CHAPTER 45

Relating structure and time in linguistics and psycholinguistics

Colin Phillips and Matthew Wagers

45.1 Linguistics and psycholinguistics

The field of psycholinguistics advertises its mentalistic commitments in its name. The field of linguistics does not. Psycholinguistic research frequently involves ingenious experimental designs, fancy lab equipment such as eye-trackers or electroencephalograms (EEGs), large groups of experimental subjects, and detailed statistical analyses. Linguistic research typically requires no specialized equipment, no statistical analyses, and somewhere between zero and a handful of cooperative informants. Psycholinguistic research is most commonly conducted in a Department of Psychology. Linguistic research is not. Some of these differences may contribute to the widespread perception, well-represented among linguists and psychologists alike, that the concerns of psycholinguistics are somehow more psychological than those of linguistics, and that psycholinguistics can be looked to for psychological validation of the constructs proposed by linguists. Although this view of the relation between the two fields gives the impression of a neat division of labor, we find it misleading, and suspect that it may have led to unrealistic expectations, and consequently to disappointments and mutual frustration.

In this commentary we focus on issues in the representation of unbounded syntactic dependencies, as a case study of what psycholinguistic methods can and cannot tell us about linguistic questions, and vice versa. This is an area where a rich linguistic literature and a sizeable body of psycholinguistic research address closely related phenomena. The most widely discussed form of unbounded dependency occurs when a noun phrase (NP) such as *which voters* in (1) appears in a position that is structurally distant from the verb that it is an argument of (e.g. *bribe* in (1)). Following standard practice, we mark the canonical position of the direct object of *bribe* with an underline or *gap* in (1), but it is a matter of great controversy whether such gap positions are a part of the mental representation of sentences like (1). After reviewing some of the competing linguistic analyses of unbounded dependency constructions in section 45.2, we discuss in section 45.3 the contributions of psycholinguistics to the question of how these dependencies are represented.

(1) Which voters did the prosecutor suspect that the candidate wanted his operatives to bribe ___ before the election?

The status of constraints on long-distance dependencies, such as the ban on dependencies that span relative clause boundaries, as illustrated in (2), is a major topic of linguistic research. In section 45.4 we discuss the effects of such constraints on language processing and their implications for the relation between linguistic and psycholinguistic models.

(2) *Which voters did the prosecutor ask [$_{NP}$ the operative [$_{RC}$ that bribed ___]] to testify against his boss?

As far as we can tell, there is no principled difference between the psychological relevance of psycholinguistic and linguistic research. Most modern linguists have serious mentalistic commitments (and those that do not are of little concern to us here).[1] The data that linguists and psycholinguists collect and the theories that they develop based on those data are all "psychological," in the sense that they aim to explain some aspect of human cognitive abilities. There are certainly differences in the issues and methods that the two fields tend to pay closest attention to; but we are unaware of reasons to think that either discipline has more direct access to psychological evidence, and we would include in this those strains of psycholinguistics that draw on cognitive neuroscience methods, as we do in our own work. Since this is not a standard position, we will briefly attempt to substantiate this claim.

There is broad agreement that mastery of a language involves at least an (unconscious) understanding of the range of possible expressions that can be represented in the language, and also an ability to identify those expressions that cannot be represented in the language. It is also agreed that since the expressive power of human languages is too large for any individual to simply memorize the possible representations of his language, a speaker must have the ability to generate, recognize, and interpret novel expressions of the language relatively quickly in order to speak and understand. Since humans are able to learn any language for which they receive adequate exposure from an early age, it is also agreed that the ability to learn language is a key component of the human capacity for language. Understanding each of these abilities is an important part of the task of explaining how human language works, and it is perhaps an accident of history or methodology that different sub-fields have emerged that focus on each of these problems.

Linguistic theory typically focuses on characterizing the representations that a speaker of a language can entertain, often allied with the question of what is a "possible human language." Theoretical linguists have generally paid less attention to questions about how these representations might be retrieved or constructed in realtime

tasks such as speaking and understanding. In contrast, questions about realtime processes have been central in adult psycholinguistics, whereas less attention has been given to the question of why certain expressions are possible and others are impossible. This focus of adult psycholinguistics upon mechanisms that are closely tied to speaking and understanding is sometimes justified by the notion that these are more directly related to common behaviors, or by the assumption that the goal of psycholinguistics is to provide "processing models." However, a look at psycholinguistic work with children casts doubt upon this rationale. Developmental psycholinguistics has devoted much attention to the question of what children can and cannot represent at different ages. Studies of childrens realtime processing and detailed studies of learning have become more prominent only recently. Much work in developmental psycholinguistics asks the same questions about children that linguists ask about adults. Therefore, the question of what a speaker can and cannot represent, and the question of how the representations are constructed in time, are presumably both psychologically respectable concerns. We suspect that the disciplinary divisions have more to do with the methodological biases of the respective fields.

There are obvious differences between the data collection methods most commonly used in linguistics and psycholinguistics. The primary data of theoretical linguistics comes from native speakers' intuitive judgements about the acceptability of sentences or the availability of specific interpretations. Such data are relatively easy to come by, making it possible to establish a large number of facts about many different languages in relatively little time. Psycholinguistic data, on the other hand, typically require a good deal more effort. In order to establish reliable generalizations about reaction times, focal brain activity, or any of a number of other common dependent measures, one needs to use specialized equipment, test large numbers of experimental items on large numbers of participants, devise ingenious ways to hide one's goals from the participants, and use complex statistical analyses to interpret the results. It can take a lot of work to establish just one fact, and it can be difficult to conduct experiments on a number of different languages.

The different data collection practices of linguistics and psycholinguistics have a clear impact upon the fields. First, they affect the empirical scope of the fields. Thanks to its low-tech methods, linguistics has amassed a large body of findings from a very diverse set of languages, including languages for which only a small number of

---

[1] We should emphasize that we are concerned in this commentary with what we take to be the "best practices" in either field. It is not difficult to find instances of careless misrepresentation of linguistic data, uninterpretable experimental designs, unwarranted inferences from brain activation patterns, etc., of which neither field would be proud. Our interest here is more in what can be learned from carefully conducted work in either field.

speakers remain. In contrast, most psycholinguistic research has been confined to a handful of closely related western European languages. Second, differences in data collection methods affect what linguists and psycholinguists spend their time on, shape what is valued in the two fields, and also affect the safeguards that the two fields place on data reliability. In psycholinguistics data collection is sufficiently difficult that great value is placed on elegant experimental designs that make it possible to establish a single fact, and many procedural safeguards are put in place in order to ensure that results are reliable. In linguistics, on the other hand, data collection methods are relatively trivial, and receive correspondingly little attention (although there is obvious value in the use of carefully controlled test sentences). Except when dealing with speakers of scarce languages, replication and verification of the empirical facts is straightforward, and hence fewer safeguards are needed to avoid the damaging effects of bogus findings. It is therefore understandable that in linguistics little value is placed on establishing individual facts. Greater value is placed on weaving together large bodies of facts into interesting general theories. The term "theoretical" in "theoretical linguistics" is all too often taken to imply that the field is somehow less concerned with empirical facts. This is unwarranted. The term merely reflects the fact that the empirical side of the field is sufficiently easy that most time is spent worrying about what the facts all mean.[2] Similarly, psycholinguists take questions of theory seriously, although such questions take up less time on a day-to-day basis.

