, the outlook is rather worse, given the lack of constraint on the relations between the language processing mechanisms, grammatical constraints, and general cognitive mechanisms. Under this view the degree of alignment between online and offline processes is surprising. There is less independent motivation for a general account of which linguistic phenomena should and should not yield illusions and misalignments. To our knowledge, there have been no attempts under a two-system view to construct a general theory of such phenomena.

5 Practical consequences for alignment

Our starting point is the fact that—at least in studies of adult native language abilities—linguists and psycholinguists have taken responsibility for understanding different types of phenomena. Linguists have typically focused on abilities reflected in 'offline' data, i.e., explicit judgments of meaning or well-formedness given by expert judges under ideal conditions with no time limits. Psycholinguists have typically focused on abilities reflected in 'online' data, generally consisting of implicit responses to meaning or well-formedness given by naïve participants and measured using time-sensitive techniques. Separate fields have grown up around the study of these "fast" and "slow" phenomena, and the practical demands of either type of research provide some historical motivation for the split. But these practical considerations should not be mistaken for motivations for a principled difference between the objects of study.

Above we have highlighted the motivations—empirical or otherwise—for two contrasting views of the relation between the theories that linguists and psycholinguists construct around the phenomena that they study. Under the two-system view, online and offline phenomena are the products of distinct-but-related cognitive systems: (at least) one system that is designed for efficient communication in real time, and another system that is task-neutral and not suited for real-time operation. Under this view, linguists and psycholinguists are investigating distinct neurocognitive systems. In contrast, under the one-system view, online and offline phenomena are merely different reflections of the behavior of a single neurocognitive system, which builds representations that are used in speaking and understanding, but can be described at different levels of abstraction, paying more or less attention to the system's fine-grained temporal operations. The difference between these views is rarely discussed, and we have been surprised to learn from discussions with (psycho)linguists that many researchers take one or other of these positions for granted, and assume that this is what everybody else assumes, contrary to fact.

It should be apparent from our discussion here that we share a bias for the one-system view, but under either view there are practical consequences for better connecting the two bodies of research, to which we now turn.

5.1 Alignment in a two-system architecture

The two-system view suggests a deceptively simple methodological approach. If grammatical theories and language processing models describe separate cognitive systems with different purposes, their properties need not be tightly related. Under this view, linguists who are concerned with offline data need not pay attention to online data and processing models, because they are irrelevant for describing the grammar system. Likewise, psycholinguists need not pay attention to offline data and grammatical theory, because they are irrelevant for describing language processing. Although we are not certain how many researchers would explicitly endorse this degree of separation between the fields, it describes relatively well the current amount of crosstalk between the two fields. However, this approach has at least three significant limitations.

First, it is clear that there are close connections between language processing systems and the grammar. As we have reviewed above, existing evidence shows that grammatical distinctions typically have immediate impact on online processes. Also, the grammar of one's native language must be learned via the mediation of the language processing systems. So a viable version of the two-system hypothesis requires a theory of how the two systems interact, drawing on both online and offline data. Explicit accounts of this interaction are currently hard to find.

Second, if the language processor and the grammar are distinct cognitive systems responsible for different kinds of phenomena, then both can be described at multiple levels of abstraction, and we ultimately will need detailed low-level accounts of how each is neurocognitively implemented. Regarding the grammar as a distinct task-neutral and process-neutral body of knowledge does not exempt the linguist from the task of specifying how the grammar is instantiated in the brain, or how that knowledge is consulted by other systems, such as the language processor. This is a very difficult challenge. Under the two-system view, the grammar is not an information processing system, but a body of static knowledge that other systems consult. There would not be an algorithm associated with the grammar—only an implementation. While research on the implementation of information processing systems has made significant progress in various cognitive domains, we still know very little about the implementation of 'static' knowledge systems, especially systems that share the grammar's property of specifying an unbounded class of possible compositional representations. There are no successful examples to follow.

Third, if the grammar and the language processor are distinct cognitive systems, then much caution is needed in attempts to use evidence from language processing to arbitrate between competing grammatical theories. There have been occasional waves of enthusiasm in theoretical linguistics about the prospect of using evidence from real-time phenomena to decide among alternative grammatical theories. The late 1980s saw a surge of interest in the use of online evidence to resolve a theoretical dispute between syntactic theories about the need for empty categories ('traces') created by movement operations (Chomsky, 1973; Sag & Fodor, 1994). Some studies reported effects of antecedent priming at putative trace/gap positions in parsing (Bever & McElree, 1988; MacDonald, 1989; Nicol, Fodor, & Swinney, 1994) or effects of pre-verbal construction of unbounded dependencies (Aoshima, Phillips, & Weinberg, 2004; Lee, 2004; Nakano, Felser, & Clahsen, 2002), and took these findings as online evidence that traces are 'psychologically real'. Other studies argued for the opposite conclusion (Pickering & Barry, 1991). But over time it became clear that the timing evidence was unlikely to be conclusive, because the competing grammatical theories do not make clear timing predictions (Gibson & Hickok, 1993; Phillips & Wagers, 2007). A similar discussion has emerged in recent years over the use of online evidence to decide among competing grammatical theories of ellipsis (Culicover & Jackendoff, 2005; Merchant, 2001), although this debate is limited by a lack of clear timing predictions, in just the same way as the earlier debate about traces (Phillips & Parker, in press). Semantic theory has recently witnessed a similar surge of interest in the use of online data to decide among competing accounts of quantification (Hackl, Koster-Hale, & Varvoutis, 2012; Szabolcsi, 2013) or scalar implicatures (Bott & Noveck, 2004; Breheny, Katsos, & Williams, 2006; Grodner, Klein, Carbary, & Tanenhaus, 2010; Huang & Snedeker, 2009; but see Lewis, 2013). It is certainly encouraging to see linguists drawing on diverse sources of evidence to constrain their theories, but it is difficult to use online data to assess theories that in a two-system architecture specifically disavow timing predictions and do not specify a clear link with language processing systems.