It is sometimes objected that the results of linguistics are less reliable or objective than those of psycholinguistics (e.g. Ferreira 2005), or that they provide less direct access to the workings of the mind or brain. Linguists are often criticized—and frequently criticize themselves—because they "do not run experiments." Aside from the mundane concern that all findings should be reported carefully and honestly, we do not see the force of this objection. Most of the linguistic literature is built upon robust acceptability judgements, and robust judgements become statistically reliable with rather small samples—unbiased agreement

among half a dozen friends will do (e.g. two-tailed sign test). One can often do without the half dozen friends, too, if one is sufficiently confident about the judgement. There are, of course, examples of errors and disagreements, and notorious cases of judgements that are subtle at best, but these are the exception rather than the rule. In our own work we often run large acceptability rating studies as controls for our on-line studies. These studies cost little effort, since the materials are independently needed for the on-line studies, but the results are almost never surprising, and are generally so robust statistically as to indicate that we tested more subjects than needed. As far as we can tell, if linguists were to replace their standard informal experiments with larger-scale acceptability judgement studies, the main consequence would be to slow the discovery of new facts.[3] Furthermore, linguistic methods are typically used to support rather direct inferences from the observed data. If a sentence is judged to be unacceptable (and various obvious controls for plausibility, memory, etc. are satisfied), then it is inferred that the sentence is not a well-formed product of the speaker"s language system. In psycholinguistics, on the other hand, we typically draw rather more indirect inferences. A 30-ms slowdown in the time that it takes to press a button may be used to infer the presence of a structural ambiguity or the need for parsing revision. A 2-microvolt positive deflection in an averaged scalp voltage may be used to infer that selective disruption is occurring in syntactic processing. These experimental methods are more appropriate than acceptability judgements for selectively targeting unconscious processes, and they certainly have more fine-grained temporal resolution, but they are not more direct windows into the mind.

Overall, we see little reason to view the concerns or the methods of either linguistics or psycholinguistics as more or less psychological in nature. This may seem obvious to some, and bizarre to others. However, we suspect that some of the misunderstanding and mutual frustration

[2] See Miller (1990) for an interesting related commentary. Miller argues that linguists and psychologists tend to have different notions of what constitutes a satisfying explanation, and that this is a source of misunderstanding. "Linguists tend to accept simplifications as explanations. […] For a psychologist, on the other hand, an explanation is something phrased in terms of cause and effect, antecedent and subsequent, stimulus and response" (p. 321).

[3] There are certainly cases where the subtlety of the judgements using standard methods raises the hope that larger-scale experimentation might provide more clarity about the data. However, throwing more subjects at a task is no guarantee of success. For example, experienced linguists are good at excluding effects of garden paths from their judgements and at constructing mental models that are relevant for evaluating the (un-)availability of quantifier scope ambiguities. Untutored experimental participants normally lack these skills, and so could add as much noise as clarity to a large-scale rating study.

that one sometimes encounters derives from the unrealistic expectation that psycholinguistics will provide psychological validation of linguistic models. Any kind of theory testing requires experimental tools that are commensurate with the hypotheses being tested, and for any given linguistic hypothesis there is no guarantee that the current tools of psycholinguistics are well suited for testing that hypothesis. An example of this that features prominently in sections 45.3 and 45.4 is that the detailed timing information provided by psycholinguistic and neurolinguistic measures is most revealing when evaluating hypotheses that make clear timing predictions.

In discussions about linguistics and psycholinguistics one encounters frequent references to the search for the "psychological reality" of linguistic constructs. We suspect that the term is unhelpful, since it contributes to the notion that psycholinguistic experiments license inferences about the mind that are inherently more privileged than the conclusions of lower-tech linguistic arguments. This in turn contributes to the notion that if a linguistic hypothesis does not clearly impact the tasks of speaking and understanding studied by psycholinguists then it is not a serious psychological hypothesis, and may discourage linguists from taking the psychological implications of their theories more seriously.

Another reason for linguists' frequent reluctance to take seriously the psychological implications of their theories may be the "competence-performance distinction" (Chomsky, 1965). At a basic level this is used to draw a distinction between what formal linguists do and do not consider their primary concern to be; but it is used in so many different ways that it may have led to more confusion than clarity. It is sometimes used to describe the necessary distinction between behavior and the mechanisms that generate behavior, or to describe the logical distinction between a declarative and procedural specification of a formal system. At other times it is used to refer to the difference between what a cognitive system could achieve with unbounded resources and what it can achieve when it is subject to real-life resource limitations. Finally, it is used to refer to a hypothesized division of labor between a cognitive system that specifies possible and impossible representations—the grammar—and distinct systems that generate or recover these representations in realtime—the parser and producer (for further discussion see Berwick and Weinberg, 1984; Phillips, 1996; 2004). This final hypothesis may contribute to a common misperception among linguists that they are investigating a cognitive system that is *necessarily*

distinct from what the psycholinguists are concerned with. This distinction is certainly possible, but it is an empirical hypothesis. It leaves a state of affairs where many linguists are committed mentalists, but are less certain of what their mentalistic commitments entail (e.g. what is the claim of a syntactic or phonological "derivation" a claim about?). This makes it more difficult to see where the concerns of linguists and psycholinguists are mutually relevant (see Boland, 2005 for another perspective on this issue).

The remainder of this chapter uses a case study of long-distance dependencies in linguistics and psycholinguistics to further illustrate the importance of tools that are commensurate with the hypotheses being tested.

45.2 Linguistic analyses of long-distance dependencies: a primer

45.2.1 Getting started

In this section we introduce some key properties of syntactic long-distance dependencies, and we compare different linguistic accounts of how they are encoded, emphasizing where the competing theories agree and where they disagree. The phenomena that we are concerned with here are variously known as "long-distance dependencies," "unbounded dependencies," "displacement," "extraction," or "movement." All but the last of these implies no commitment to a particular linguistic analysis. The term "movement" is generally associated with transformational grammar analyses, and we use it here only in that context. The term "long-distance dependency" often refers to a broader class of syntactic phenomena including antecedent–pronoun relations, but we primarily use it here in a sense that is interchangeable with the other terms. We will also use the standard psycholinguistic terms "filler" and "gap" to refer to the components of the dependencies (Fodor, 1978), with no commitment to a specific theoretical account intended.

In order to understand the importance of long-distance dependencies in language it is helpful to highlight the fact that local linguistic dependencies are (i) pervasive and (ii) in competition with one another. Many relations in natural language syntax appear in highly local configurations, as the examples in (3) illustrate:

(3) (a) THEMATIC DEPENDENCIES
   Marcel$_{AGENT(x)}$ memorized$_{f(x)(y)}$ a poem$_{THEME(y)}$.

(b) CASE ASSIGNMENT

His father often rebuked$_{ACC}$ him$_{ACC}$.

(c) AGREEMENT

The critics$_{3PL}$ were$_{3PL}$ initially unkind.

(d) SCOPE

Albert wonders [$_{CP}$ who [$_{C'}$ Gilbert loves]].

Often a dependent element can participate in several relationships in one configuration—for example, a direct object can receive both its case and its thematic role in a sisterhood configuration with the verb. At other times, those relationships place competing configurational demands on elements. In such cases, one or more of the dependencies must be satisfied from a non-local position. For example, in the passive construction below, the subject *the doctor* participates in a local case/agreement relation with the auxiliary verb, while it bears the thematic role most typical of a head-complement configuration with the verb *consulted*, which we find in the corresponding active construction.

(4) (a) PASSIVE

[*The doctor*$_{THEME}$] was frequently *consulted* by the diplomat.

(b) ACTIVE

The diplomat frequently *consulted* [*the doctor* $_{THEME}$]

We would like a theory of syntactic dependencies to explain how the thematic relationship between the predicate *consult* and its argument *the diplomat* is expressible in two different phrase structure configurations. Passive constructions reflect one class of displacement that retains a relatively local flavor: similar phenomena include Raising, Control, and Exceptional Case Marking. Other displacement phenomena establish relations between indefinitely distant elements. We refer to such dependences as "unbounded" or "long-distance" dependencies. Consider *wh*-interrogatives in English, as in (5).

(5) (a) The teacher said that the police falsely accused the students of the crime.

(b) The teacher said that the students were falsely accused ___ of the crime.

(c) Which students did the teacher say ___ were falsely accused ___ of the crime?

(d) The teacher said which students ___ were falsely accused ___ of the crime.

The NP *the students* is an embedded direct object in (5a), where it appears in a local relation with the verb *accuse* that assigns its thematic role. The same NP is a passive subject in (5b), where it appears in a local relation with an auxiliary

verb that it agrees with in number. In (5c, d) the NP *which students* still receives a thematic role from the same verb and governs agreement on the same auxiliary, but appears locally to neither. In these examples the position of the *wh*-phrase marks scope, i.e. whether the sentence is a direct or indirect question. The examples show that multiple syntactic relations *can* be satisfied locally, but an element normally appears in only one local configuration at a time. The other relations require non-local dependencies.

Long-distance dependencies with gaps are established in a number of other cases, such as relativization, topicalization, comparatives, and adjective–*though* constructions (6–9).

(6) RELATIVE CLAUSES

The aristocrat hired *a young maid who* he realized ___ would become his closest confidante.

(7) TOPICALIZATION

These chapters, most critics agree you can safely skip ___.

(8) COMPARATIVES

The first draft was much *longer* than anyone had suspected it to be ___.

(9) ADJECTIVE-"THOUGH"

Sophisticated though he thought his friends were ___, they failed to catch the obscure allusion.

These instances of syntactic action-at-a-distance are recognized in all theories of syntax; but different theories have different means of encoding these phenomena, and there has been much interest in finding linguistic and psycholinguistic evidence that might choose among the competing theories.

45.2.2. Competing accounts of long-distance dependencies.

Long-distance dependencies create a separation between the position where a phrase is pronounced and the verb (or other head) which determines its thematic role. Here we review several mechanisms for analyzing this separation, with an emphasis on how different theories encode long-distance dependencies, rather than on the empirical merits of their respective analyses.

We should emphasize that there is little real disagreement between theories about the notion that sentences involve multiple levels of representation. Where theories diverge is on the question of what information these different levels of representation contain, and how they are related to one another. Transformational Grammar models are famous for the claim that there are

multiple levels of representation that are specifically syntactic, and that they are related to one another by movement operations that convert each successive structure into the next; but other theories also adopt multiple levels of representation, which are sometimes non-syntactic in nature, and are related to one another by various mechanisms. All theories that we are aware of take advantage of different levels of representation in capturing the various phenomena associated with long-distance dependencies. For example, when a pronoun or reflexive element is contained inside a displaced NP, it generally retains the coreference possibilities that it would have if the NP were not displaced ("reconstruction" effects). In some theories this parallel between displaced and non-displaced NPs is captured in syntactic terms, in others this can be captured in terms of a semantic or argument structure level of representation.

45.2.2.1. Transformational accounts

The earliest models in transformational generative grammar (TGG; Chomsky, 1957; 1965) accounted for displacement in a purely derivational manner. The surface word order of a sentence was taken to be derived by first forming an underlying phrase structure, generated by rewriting phrase structure rules, and by then applying successive structural transformations to this initial representation. In the development of transformational grammar identified with Chomsky's *Aspects* model (Chomsky 1965), the initial structure or "deep structure" was taken to encode the thematic relations of a sentence, and was also taken to be the primary encoding of sentence meaning. Application of transformations yielded a "surface structure" representation that served as the primary interface with phonological systems. When an argument was moved by a transformational operation, its initial local relation to its thematic role assigner was not preserved. In these accounts the relationship between a filler and its gap was encoded in the underlying representation and the derivational history, but crucially not in the surface structure.

45.2.2.2. Transformations with traces

In the early TGG models the relationship between a displaced argument and its predicate was encoded in the same way as the relationship between other predicates and arguments—through local phrase structure composition in the generation of deep structure—but this configuration was not retained in surface structures. The introduction of phonologically null

categories, or *traces*, into TGG models in the 1970s (Chomsky, 1973; Fiengo, 1977) effectively endowed surface structures with a record of the transformations that had taken place. Displacement operations continued to be captured by means of transformational operations, but filler-gap dependencies could now be encoded in surface structure representations, as in the example in (10). Thus, a grammar that encodes long-distance dependencies using phonologically empty categories is in no way logically dependent on a transformational system.

(10) [Which letter]$_i$ did Marcel write t_i to his mother?

While traces, and empty constituents generally, have played an important role in many TGG models, like Government and Binding (GB) theory (Chomsky, 1981), Tree-Adjoining Grammar (Kroch and Joshi, 1985; Frank, 2002), and Minimalist Program models (Chomsky, 1995), they have also featured in some versions of Generalized Phrase Structure Grammar (GPSG: Gazdar et al., 1985) and Head-driven Phrase Structure Grammar (HPSG: Pollard and Sag, 1994), approaches that explicitly reject transformational derivations. Indeed, much of the work that traces accomplish in classical GB theory results from their interaction with representational well-formedness constraints, as opposed to derivational conditions on transformations. Recognition of this point has led a number of syntacticians working in the TGG tradition to propose theories that use empty categories but lack transformational derivations (e.g. Koster, 1978; Rizzi, 1986; Brody, 1995).

Some recent proposals in the context of the Minimalist Program (Chomsky 1995) have argued that traces should be replaced with a notion of unpronounced copies of the displaced phrase. However, these proposals retain the crucial feature of all transformational theories, namely that a predicate is related to a displaced argument in exactly the same way it is related to a non-displaced argument—by local phrase structure relations. It is this property that has led to a search for decisive psycholinguistic evidence.

45.2.2.3 Path marking with category labels

We have already seen that a long-distance dependency between a displaced phrase and an empty category can be encoded in a non-transformational grammar. However, a long-distance dependency of this kind cannot be directly generated in a grammar which uses standard syntactic categories and restricts itself to only context-free phrase

structure rules, since there can be no rule that directly relates the filler and the (indefinitely distant) gap. However, if we distinguish a category that dominates a gap and one that does not, then we can encode the dependency using context-free phrase structure rules. Call the category dominating a verb and a gap VP_{GAP}, and its dominating category S_{GAP}. If we admit a rule that rewrites S as a WH phrase and an S_{GAP} (11), then we can create a chain of local links between a displaced constituent and a gap position. In effect, the category label encodes a GAP feature that is passed through the tree between the wh-phrase and the gap, across a potentially unbounded distance.

(11) $S \rightarrow WH\ S_{GAP}$
  $S_{GAP} \rightarrow NP\ VP_{GAP}$
  $VP_{GAP} \rightarrow V\ GAP$

However, once the GAP-feature passing mechanism is introduced, one could take the next step and make a lexical distinction between those verbs that combine with an overt constituent and those capable of linking to a higher constituent via the GAP-feature passing mechanism (12). This long-distance linking mechanism raises the possibility of doing without gaps altogether, and relying instead on the passing of GAP-features. Current analyses in HPSG exemplify this approach (Pollard and Sag, 1994; Sag et al., 2003).

(12) $VP_{GAP} \rightarrow V_{GAP}$

For concreteness, consider the two sentences in (13), the first a multi-clause declarative, the other a corresponding topicalization.