5.2 Alignment in a one-system architecture

Under a one-system view, grammatical theories and psycholinguistic models describe a single system at different levels of abstraction. This system links strings of sounds or symbols to complex conceptual representations. Grammatical theories describe the general properties of this linking function: how the linear strings are related to hierarchical structures, and how those structures relate to meanings. Psycholinguistic models describe how mental processes implement that linking function using available cognitive operations, including information about the time course of such processes and specifying how the system operates under situations of uncertainty.

A one-system architecture solves some of the challenges that we outlined for the two-system approach. It avoids the need to specify how the language processor and the grammar interact: if they are the same system, then they do not need to interact. It also avoids the need to provide a separate description of how the grammar is implemented neurocognitively—we can rely on the psycholinguists and neurolinguists for that. However, important challenges remain for this approach.

First, in a one-system architecture it is important to understand the cases of apparent misalignment between online ("fast") and offline ("slow") responses. If these truly are products of the same system, then any contrasts in the representations that are implicated by these different data types should be few in number and they should be predictable from independent cognitive constraints, or from the temporal unfolding of the system's computations. This is what we have sketched in outline in §4.3 above.

Second, in a one-system architecture where the grammar is understood as a higher-level description of the representations that a language processing system works with, we need to fill the same theoretical gap that the two-system view must also fill. If there are differences between the data obtained using fast and slower measures, then we need an explicit process model that explains how the offline judgments arise. It is insufficient for high-level grammatical theories to provide accounts only of offline judgments, and for more process-oriented psycholinguistic models to restrict their attention to accounts of phenomena that happen quickly. In other words, we need a psycholinguistic model of slow linguistic processes.

Third, in order to maintain a one-system view, in which the grammar is simply regarded as a more abstract characterization of what the language processing system does, it is necessary to show that the same mechanism can carry out comprehension and production. It is clear that comprehension and production have much in common—we speak and understand the same native language—so there must be something that is shared between the two tasks. This is captured easily in a two-system view, in which distinct language processing mechanisms for comprehension and production both draw on a separate grammatical system. But in a one-system architecture the only way to capture this commonality is to assume that comprehension and production are achieved by a single system that is put to work in different situations, where either the message is known and a corresponding external form must be found (‘production’) or where an external form is provided and a corresponding message must be inferred (‘comprehension’). Conflating comprehension and production systems is not straightforward, although some interesting initial attempts have been made (Kempen, in press; Kempen, Olsthoorn, & Sprenger, 2012).

Fourth, the remarks about the difficulty of using online data to arbitrate disputes among grammatical theories apply to the one-system view in a very similar fashion to the two-system view. Online evidence, which generally provides information on the timing and/or difficulty of constructing specific representations, can only be theoretically decisive to the extent that it addresses theories that make clear predictions about timing or difficulty. Most grammatical theories do not do this.

5.3 A note about abstraction and ‘Marr levels’

In his landmark book on vision, Marr (1982) argues that any information processing system—including human cognitive systems—can be described at three different levels. The *computational theory* describes the purpose of the system: what the inputs and outputs are, and why. The *algorithm* describes the specific representations and operations that are used to accomplish the purpose described by the computational theory. The *implementation* describes the hardware underlying the algorithm.

Following Marr himself, theories of grammar are often described as computational-level theories. Let's take syntax as an example. A computational-level syntactic theory is a description of the inputs and outputs—word strings and hierarchical structures—and the logic of the (reversible) relation between them. Psycholinguistic models are often regarded as algorithms: they specify the representations and operations responsible for accomplishing the translation between word strings and hierarchical structures. At the level of implementation, we would need a description of how these algorithms are instantiated using neural mechanisms.

It is certainly useful to think about language and other cognitive systems at multiple levels of abstraction. In fact, we agree with Marr that it is actually impossible to fully understand a cognitive system without doing so. However, we have some concerns about how Marr's levels are used in discussions of the language system.