(13) (a) The gossip columnist knew the publisher rejected the dilettante's manuscript.
  (b) The dilettante's manuscript, the gossip columnist knew the publisher rejected.

Sag et al. (2003) assume that the embedded verbs in the two examples have different but related feature specifications. In both cases the verb *rejected* is specified as taking two arguments. In (13a) one of those arguments belongs to the verb's complement list, COMPS; however, in (13b), that same argument has been moved to the verb's GAP list (a lexical feature that contains information about missing arguments) and the verb's COMPS list is null. Figure 45.1 is a schematic HPSG representation of sentence (13b). The contents of the GAP feature are inherited by successive phrase structure nodes that dominate the verb until the argument listed in the GAP list can be bound to a corresponding displaced argument. In Figure 45.1 the completion of the dependency is marked by the fact that the GAP feature is empty in the top-level S node.

Hence, approaches like HPSG provide the tools to syntactically mark a path that connects structurally distant participants in a dependency, and therefore allow the encoding of non-local predicate-argument relations, reducing the need for empty categories.[4] The key difference between this approach and theories that use empty categories to represent long-distance dependencies lies in whether predicate–argument relations are syntactically encoded in a uniform fashion.[5]

45.2.2.4 Beyond constituency

In the foregoing frameworks, the challenge is to encode non-local dependencies within the notion of constituency offered by standard phrase structure grammars. Those approaches have devised systems that relate arguments with their non-constituent predicates, either through identity with a sister of the predicate, as in trace-based theories, or through feature inheritance and matching. Alternately one might extend the notion of what counts as a constituent. Combinatory Categorial Grammar (CCG: Steedman, 2000) exemplifies this approach. CCG is a species of categorial grammar (Ajudkiewicz, 1935; Bar-Hillel, 1953), a lexicalized formalism that assigns to expressions the syntactic types either of a function or an argument. The syntactic type controls the combinatory possibilities of a given expression and it is an idiosyncratic property of individual lexical items. Predicates missing arguments in canonical positions can be established as constituents, by means of rules like Functional Composition and Type Raising. In this way CCG could be viewed as sharing the feature inheritance property of HPSG analyses, albeit in a derivational fashion. In another approach, Lexical-Functional Grammar (Bresnan, 2000) encodes dependencies via mappings between "c-structure" (constituent structure) and other levels of representation, such as "f-structure" (function structure, which represents grammatical roles like

---

[4]As should be clear from the discussion here, the use of sequences of local feature passing relations to encode long-distance dependencies reduces the need for empty categories, but does not exclude their use. Accordingly, one finds a number of transformational theories that exploit the equivalent of local feature passing mechanisms (e.g. Kayne, 1984; Manzini, 1992). These mechanisms have proven to be particularly useful for capturing constraints on long-distance dependencies in which the dependency is blocked by an element that intervenes between the filler and the gap.
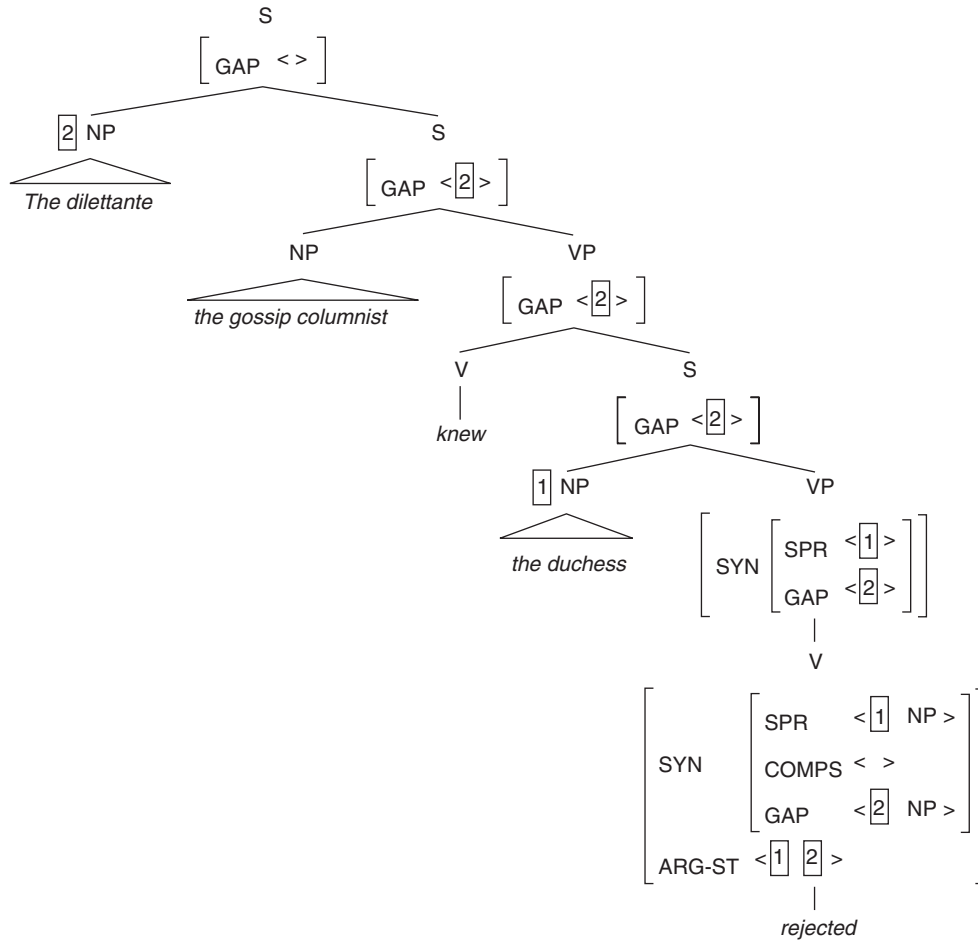
Figure 45.1 Schematized HPSG representation of sentence (13b). ARG-ST: argument structure list; COMPS: complements list (for the object arguments); SPR: specifier list (for the subject argument); GAP: gap list (for missing arguments).

Subject, Object, etc.) and "a-structure" (argument structure, which represents argument/thematic roles).

Because the experimental studies discussed in section 45.3 focus on whether or not long-distance dependencies involve traces, we do not detail the analyses of long-distance dependencies found in CCG, LFG, etc., for which there are many readable introductions (CCG: Steedman and Baldridge, 2003; LFG: Bresnan, 2001). While there are a number of different formal accounts of long-distance dependencies, it is important not to lose sight of the fact that many central insights are shared across frameworks.

45.3 **Long-distance dependencies and the status of gaps**

Although the representations posited by grammatical theories are largely motivated by distributional analyses based upon native speaker intuitions, there has been recurring interest in whether psycholinguistic evidence can be brought to bear on theoretical controversies. In the case of long-distance dependencies in particular, psycholinguistics has sometimes been viewed as a kind of appellate court that might decide in favor

of one class of analyses or another. In this section and the next we discuss psycholinguistic findings relevant to the status of gaps and to constraints on long-distance dependencies, respectively, and argue that psycholinguistic arbitration has been most effective when its tools are commensurate with the linguistic hypotheses being tested.

Psycholinguistic evidence on the status of gaps has consisted principally of information about the time course of long-distance dependency construction and the timing of semantic activation of displaced phrases.[6]

It is by now relatively uncontroversial that the parser completes long-distance dependencies without waiting for unambiguous evidence for the position of the gap. Having identified a displaced filler, the parser posits a gap at the first position that might allow satisfaction of the filler's thematic requirements. This corresponds to what Fodor (1978) describes as a "filler-driven" parsing mechanism, contrasting with a "gap-driven" alternative. An important line of evidence comes from the "filled-gap effect," a temporary disruption in reading times upon encountering an NP where a gap had been expected (Crain and Fodor, 1985; Stowe, 1986). Stowe compared self-paced reading times for sentences containing a displaced *wh*-phrase like (14a) with closely matched sentences that lacked displacement (14b). She found increased reading times at the direct object NP *us* in (14a) relative to (14b) and interpreted this

as a surprise effect, resulting from interpreting the *wh*-phrase as a displaced direct object as soon as the verb *bring* is reached, and before finding direct evidence for a direct object gap. This approach of forming linguistic dependencies before key bottom-up information is available is commonly referred to as "active" dependency formation.

(14) (a) My brother wanted to know who Ruth will bring *us* home to at Christmas
(b) My brother wanted to know if Ruth will bring *us* home to Mom at Christmas

Active construction of filler-gap dependencies has been observed in many languages, including Dutch (Frazier, 1987; Frazier and Flores D'Arcais, 1989; Kaan, 1997), Russian (Sekerina, 2003), Hungarian (Radó, 1999), Italian (de Vincenzi, 1991), German (Schlesewsky et al., 2000), and Japanese (Aoshima et al., 2004). Furthermore, evidence for active dependency formation comes from a number of paradigms, include event-related potentials (ERPs: Garnsey et al., 1989; Kaan et al., 2000; Phillips et al., 2005); plausibility measures in eye-tracking or self-paced reading (Traxler and Pickering, 1996; Phillips 2006) and in the "stops making sense" task (Tanenhaus et al., 1985); cross-modal lexical priming (Nicol and Swinney, 1989; Nicol et al., 1994); and head-mounted eye-tracking (Sussman and Sedivy, 2003). We review examples from these paradigms below.

The filled gap effect provides information about the timing of long-distance dependency formation, but it is compatible with differing accounts of how long-distance dependencies are encoded. The filled gap effect in a sentence like (14a) shows only that the long-distance dependency is completed *at some time before* the overt direct object NP *us*. If we suppose that the dependency is formed at the position of the verb *bring*, then the timing evidence is compatible with a trace-based account in which an empty category is constructed as the sister of the verb *bring* as soon as the verb is reached, and is also compatible with a trace-free account in which a link is forged between the filler and the verb as soon as the verb is reached. Either account predicts a surprise effect at the following overt object NP.

Findings from other techniques provide similar evidence on the timing of dependency formation. For example, in a plausibility manipulation paradigm, Traxler and Pickering (1996) recorded eye movements while participants read sentences like (15).

---

[6] In this context one sometimes encounters discussion of the rise and fall of the Derivational Theory of Complexity (DTC) in the 1960s as evidence in favor of gap-less theories. We will leave aside this literature here (for contrasting accounts see Townsend and Bever, 2001; Phillips, 1996), since we see it as orthogonal to questions about the status of gaps. The DTC was the hypothesis that the "perceptual complexity" of a sentence was directly proportional to the number of steps in its transformational derivation. A common argument is that the DTC was resoundingly disconfirmed in the late 1960s and that this therefore argues against transformational theories of grammar. This argument is not relevant to our current concerns, for at least two reasons. First, DTC was a hypothesis about transformations, and as we have emphasized, empty categories and transformational rules are independent syntactic constructs. Second, what the DTC studies are purported to have shown is that it is hard to find an experimental measure that is proportional to the total number of transformational operations in the derivation of a sentence. The search for such a measure may be seen as a substitute for measures of individual parsing operations, which were hard to obtain with the tools of the 1960s. Today we still have no global measure that co-varies with the number of transformational operations in a sentence, but this point is moot, since more sensitive psycholinguistic measures have made it easier to track individual parsing operations.

(15) (a) That's the pistol with which the heart-
less killer *shot* the hapless man yesterday
afternoon ____.
(b) That's the garage with which the
heartless killer *shot* the hapless man
yesterday afternoon ____.

(15a) has a perfectly sensible, plausible inter-
pretation, whereas (15b) is semantically anom-
alous, due to the predicate *shoot* taking *garage* as
an instrument argument. Traxler and Pickering
show that this anomaly is detected as soon as the
verb is reached. This indicates that the parser has
formed the dependency at least by this point.
A closely related result can be found in an ERP
study by Garnsey et al. (1989), who varied the
plausibility of the filler-verb combination in an
embedded question in sentences like (16), and
observed detection of the semantic anomaly at the
verb, as indexed by the N400 evoked response.

(16) The businessman knew which {customer |
article} the secretary called ____ at home.

A series of recent ERP studies in English pro-
vide a different index of long-distance depend-
ency completion. Processing of the verb that
allows completion of a *wh*-dependency elicits a
posterior positivity relative to the same verb in a
sentence without a *wh*-dependency (17ab: Kaan
et al., 2000). Kaan and colleagues use this finding
to suggest that the P600 is an index of "syntactic
integration difficulty" in general. Although the
interpretation of this effect remains uncertain
(cf. Fiebach et al., 2002; Phillips et al., 2005), its
timing again shows that long-distance depend-
encies are formed as soon as an appropriate verb
is encountered.

(17) (a) NO WH-DEPENDENCY
Emily wondered whether the performer
in the concert had *imitated* a pop star
for the audience's amusement.
(b) WH-DEPENDENCY
Emily wondered which pop star the
performer in the concert had *imitated*
for the audience's amusement.

Pickering and Barry (1991) use the fact that
filler-verb relations are constructed immediately
at the verb to argue that empty categories cannot
be mediating the filler-verb relation. Pickering
and Barry pay particular attention to cases where
the verb and the putative empty category are
separated by another constituent. Their primary
argument is based on examples like those in (18).
The double object construction in (18a) is notice-
ably difficult to process, presumably because the
length and complexity of the recipient argument

of the verb *give* makes it harder to associate the
theme argument *a prize* with the verb. However,
displacement of the theme argument (18b) makes
it considerably easier to process. Pickering and
Barry argue that a trace-based account incorrectly
predicts that (18b) should cause the same pro-
cessing load as (18a), since the representation of
(18b) would contain a trace in the same location
as *a prize* in (18a). On the other hand, a trace-
free theory could relate the displaced argument
the prize in (18b) to the verb *give* as soon as the
verb is reached, accounting for the reduced pro-
cessing difficulty. A similar argument can be
constructed based on Traxler and Pickering's
eye-tracking study, illustrated in (15).

(18) (a) We gave [every student capable of
answering every single tricky question
on the details of the new and extremely
complicated theory about the causes
of political instability in small nations
with a history of military rulers]
[*a prize*]
(b) That"s *the prize* that we gave [every
student capable of answering every
single tricky question on the details of
the new and extremely complicated
theory about the causes of political
instability in small nations with a
history of military rulers]

The argument developed by Pickering and
Barry involves mapping a representational claim
onto a timing prediction. They assume that if
fillers are linked to verbs through the mediation of
an empty category, then the timing of this opera-
tion should coincide with the linear position of
the empty category in the sequence. Their argu-
ment is therefore only as strong as the timing
prediction. As pointed out in various replies
(Gibson and Hickok, 1993; Gorrell, 1993; Crocker,
1994) the parser might easily construct an empty
category position in advance of its linear posi-
tion, such as by projecting argument positions
as soon as a verb is reached. If this assumption
about the parser is adopted, then the predicted
timing contrast between the competing theories
is neutralized.