First, it is misguided to justify a division of labor between linguists, psycholinguists, and neuroscientists by appealing to the three levels. Marr argues that no single level of description can be well understood without reference to the others. This point on its own suggests that isolating the different perspectives on the language system will not be helpful. To take an obvious example, a complete map of a human neural network and a description of all its activities would be useless without accompanying descriptions of the algorithms being implemented and the function of the system as a whole. On the other side, linguists should not appeal to Marr as a refuge from explaining the mentalistic commitments of their theories. So, for example, many syntacticians build theories in which sentence structures are assumed to be constructed via bottom-to-top derivations and generally right-to-left movement or copying operations.

There is very little discussion of the mentalistic commitments of these claims. The question cannot be avoided by appeal to Marr's computational level, for two reasons. First, a complete understanding of the system requires linking together descriptions at different levels of abstraction. There is no reason that theoretical linguists should not contribute to this effort by at least developing linking-ready theories. Second, as we discuss below, most grammatical theories already go beyond what could be considered a computational level description and bleed into the algorithmic level. (For more discussion of derivations in linguistic theory and sentence processing, see Phillips & Lewis, 2013.)

A related point that Marr emphasizes repeatedly is that any given phenomenon may be best explained at only one of the levels. If research at each level is conducted in isolation, then researchers at one level may end up hammering away fruitlessly at a problem that would be more easily solved at another level. Such cases have already been encountered in research on the language system. For example, various generalizations about syntactic constraints have been argued to be explainable if we move from the standard time-independent level of abstraction to a slightly lower level of abstraction that specifies the temporal order of structure building operations (e.g., Chesi, 2007; Kempson, Meyer-Viol, & Gabbay, 2001; Phillips, 2003; Shan & Barker, 2006). Others have argued that some grammatical generalizations turn out to be epiphenomena once we take account of independent constraints on memory and cognitive control that are typically the purview of theories of language processing (Anderson, 2004; Hofmeister & Sag, 2010; Kluender & Kutas, 1993; Steiner, 2009), though such arguments are not uncontroversial (Phillips, 2013b; Sprouse, Wagers, & Phillips, 2012).

A different kind of problem we observe in discussions of Marr's levels is the emphasis on "three". Although the division into three distinct levels is convenient, it is really a simplification of a continuum of abstraction. At the "bottom" extreme, you have a description of a physical implementation of that cannot be made more concrete. At the "top" extreme, you have a computational description that cannot be made more abstract. In the middle, there is a wide range of possible descriptions at different levels of abstraction. In some cases, the simplification of this continuum into three discrete levels may be misleading. For example, despite assumptions to the contrary, most grammatical theories do not operate at the highest possible level of abstraction. They specify more than the inputs and outputs of the system, instead providing detailed descriptions of the internal components of the grammar. Much of what syntacticians spend their time studying is the many structures that are not allowed by the system. Their descriptions of the constraints on possible structures cannot be considered highly abstract characterizations of *what* the system does—these explanations clearly concern the *how*. If a grammatical theory claims that specific parts of human grammar are subject to cross-linguistic variation, or need to be learned by children, then this already is a more detailed description of the system which goes beyond the specification of inputs and outputs as required by a computational-level theory.

Pietroski and colleagues (2009) provide a useful example from semantics. The meaning of the quantifier 'most' is something like "more than half". So the meaning of (24) is something like "more than half of the dots are blue."

(24) Most of the dots are blue.

In a traditional semantic theory, the input would be some representation of the sentence that includes lexical and syntactic information, and the output would be the set of all situations that would make the sentence true (assuming a truth-based semantic system). In practice, most semanticists go further: they write down a mathematical representation of the set of situations satisfying the truth conditions. In the case of (24), there are multiple ways to do that, including those in (25)-(27) (stated pseudo-mathematically for convenience). All of these methods would yield the same result (i.e., the same set of situations where the sentence is true), and many semanticists refuse to commit to the mental implications of these logical formulas. In fact, they ought to be taken seriously as hypotheses about mental algorithms. Using one formula over another is a claim about *how* the output is computed, not just *what* the output is.[8]

(25) (number of blue dots) > (number of dots)/2

(26) (number of blue dots) > (number of dots) – (number of blue dots)

(27) (number of blue dots) > (number of non-blue dots)

Similarly, psycholinguistic models do not operate at a well-defined level of analysis corresponding to Marr's algorithm level, instead operating at varying degrees of abstraction. To take just one example, a psycholinguistic model of pronoun resolution might simply specify that upon encountering a pronoun an antecedent is retrieved from memory. Or it might offer a more detailed account of that retrieval process, couched in terms of a specific model of memory encoding and access. Or it might go a step further and describe how specific components of the retrieval process engage different cortical areas at specific times, on a scale of tens to hundreds of milliseconds. All of these can be useful ways of characterizing pronoun resolution, and all are algorithmic in some fashion, but they involve different degrees of abstraction.

Properties of the system at each level of description constrain the properties of other levels. The algorithm for implementing a given computation is shaped by both its purpose and its implementation. It is therefore entirely appropriate that psycholinguistic models are influenced by more abstract grammatical theory. However, the influence should flow in the other direction as well. To the extent that grammatical theories incorporate algorithms as well as computational-level theories, they must be sensitive to the constraints of their implementation. We have argued elsewhere that it is an empirical question how *implementation independent* or *dependent* any abstract property of the language system is, and that this

[8] Based on evidence from verification tasks, the right answer seems to be (26) (Lidz, Pietroski, Halberda, & Hunter, 2011).

impacts whether the abstract description should be understood as a matter of convenience or as having independent neurocognitive status. (Phillips & Lewis, 2013).