Using related logic, one might look to the
processing of head-final languages in search of
decisive evidence on the status of empty cate-
gories. However, the arguments in this area have
the same limitations as Pickering and Barry's
argument, but in the opposite direction. In head-
final languages like Japanese, all arguments canon-
ically appear before the verb. Therefore, in a
trace-based representation of filler-gap depend-
encies in such languages, the position of the

empty category appears before the verb. In the spirit of Pickering and Barry's argument, one might suppose that a trace-based representation would allow filler-gap dependencies to be completed before the verb is reached, whereas a trace-free theory would delay completion of the dependency until the verb. And, indeed, Aoshima and colleagues have presented evidence for pre-verbal dependency completion using a Japanese adaptation of the filled gap effect paradigm (Aoshima et al., 2004). They show that, when a fronted dative NP is processed, a dative NP in an embedded clause engenders a reading-time slow-down (with respect to an embedded dative NP in a structure without a long-distance dependency). Based on this evidence, Aoshima and colleagues conclude that filler-gap dependencies can be formed in advance of the verb in Japanese, and also suggest that this favors a trace-based representational model. However, this argument in favor of empty categories has exactly the same weakness as Pickering and Barry's argument against empty categories. If the parser for Japanese allows the verb position to be constructed before the overt verb is reached, then direct filler-verb relations may be constructed in advance of the verb. Similar concerns apply to the timing-based argument for traces presented by Lee (2004), who demonstrates a filled gap effect in subject positions in English.

The experimental paradigms discussed so far all provide useful information about the timing of dependency completion, yet they yield few solid conclusions about what representations are being constructed. One limitation of these measures is that they often rely on semantically-based effects to draw inferences about syntactic computations. Implausibility detection measures implicate semantic representations; the filled-gap effect is ambivalent between a syntactic or semantic explanation; and the ERP P600 effects at best suggest that something syntactic happens when a verb is processed following a displaced phrase, although even this interpretation is not certain. A final potential source of evidence comes from studies of filler reactivation.

Drawing on evidence from on-line cross-modal lexical priming tasks (Nicol and Swinney 1989; Nicol et al., 1994) and off-line probe recognition scores (McElree and Bever, 1989), it has been argued that fillers are lexically reactivated at the gap site. This phenomenon might be taken as evidence that displaced constituents combine with verbs in the same way that arguments in canonical positions combine with verbs, and by extension could be presented as evidence in favor of a trace-based analysis. When these findings first appeared there was considerable interest in the possibility that they might constitute psycholinguistic confirmation of the "psychological reality" of traces (and in certain linguistic circles one still hears them discussed in such terms). However, they are subject to the now familiar limitations. First, the reactivation of a lexical code at the verb or gap site does not clearly favor one representational approach over another. Reactivation of the lexical code of the filler may reflect construction of an empty category, or may equally reflect construction of a direct link from a verb to the filler. Gap-site priming effects have also been reported in preverbal positions in head-final languages (German scrambling: Clahsen and Featherston, 1999; Japanese scrambling: Nakano et al., 2002). However, just as with the Japanese filled gap effect, such arguments depend on the questionable assumption that verb positions are constructed only when the overt verb is reached in such languages. Second, the reactivated code need not be strictly lexical, since only the contextually relevant meaning of an ambiguous filler is reactivated at the gap site (Love and Swinney, 1996). Consequently, we are ultimately left with further information on the time-course of semantic interpretation that is consistent with multiple syntactic accounts.

In the absence of theoretically decisive timing arguments, it is sometimes claimed that considerations of parsimony should favor a trace-free theory (Pickering, 1993) or that the behavioral results most directly implicate a trace-free account (Sag and Fodor, 1994). Why posit multiple levels of syntactic representation if one level will suffice? Why appeal to phonologically empty categories if we can do without them? We find such arguments to be somewhat disingenuous, for a couple of reasons. First, competing theories agree on the need for multiple levels of representation for sentences, but disagree on the issue of how many of these levels are syntactic or semantic and how the levels are related to one another. Phenomena that are explained in terms of empty categories in one theory must be accounted for using other machinery in a theory that lacks empty categories. Simplifying one's syntax often leads to complications at other levels of representation (cf. Jackendoff 2002: 144–8). Second, while we acknowledge the need for caution in proliferating the inventory of empty categories in the grammar, it strikes us as odd to claim that phonologically empty syntactic formatives are inherently objectionable. Perceptual processes, whether in language or in other domains, involve the construction of many kinds of mental object that do not correspond to a clearly defined sensory stimulus.

In speech perception we detect segments that are masked or absent in the input; in vision we perceive objects that are not present in the distal stimulus; in syntax we perceive combinatorial structure that is not encoded in the phonological form of a sentence. Taken in this context, the notion of empty categories is rather banal. This does not, of course, entail that they are needed, merely that one should continue to search for good empirical arguments rather than falling back on questionable claims of parsimony.

In sum, a constellation of findings spanning multiple experimental approaches have converged in support of the idea that long-distance dependency formation is a rapid, top-down process. Moreover it is an active process, occurring as soon as there is sufficient information to posit a gap, often at the point of encountering the verb, and in head-final languages in advance of the verb. However, this is a timing result that does not clearly correlate with particular ways of encoding the relation between a verb and a displaced argument. In this area psycholinguistic tools have so far proven inconclusive. We see this more as a practical failure than a principled one. More ingenious methods of probing syntactic representations might yet succeed in distinguishing the competing theories.

45.4 Experimental studies of constraints on dependencies

A possible moral of the previous section is that psycholinguistic measures of timing can be used to resolve representational controversies only to the extent that the competing representational accounts yield clear predictions about timing. In this section we consider two issues that may be better suited to psycholinguistic testing, specifically because they involve clearer timing predictions. By exploring how realtime language processing is affected by constraints on long-distance dependencies, we can address whether linguistic and psycholinguistic models should be viewed as accounts of the same underlying mechanisms, and also the extent to which grammatical constraints might be reducible to constraints on language processing.

45.4.1 Island constraints

Although filler-gap dependencies may span long distances, they are also subject to a number of restrictions that have attracted substantial interest in linguistics since classic studies in the mid-1960s (Chomsky, 1964; Ross, 1967). Following terminology introduced by Ross, contexts that block filler-gap dependencies are widely known as "islands." Syntactic islands include relative clauses (19a), *wh*-clauses (19b), factive clauses (19c), subjects (19d), adjuncts (19e), and coordinate structures (19f).

(19) (a) *What did the agency fire the official that recommended ___?
(b) *Who do you wonder whether the press secretary spoke with ___?
(c) *Why did they remember that the corrupt CEO had been aquitted ___?
(d) *What did the fact that Joan remembered ___ surprise her grandchildren?
(e) *Who did Susan watch TV while talking to ___ on the phone?
(f) *What did the Senate approve ___ and the House reject the bill?

There have been numerous attempts to capture the common property or properties that underlie these and other island constraints. For example, Chomsky's (1973) Subjacency Constraint captured the effects of a number of islands under a constraint that blocks filler-gap dependencies that cross two or more bounding nodes (NP or S) in one step. A number of good summaries of different formal accounts of islands are available (e.g. Manzini, 1992; Culicover, 1997). However, our concern here is less with adjudicating competing formal accounts, and more with how these constraints impact the relation between linguistic and psycholinguistic models.

There is widespread skepticism over the issue of whether linguistic and psycholinguistic models are concerned with the same mental phenomena. Psycholinguists are often suspicious of linguists' obsession with ephemeral constructions which rarely occur in real life situations. Meanwhile, in linguistics there is a long-standing tradition of distancing theoretical models from claims about realtime processes. In an influential statement of objectives for the field Chomsky states: "When we say that a sentence has a certain derivation with respect to a particular generative grammar, we say nothing about how the speaker or hearer might proceed, in some practical or efficient way, to construct such a derivation" (Chomsky 1965: 9). In support of this position it has been claimed that the parser initially builds coarse-grained representations that lack the detail required of the grammar (Townsend and Bever, 2001) and that language is not "readily usable" (Chomsky and Lasnik, 1993: 18). In this context, it is relevant to ask whether realtime language processes are sensitive to constraints on filler-gap dependencies and other phenomena that are central concerns of grammatical theory. To the extent that these are reflected in realtime language processing

mechanisms, there is more reason to think that linguists and psycholinguists are concerned with the same mental representations.

45.4.2 The timing of island constraints

A number of studies have addressed the impact of island constraints on realtime language processing. Most published studies on this topic have concluded that island constraints do impact language processing, and in doing so have shown different ways in which island constraints may be reflected in comprehension processes.

One line of research asks whether the parser suspends its normal "active" search for gaps in positions where this would lead to an island constraint violation. Typically, the logic of these studies is to show that a manipulation that yields a measurable experimental effect when a well-formed gap is posited yields a null effect at comparable positions inside syntactic islands. For example, Stowe (1986: experiment 2) followed up on her demonstration of the filled gap effect (FGE) by showing that the FGE is not observed inside a syntactic island. The NP *Greg's* in (20a) is the object of a complement PP, and thus occupies a potential gap site for the fronted wh-phrase, as in a sentence like (22a). This NP was read more slowly in (20a) than in a control condition that lacked *wh*-fronting (20b), an FGE that suggests that the parser actively posited a gap site in the prepositional object position. In contrast, the NP *Greg's* in (21a) is embedded inside a subject NP. Subjects are typically islands for *wh*-fronting, and hence the NP *Greg"s* does not occupy a potential grammatical gap site, as shown by the unacceptability of (22b). Stowe found no FGE at this NP, suggesting that the parser made no attempt to posit a gap inside the island. Similar findings about the disappearance of the FGE in island environments have been reported in French (Bourdages, 1992) and Japanese (Yoshida et al., 2004).

(20) (a) The teacher asked what the team laughed about Greg's older brother fumbling.
    (b) The teacher asked if the team laughed about Greg's older brother fumbling the ball.
(21) (a) The teacher asked what the silly story about Greg's older brother was supposed to mean.
    (b) The teacher asked if the silly story about Greg's older brother was supposed to mean anything.
(22) (a) The teacher asked what the team laughed about ___.

    (b) *The teacher asked what the silly story about ___ was supposed to mean.

Other studies have made a related argument using the plausibility manipulation paradigm introduced above. Traxler and Pickering (1996) showed that manipulation of the semantic plausibility of a filler–verb combination elicited an immediate reading-time slowdown at the verb in examples like (23), but no corresponding slowdown at the same verb when it appeared inside a relative clause (24), again suggesting immediate effects of island constraints.[7]

(23)   *Preamble: Waiting for a publishing contract*
    The big city was a fascinating subject for the new book.
    (a) We like the book that the author wrote unceasingly and with great dedication about while waiting for a contract.

We like the city that the author wrote unceasingly and with great dedication about while waiting for a contract.

(24) (a) We like the book that the author who wrote unceasingly and with great dedication saw while waiting for a contract.
    (b) We like the city that the author who wrote unceasingly and with great dedication saw while waiting for a contract.

In contrast to the studies that have shown the absence of active dependency completion effects in island environments, a number of other studies have demonstrated processing disruption when the search for a gap encounters the boundary of a syntactic island. Three different ERP studies have measured the effect of encountering an island boundary while searching for a gap site for a filler. McKinnon and Osterhout (1996) compared ERP responses elicited by sentences like (25a), containing an illicit extraction from a *when*-clause, with closely matched sentences that lack an island constraint violation (25b). The P600 response characteristic of responses to syntactic anomalies was elicited at the word *when*, indicating that comprehenders are immediately sensitive to island domains while processing filler-gap dependencies. However, this effect is open to multiple interpretations: the P600 response may

---

[7] Wagers and Phillips (2006) demonstrate island sensitivity without relying on a null effect. They show plausibility sensitivity in the second conjunct of a coordinate structure, indicating sensitivity to the Coordinate Structure Constraint.

reflect calculation of ill-formedness in a formal account of island constraints, or it may instead reflect disruption to the process of searching for a gap by the presence of a second *wh*-phrase. These two possible interpretations are represented in two other ERP studies of island effects that have elicited left anterior negativity (LAN) effects one word after the beginning of an island domain (Neville et al.,1991; Kluender and Kutas, 1993).

(25) (a) *I wonder which of his staff members the candidate was annoyed *when* his son was questioned by.
     (b) I wonder whether the candidate was annoyed *when* his son was questioned by his staff member.

Related effects can be found in a study by McElree and Griffith (1998) using a "speed–accuracy tradeoff" (SAT) paradigm in which participants were trained to give acceptability judgements immediately upon hearing a tone that occurs at specific intervals after the end of a test sentence. Unsurprisingly, when the tone appears very shortly after the sentence, accuracy is low, and when the tone appears after a longer delay, accuracy is higher. The interest of the SAT paradigm is that it allows the researcher to precisely track the time-course of increases in accuracy. McElree and Griffith show that sensitivity to an island violation begins to emerge almost immediately after the verb in sentences with relative clause island violations such as (26). This effect may reflect detection of the island violation at the verb position, but it more likely reflects detection of the violation at the preceding word *who* that begins the relative clause, which appeared only 250 ms earlier.

(26) *It was the essay that the writer scolded the editor who admired.

If island constraints can be held responsible for the disappearance of filled gap effects and plausibility effects inside islands, and for the various effects of island-boundary detection, then we can conclude that constraints on filler-gap dependencies have a more or less immediate impact upon language comprehension processes.[8] This, and related experimental findings showing the immediate effects of other grammatical constraints (e.g. binding constraints: Nicol and Swinney, 1989; Sturt, 2003; Kazanina et al., 2006),

suggest that the parser constructs the same kinds of representation that linguists are concerned with, and make it more difficult for linguists to argue that realtime processes are irrelevant to their concerns. The findings are at least consistent with the stronger position that the grammar is a realtime structure building mechanism (e.g. Phillips, 1996; 2004; Kempson eet al., 2001), but it by no means entails this view.

45.4.3 The origin of island constraints

Findings about the immediacy of island constraints do not show that the constraints are "psychologically real" in the sense that they lend greater psychological respectability to the formal linguistic description of the constraint. The experimental findings indicate that the same properties that account for the unacceptability of island violations also affect realtime comprehension processes, but do not indicate whether island constraints are more appropriately viewed as formal constraints on structures or as the products of independent constraints on memory, focus, or any other factors that might affect language processing.

In contrast to formal accounts of island constraints, it has often been suggested that island constraints may ultimately derive from limitations on realtime language processing. Some accounts assume that the island constraints are grammaticized, but ultimately owe their presence in grammars to constraints on language processing (Fodor, 1978; 1983; Berwick and Weinberg, 1984; Hawkins, 1999), whereas other accounts assume that island constraints are genuine epiphenomena that are not explicitly represented in a speaker's grammar (Deane, 1991; Pritchett, 1991; Kluender and Kutas, 1993). For example, some accounts have proposed that the impossibility of extraction from a subject NP (27) may reflect the order of structure building operations in language processing.

(27) *Who did [$_{NP}$ the news about ___] surprise everybody?