6 Conclusion

We are very encouraged to encounter renewed interest in bridging the divide between grammatical theories and psycholinguistic theories. The divide appears to have grown up in large part based on practical differences (laboratory methods vs. simpler data gathering), which in turn led to separate focus on online vs. offline data. We think that successfully bridging the two fields requires more than simply using psycholinguistic notions to explain away traditional grammatical generalizations, or using online evidence to answer old questions about competing grammatical theories. Instead, closing the gap between grammatical theories and language processing models requires answers to two classes of questions that are rarely investigated. First, it is important to identify the relation between the cognitive systems that the two fields are studying—are they separate cognitive systems (a “two-system architecture”) or are they simply different descriptions of the same cognitive system (a “one-system architecture”). In our assessment, the existing empirical evidence favors the second of these positions. Second, for both of these architectures we have laid out a series of theoretical and empirical challenges that should be addressed in order to properly understand the relation between theories of fast and slow phenomena in language.

7 Acknowledgments

Preparation of this paper was supported in part by NSF #BCS-0848554 to CP, by NSF IGERT #DGE-0801465 to the University of Maryland, and by a University of Maryland Flagship Fellowship to SL. For useful discussion we are indebted to Brian Dillon, Janet Fodor, Norbert Hornstein, Dave Kush, Bradley Larson, Jeff Lidz, and Amy Weinberg.

8 References

Aho, A. V., Lam, M. S., Sethi, R., & Ullman, J. D. (2006). *Compilers: principles, techniques, and tools (2nd edition)*. Boston: Addison-Wesley.

Anderson, C. (2004). *The structure and real-time comprehension of quantifier scope ambiguity*. PhD dissertation, Northwestern University.

Aoshima, S., Phillips, C., & Weinberg, A. (2004). Processing filler-gap dependencies in a head-final language. *Journal of Memory and Language, 51*, 23-54.

Aoshima, S., Yoshida, M., & Phillips, C. (2009). Incremental processing of coreference and binding in Japanese. *Syntax, 12*, 93-134.

Badecker, W., & Straub, K. (2002). The processing role of structural constraints on the interpretation of pronouns and anaphora. *Journal of Experimental Psychology: Learning, Memory, and Cognition , 28*, 748-769.

Bader, M., Meng, M., & Bayer, J. (2000). Case and reanalysis. *Journal of Psycholinguistic Research, 29*, 37-52.

Berwick, R. C., Friederici, A. D., Chomsky, N., & Bolhuis, J. J. (2013). Evolution, brain, and the nature of language. *Trends in Cognitive Sciences, 17*, 89-98.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 279-362). New York: Wiley.

Bever, T. G. & McElree, B. (1988). Empty categories access their antecedents during comprehension. *Linguistic Inquiry, 19*, 35-43.

Bickerton, D. (2003). Symbol and structure: A comprehensive framework for language evolution. In M. H. Christiansen & S. Kirby (eds.), *Language Evolution* (pp. 77-93). Oxford: Oxford University Press.

Bock, J. K. & Cutting, J. C. (1992). Regulating mental energy: performance units in language production. *Journal of Memory and Language, 31*, 99-127.

Bock, K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology , 23*, 45-93.

Bott, L. & Noveck, I. A. (2004). Some utterances are underinformative: the onset and time course of scalar inferences. *Journal of Memory and Language, 51*, 437-457.

Bourdages, J. S. (1992). Parsing complex NPs in French. In H. Goodluck, & M. S. Rochemont (Eds.), *Island Constraints: Theory, Acquisition and Processing* (pp. 61-87). Dordrecht: Kluwer Academic.

Breheny, R., Katsos, N., & Williams, J. (2006). Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences. *Cognition, 100*, 434-463.

Bresnan, J. (1978). A realistic transformational grammar. In G. Miller, M. Halle, & J. Bresnan (eds.), *Linguistic theory and psychological reality* (pp. 282-390). Cambridge, MA: MIT Press.

Büring, D. (2005). *Binding theory*. Cambridge, UK: Cambridge University Press.

Chesi, C. (2007). Five reasons for building phrase structures top-down from left to right. *Nanzan Linguistics, Special Issue 3*, 71-105.

Chomsky, N. (1973). Conditions on transformations. In S. Anderson & P. Kiparsky (eds.), *A festschrift for Morris Halle* (pp. 232-286). New York: Holt, Rinehart, & Winston.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.

Chow, W.Y., Lewis, S., & Phillips, C. (in prep). Immediate structural constraints on pronoun resolution.

Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology , 42*, 368-407.

Clackson, K., Felser, C., & Clahsen, H. (2011). Children's processing of reflexives and pronouns in English: evidence from eye-movements during listening. *Journal of Memory and Language, 65*, 128-144.

Clifton, C. Jr. & Frazier, L. (1989). Comprehending sentences with long-distance dependencies. In M. K. Tanenhaus & G. N. Carlson (eds.), *Linguistic structure in language processing* (pp. 273-317). Dordrecht: Kluwer.

Clifton, C. Jr., Frazier, L., & Deevy, P. (1999). Feature manipulation in sentence comprehension. *Rivista di Linguistica, 11*, 11-39.

Clifton, C. Jr., Kennison, S., & Albrecht, J. (1997). Reading the words *her*, *him*, and *his*: implications for parsing principles based on frequency and on structure. *Journal of Memory and Language, 36*, 276-292.

Condry, K. & Spelke, E. (2008). The development of language and abstract concepts: the case of natural number. *Journal of Experimental Psychology: General, 137*, 22-38.

Cowart, W. & Cairns, H. S. (1987). Evidence for an anaphoric mechanism within syntactic processing: Some reference relations defy semantic and pragmatic constraints. *Memory and Cognition, 15*, 318-331.

Crain, S. & Fodor, J. D. (1987). Sentence matching and overgeneration. *Cognition, 26*, 123-169.

Culicover, P. & Jackendoff, R. (2005). *Simpler syntax*. Oxford: Oxford University Press.

de Villiers, J. G., & de Villiers, P. A. (2009). Complements enable representation of the contents of false beliefs: the evolution of a theory of theory of mind. In S. Foster-Cohen (Ed.), *Language Acquisition.* New York: Palgrave Macmillan.

de Villiers, J. G., & Pyers, J. E. (2002). Complements to cognition: a longitudinal study of the relationship between complex syntax and false-belief understanding. *Cognitive Development , 17*, 1037-1060.

de Villiers, J. G. (2007). The interface of language and theory of mind. *Lingua , 117*, 1858-1878.

Dillon, B. W. (2011). *Structured access in sentence comprehension*. PhD dissertation, University of Maryland.

Dillon, B., Mishler, A., Sloggett, S., & Phillips, C. (2013). Contrasting interference profiles for agreement and anaphora: experimental and modeling evidence. *Journal of Memory and Language, 69*, 85-103.

Drenhaus, H., Saddy, D., & Frisch, S. (2005). Processing negative polarity items: when negation comes through the back door. In S. Kepser & M. Reis (eds.), *Linguistic evidence: empirical, theoretical, and computational perspectives* (pp. 145-165). Berlin: de Gruyter.

Edelman, S. & Christiansen, M. (2003). How seriously should we take Minimalist syntax? *Trends in Cognitive Sciences, 7*, 60-61.

Ferreira, F. (2005). Psycholinguistics, formal grammars, and cognitive science. *The Linguistic Review, 22,* 365-380.

Ferreira, F., Ferraro, V., & Bailey, K. G. D. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science, 11,* 11-15.

Ferreira, F. & Patson, N. (2007). The ‘good enough’ approach to language comprehension. *Language and Linguistics Compass, 1,* 71-83.

Fodor, J. A., Bever, T. G., & Garrett, M. F. (1974). *The psychology of language.* New York: McGraw-Hill.

Francis, W. N. (1986). Proximity concord in English. *Journal of English Linguistics, 19*, 309-317.

Franck, J., Vigliocco, G., & Nicol, J. (2002). Attraction in sentence production: the role of syntactic structure. *Language and Cognitive Processes, 17*, 371-404.

Frank, S. L., Bod, R., & Christiansen, M. H. (2012). How hierarchical is language use? *Proceedings of the Royal Society B, 279*, 4522-4531.

Frazier, L. (1985). Syntactic complexity. In D. Dowty, L. Karttunen, & A. Zwicky (eds.), *Natural language processing: psychological, computational, and theoretical perspectives* (pp. 129-189). Cambridge, UK: Cambridge University Press.

Freedman, S. E. & Forster, K. I. (1985). The psychological status of overgenerated sentences. *Cognition, 19*, 101-131.

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: effects of semantic, morphological, and syntactic violations. *Cognitive Brain Research, 1*, 183-192.

Giannakidou A. (2011). Negative and positive polarity items. In K. von Heusinger, C. Maienborn, & P. Portner (eds.), *Semantics: an international handbook of natural language meaning* (pp. 1660-1712). Berlin: de Gruyter.

Gibson, E. & Fedorenko, E. (2013). The need for quantitative methods in syntax and semantics research. *Language and Cognitive Processes, 28*, 88-124.

Gibson, E. & Hickok, G. (1993). Sentence processing with empty categories. *Language and Cognitive Processes, 8*, 147-161.

Gibson, E. & Thomas, J. (1999). Memory limitations and structural forgetting: the perception of complex ungrammatical sentences as grammatical. *Language and Cognitive Processes, 14*, 225-248.

Gimenes, M., Rigalleau, F., & Gaonach, D. (2009). When a missing verb makes a French sentence more acceptable. *Language and Cognitive Processes, 24*, 440-449.

Grodner, D. J., Klein, N. M., Carbary, K. M., & Tanenhaus, M. K. (2010). "Some", and possibly all, scalar inferences are not delayed: evidence for immediate pragmatic enrichment. *Cognition, 116*, 42-55.

Grune, D, & Jacobs, C. J. H. (2008). *Parsing techniques: a practical guide*. New York: Springer.

Hackl, M., Koster-Hale, J., & Varvoutis, J. (2012). Quantification and ACD: evidence from real-time sentence processing. *Journal of Semantics, 29*, 145-207.

Hermer, L. & Spelke, E. (1996). Modularity and development: the case of spatial reorientation. *Cognition, 61,* 195-232.

Hermer-Vazquez, L., Spelke, E., & Katsnelson, A. S. (1999). Sources of flexibility in human cognition: dual-task studies of space and language. *Cognitive Psychology, 39*, 3-36.

Hermer-Vazquez, L., Moffet, A., & Munkholm, P. (2001). Language, space, and the development of cognitive flexibility in humans: the case of two spatial memory tasks. *Cognition, 79*, 263-299.

Hofmeister, P. & Sag, I. A. (2010). Cognitive constraints and island effects. *Language, 86*, 366-415.

Huang, Y. & Snedeker, J. (2009). Online interpretation of scalar quantifiers: insight into the semantics-pragmatics interface. *Cognitive Psychology, 58*, 376-415.

Hussey, E. K. & Novick, J. (2012). The benefits of executive control training and the implications for language processing. *Frontiers in Psychology, 3,* 1-14.

Hyde, D. C., Winkler-Rhoades, N., Lee, S.-A., Izard, V., Shapiro, K. A., & Spelke, E. S. (2011). Spatial and numerical abilities without a complete natural language. *Neuropsychologia, 49,* 924-936.

Jakob, F. (1977). Evolution and tinkering. *Science, 196*, 161-171.

Just, M. A. & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99,* 122-149.

Kaan, E. (2007). Event-related potentials and language processing: a brief overview. *Language and Linguistics Compass, 1*, 571-591.

Kazanina, N., Lau, E. F., Lieberman, M., Yoshida, M., & Phillips, C. (2007). The effect of syntactic constraints on the processing of backward anaphora. *Journal of Memory and Language, 56*, 384-409.

Kazanina, N. & Phillips, C. (2010). Differential effects of constraints in the processing of Russian cataphora. *Quarterly Journal of Experimental Psychology, 63*, 371-400.

Kempen, G. (in press). Prolegomena to a neurocomputational architecture for human grammatical encoding and decoding. *Neuroinformatics.*

Kempen, G., Olsthoorn, N., & Sprenger, S. (2012). Grammatical workspace sharing during language production and language comprehension: evidence from grammatical multitasking. *Language and Cognitive Processes, 27*, 345-380.

Kempson, R., Meyer-Viol, W., & Gabbay, D. (2001). *Dynamic syntax: the flow of language understanding.* Malden, MA: Blackwell.

Kennison, S. (2003). Comprehending the pronouns *her*, *him*, and *his*: implications for theories of referential processing. *Journal of Memory and Language, 49*, 335-352.

King, J., Andrews, C., & Wagers, M. (2012). Do reflexives always find a good antecedent for themselves? Poster at the 25th annual CUNY Conference on Human Sentence Processing. New York, NY.

Kluender, R. & Kutas, M. (1993). Subjacency as a processing phenomenon. *Language and Cognitive Processes, 8*, 373-433.

Kush, D. (2013). *Respecting relations: memory access and antecedent retrieval in incremental sentence processing.* PhD dissertation, University of Maryland.

Ladusaw, W. (1996). Negation and polarity items. In S. Lappin (ed.), *The handbook of contemporary syntactic theory* (pp. 321-341). Oxford: Blackwell.

Lago, S. & Phillips, C. (in prep.). Agreement attraction effects in Spanish comprehension.

Ledoux, K., Gordon, P. C., Camblin, C. C., & Swaab, T. Y. (2007). Coreference and lexical repetition: mechanisms of discourse integration. *Memory and Cognition, 35*, 801-815.

Lee, M.-W. (2004). Another look at the role of empty categories in sentence processing (and grammar). *Journal of Psycholinguistic Research, 33*, 51-73.

Levelt, W. J. M. (1974). *Formal grammars in linguistics and psycholinguistics* (3 volumes). The Hague: Mouton.

Lewis, S. (2013). *Pragmatic enrichment in language processing and development*. PhD dissertation, University of Maryland.

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science , 29*, 1-45.

Lidz, J., Pietroski, P., Halberda, J., & Hunter, T. (2011). Interface transparency and the psychosemantics of *most*. *Natural Language Semantics*, *19*, 227-256.

Linebarger, M. (1987). Negative polarity and grammatical representation. *Linguistics and Philosophy, 10*, 325-387.

MacDonald, M. C. (1989). Priming effects from gaps to antecedents. *Language and Cognitive Processes, 4*, 1-72.

MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology, 24*, 56-98.

Marr, D. (1982). *Vision*. San Francisco: W. H. Freeman.

McElree, B., Foraker, S., & Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language* , 67-91.

McElree, B. & Griffith, T. (1998). Structural and lexical constraints on filling gaps during sentence comprehension: a time-course analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 432-460.

Merchant, J. (2001). *The syntax of silence: sluicing, islands, and identifying in ellipsis*. Oxford: Oxford University Press.

Miller, G. A. & Chomsky, N. (1963). Finitary models of language users. In R. D. Luce, R. R. Bush, & E. Galanter (eds.), *Handbook of mathematical psychology, volume 2* (pp. 419-492). New York: Wiley.

Miller, G. A. & McKean, K. (1964). A chronometric study of some relations between sentences. *Quarterly Journal of Experimental Psychology, 16*, 297-308.

Neville, H., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically-based sentence processing classes: evidence from event-related brain potentials. *Journal of Cognitive Neuroscience, 3*, 151-165.

Nakano, Y., Felser, C., & Clahsen, H. (2002). Antecedent priming at trace positions in Japanese long-distance scrambling. *Journal of Psycholinguistic Research, 31*, 531-571.

Nicol, J. L., Fodor, J. D., & Swinney, D. (1994). Using cross-modal lexical decision tasks to investigate sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1-10.

Nicol, J. & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research, 18*, 5-19.

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.

Novick, J., Hussey, E. K., Teubner-Rhodes, S., Harbison, J. I., & Bunting, M. (2013). Clearing the garden path: Improving sentence processing through executive control training. *Language and Cognitive Processes*.

Omaki, A. & Schulz, B. (2011). Filler-gap dependencies and island constraints in second language sentence processing. *Studies in Second Language Acquisition, 33*, 563-588.

Osterhout, L. & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31*, 785-806.

Pablos, L., Ruijgrok, B., Doetjes, J., & Cheng, L. (2012). Processing cataphoric pronouns in Dutch: an ERP study. Talk at the GLOW workshop on Timing in Grammar. Potsdam, Germany.

Patil, U., Vasishth, S., & Lewis, R. (2011). Early retrieval interference in syntax-guided antecedent search. Talk at the 24th annual CUNY Conference on Human Sentence Processing. Stanford, CA.

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

Phillips, C. (1996). *Order and structure*. PhD dissertation, MIT.

Phillips, C. (2003). Linear order and constituency. *Linguistic Inquiry, 34*, 37-90.

Phillips, C. (2006). The real-time status of island phenomena. *Language , 82*, 195-823.

Phillips, C. (2010). Should we impeach armchair linguists? In S. Iwasaki, H. Hoji, P. Clancy, & S.-O. Sohn (eds.), *Japanese-Korean Linguistics, 17* (pp. 49-64). Stanford, CA: CSLI Publications.

Phillips, C. (2013a). Parser-grammar relations: we don't understand everything twice. In M. Sanz, I. Laka, & M. Tanenhaus (eds.), *Language down the garden path: the cognitive and biological basis for linguistic structure* (pp. 294-315). Oxford University Press.

Phillips, C. (2013b). Some arguments and non-arguments for reductionist accounts of syntactic phenomena. *Language and Cognitive Processes, 28*, 156-187.

Phillips, C. & Lewis, S. (2013). Derivational order in syntax: evidence and architectural consequences. *Studies in Linguistics, 6*, 11-47.

Phillips, C. & Parker, D. (in press). The psycholinguistics of ellipsis. *Lingua.*

Phillips, C. & Wagers, M. W. (2007). Relating structure and time in linguistics and psycholinguistics. In G. Gaskell (ed.), *The Oxford Handbook of Psycholinguistics* (pp. 739-756). Oxford University Press.

Phillips, C., Wagers, M. W., & Lau, E. F. (2011). Grammatical illusions and selective fallibility in real-time language comprehension. In J. Runner (ed.), *Experiments at the interfaces (Syntax and Semantics, vol. 37)* (pp. 153-186). Bingley, UK: Emerald.

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian Indigene group. *Science, 306*, 499-503.

Pickering, M. J. & Barry, G. D. (1991). Sentence processing without empty categories. *Language and Cognitive Processes, 6*, 229-259.

Pickering, M., Barton, J. S., & Shillcock, R. (1994). Unbounded dependencies, island constraints, and processing complexity. In C. Clifton, L. Frazier, & K. Rayner (eds.), *Perspectives on sentence processing* (pp. 199-224). London: Lawrence Erlbaum.

Pietroski, P., Lidz, J., Hunter, T., & Halberda, J. (2009). The meaning of 'most': Semantics, numerosity and psychology. *Mind & Language, 24*, 554-585.

Pinker, S. & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences, 13*, 707-727.

Prince, A. & Smolensky, P. (2004). *Optimality theory: constraint interaction in generative grammar*. Oxford: Blackwell.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. New York, NY: Longman.

Runner, J. T., Sussman, R., S., & Tanenhaus, M. K. (2006). Processing reflexives and pronouns in picture noun phrases. *Cognitive Science, 30*, 193-241.

Sag, I. A. & Fodor, J. D. (1994). Extraction without traces. In R. Aronovich, W. Byrne, S. Preuss, & M. Senturia (eds.), *Proceedings of the 13th annual meeting of the West Coast Conference on Formal Linguistics* (pp. 365-384). Stanford: CSLI Publications.

Sag, I. A. & Wasow, T. (2011). Performance-compatible competence grammar. In R. Borsley & K. Borjars (eds.), *Non-transformational syntax: formal and explicit models of grammar.* Oxford: Wiley-Blackwell.

Shan, C.-C. & Barker, C. (2006). Explaining crossover and superiority as left-to-right evaluation. *Linguistics and Philosophy, 29*, 91-134.

Sloggett, S. (2013). Case licensing in processing: evidence from German. Poster at the 26th annual CUNY Conference on Human Sentence Processing. Columbia, SC.

Spelke, E. S. (2003). What makes us smart? Core knowledge and natural language. In D. Gentner & S. Goldin-Meadow (eds.), *Language in Mind* (pp. 277-311). Cambridge, MA: MIT Press.

Sprouse, J. & Almeida, D. (2012). Assessing the reliability of textbook data in syntax: Adger's *Core Syntax*. *Journal of Linguistics, 48*, 609-652.

Sprouse, J. & Lau, E. F. (2013). Syntax and the brain. In M. den Dikken (ed.), *The Cambridge Handbook of Generative Syntax*. Cambridge University Press.

Sprouse, J., Schütze, C. T., & Almeida, D. (in press). A comparison of informal and formal acceptability judgments using a random sample from Linguistic Inquiry 2001-2010. *Lingua.*

Sprouse, J., Wagers, M. W., & Phillips, C. (2012). A test of the relation between working memory capacity and syntactic island effects. *Language, 88,* 82-123.

Stabler, E. P. (2013). Two models of minimalist, incremental syntactic analysis. *Topics in Cognitive Science*. DOI: 10.1111/tops.12031

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). Reponse time distributional evidence for distinct varieties of number attraction. *Cognition, 114*, 447-454.

Steiner, I. (2009). Partial agreement in German: a processing issue? In S. Featherston & S. Winkler (eds.), *The fruits of empirical linguistics*, vol. II (pp. 239-260). Berlin: de Gruyter.

Stowe, L. A. (1986). Parsing WH-constructions: evidence for on-line gap location. *Language and Cognitive Processes, 3*, 227-245.

Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language, 48,* 542-562.

Sturt, P. (2007). Semantic re-interpretation and garden path recovery. *Cognition , 105*, 477-488.

Swaab, T. Y., Camblin, C. C., & Gordon, P. C. (2004). Reversed lexical repetition effects in language processing. *Journal of Cognitive Neuroscience, 16*, 715-726.

Szabolcsi, A. (2013). Quantification and ACD: what is the evidence from real-time processing evidence for? A response to Hackl et al. (2012). *Journal of Semantics*. Published online 1/28/13, DOI: 10.1093/jos/ffs025

Townsend, D. J. & Bever, T. G. (2001). *Sentence comprehension: the integration of habits and rules*. Cambridge, MA: MIT Press.

Traxler, M. J. & Pickering, M. J. (1996). Plausibility and the processing of unbounded dependencies. *Journal of Memory and Language, 35*, 454-475.

Trotzke, A., Bader, M., & Frazier, L. (2013). Third factors and the performance interface in language design. *Biolinguistics, 7*, 1-34.

Vasishth, S., Brüssow, S., Lewis, R., & Drenhaus, H. (2008). Processing polarity: how the ungrammatical intrudes on the grammatical. *Cognitive Science, 32*, 685-712.

Wagers, M. W. (2008). *The structure of memory meets memory for structure in linguistic cognition*. PhD dissertation, University of Maryland.

Wagers, M., Lau, E., & Phillips, C. (2009). Agreement attraction in comprehension: representations and processes. *Journal of Memory and Language , 61*, 206-237.

Wagers, M. W. & Phillips, C. (in press). Going the distance: memory and decision making in active dependency construction. *Quarterly Journal of Experimental Psychology.*

Wang, L., Bastiaansen, M., Yang, Y., & Hagoort, P. (2012). Information structure influences depth of syntactic processing: event-related potential evidence for the Chomsky Illusion. *PLOSone, 7*, 1-9.

Wasow, T. & Arnold, J. (2005). Intuitions in linguistic argumentation. *Lingua, 115*, 1481-1496.

Wellwood, A., Pancheva, R., Hacquard, V., & Phillips, C. (submitted). Deconstructing a comparative illusion. Ms. University of Maryland and University of Southern California.

Whitney, C. S. (2004). *Investigations into the neural basis of structured representations*. PhD dissertation, University of Maryland.

Xiang, M., Dillon, B., & Phillips, C. (2009). Illusory licensing effects across dependency types: ERP evidence. *Brain and Language, 108*, 40-55.

Yoshida, M., Aoshima, S., & Phillips, C. (2004). Relative clause prediction in Japanese. Talk at the 17th Annual CUNY Conference on Human Sentence Processing. College Park, MD.

Zukowski, A. & Larsen, J. (2004). The production of sentences that we fill their gaps. Poster at the 17th Annual CUNY Conference on Human Sentence Processing. College Park, MD.