Pritchett (1991) suggests that the islandhood of subjects is a natural consequence of his "head-driven" parsing architecture, which allows the parser to start building a phrase only once the head of that phrase has been encountered. Since subject NPs precede the head of the phrase that they are a part of, e.g. an auxiliary or a verb, the head-driven architecture prevents subject NPs from being immediately attached into the parse tree. This, in turn, delays the completion of a filler-gap dependency into the subject NP,

[8] There is a small number of studies whose results suggest that gaps are posited inside syntactic islands in real time (Pickering et al., 1994: expt. 1; Clifton Frazier, 1989), but these results are open to alternative explanations.

and Pritchett suggests that this is what underlies the unacceptability of subject island violations. A related mechanism is responsible for subject island effects in a study by Hawkins (1999).

Attempts to derive island constraints from constraints on realtime structure building share a simple prediction, which can be tested using psycholinguistic methods. If the unacceptability of a gap in a given location is due to the parser's difficulty or inability to construct a gap in that location, then speakers should indeed find it difficult or impossible to construct such gaps during realtime processing. Phillips (2006) describes a test of this prediction, taking advantage of the phenomenon of "parasitic gaps" (Engdahl, 1983; Culicover, 2001)—constructions in which otherwise ill-formed gaps are rendered acceptable when they appear in a sentence with an additional well-formed gap. (28a) is another illustration of the islandhood of subject NPs. In this case, the illicit gap is inside an infinitival clause that is the complement of a subject NP. When the illicit gap in (28a) is combined with the acceptable direct object gap in (28b), the result is an acceptable sentence (28c). (These judgements have been confirmed in controlled rating studies.) The first gap in (28c) is referred to as a "parasitic gap," since its well-formedness relies upon the presence of another gap.

(28) (a) *What did the attempt to repair ___ ultimately damage the car?
      (b) What did the attempt to repair the car ultimately damage ___?
      (b) What did the attempt to repair ___pg ultimately damage ___?

The phenomenon of parasitic gaps is interesting in its own right, but parasitic gap examples like (28c) are particularly interesting from the perspective of realtime processing, since the illicit gap precedes the gap that licenses it. If the parser actively posits gaps in all positions where a well-formed gap might appear, then it should be able to create a gap upon reaching the embedded verb *repair* in sentences like (28), and should then seek an additional licensing gap in order for the sentence to be well-formed. If, on the other hand, the parser more strictly avoids positing gaps inside islands, then the parser should never construct an illicit gap like the one shown in (28a). This would imply that well-formed constructions like (28c) must be parsed in a non-incremental fashion, constructing the parasitic gap only after the licensing gap has been confirmed.

Using an implausibility detection paradigm similar to Traxler and Pickering (1996), Phillips (2006) shows that active gap creation occurs in potential parasitic gap environments just as in more familiar environments. A slowdown reflecting implausibility detection occurs immediately at the underlined verb when it appears inside an island that supports parasitic gaps, as in (29), where the subject NP contains an infinitival complement clause. This suggests that speakers actively created a gap inside the subject NP, despite its islandhood. No corresponding slowdown is observed in islands that do not support parasitic gaps. The finite relative clause in (30a) creates an island for filler-gap dependencies, but unlike the examples in (28) this violation cannot be "rescued" by combination with a well-formed gap (30b, c). The lack of a plausibility effect implies that the parser failed to posit a gap in this environment, and therefore that the parser constructs gaps inside islands in precisely the environments where the grammar of parasitic gaps makes this possible.

(29) The school superintendent learnt {which schools/which high school students} the plan to *expand* …
(30) (a) *What did the reporter that criticized ___ eventually praise the war?
      (b) What did the reporter that criticized the war eventually praise ___?
      (c) *What did the reporter that criticized ___ eventually praise ___?

These findings are directly relevant to claims that island phenomena can be reduced to effects of difficulty in the processing of filler-gap dependencies. The unacceptability of the subject island violation in (28a) cannot be due to difficulty in realtime gap creation, since the experimental results show that speakers readily create a filler-gap dependency into the subject NP. If the unacceptability of the gap in (28a) is not reducible to processing constraints, then this lends credence to a formal account of the island constraint. Of course, this argument does not necessarily extend to other types of island, even including other types of subject island like (30a). Nevertheless, it would seem odd to claim that highly unacceptable islands, such as the extraction from a relative clause in (30a), are grammatically well-formed and are epiphenomena of constraints on processing, whereas less severe violations like (28a) are grammatically ill-formed.

In sum, the studies reviewed in this section suggest that it is possible to use psycholinguistic results to learn about the form of the grammar, even in the same domain of long-distance dependencies that proved to be more difficult to test in section 45.3. The difference is that in this section we have been considering questions about the grammar that have direct timing consequences, and hence are well-suited for psycholinguistic testing.

45.5 **Conclusion**

We see no principled reason why the fields of linguistics and psycholinguistics should not have an ongoing and mutually beneficial interaction. This reflects, in part, the fact that we struggle to draw a clear distinction between the two areas. Both fields have serious mentalistic commitments, and we see neither as having more privileged access to the psychological mechanisms of language. In cases where linguists are unwilling to take their mentalistic commitments sufficiently seriously, or where psycholinguists are dismissive of the complexities that linguists spend their time worrying about, we will continue to find skepticism and suspicion from both directions. The relation between the fields has sometimes been viewed in hierarchical terms, according to which linguistics should look to psycholinguistics as a court of arbitration for its disputes, but not vice versa. We find this view, and the related notion of the "psychological reality" of linguistic constructs, to be somewhat unhelpful. The hypotheses that linguists develop on the basis of distributional analyses of informant judgements are just as psychological as hypotheses developed on the basis of analyses of complex reaction time or eye-gaze data. It is conceivable that when linguists investigate acceptability judgements they are studying a cognitive system that is distinct from the processing systems with which psycholinguists are more commonly concerned; but we should stress that this distinction is an empirical hypothesis, and one that has received very little direct testing. Therefore, in the absence of good evidence to the contrary, we assume that linguists and psycholinguists are exploring the same cognitive system, albeit with different tools. We take the case study of long-distance dependencies to show that the prospects for influence from psycholinguistics to linguistics (and vice versa) are good, and are subject to merely practical limitations. One field can successfully influence the other only when its tools are commensurate with the hypotheses that are being tested. This should come as no surprise.

Acknowledgements

References

Ajdukiewicz, K. (1935) Die syntaktische Konnexität. *Studia Philosophica*, 1: 1–27. English translation in S. McCall (ed.), *Polish Logic 1920-1939*, pp. 207–31. Oxford University Press, Oxford.

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

Bar-Hillel, Y (1953) A quasi-arithmetical notation for syntactic description. *Language*, 29: 47–58.

Berwick, R., and Weinberg, A. S. (1984) *The Grammatical Basis of Linguistic Performance*. MIT Press, Cambridge, Mass.

Boland, J. E. (2005) Cognitive mechanisms and syntactic theory. In A. Cutler (ed.,), *Twenty-First Century Psycholinguistics: Four Cornerstones*, pp. 23–42. Erlbaum, Mahwah, NJ.

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

Bresnan, J. (1978) A realistic transformational grammar. In J. Bresnan, M. Halle, and G– Miller (eds.), *Linguistic Theory and Psychological Reality*, pp. 1–59. MIT Press, Cambridge, Mass.

Bresnan, J. (2001) *Lexical-Functional Syntax*. Blackwell, Oxford.

Brody, M. (1995) *Lexico-logical Form*. MIT Press, Cambridge, Mass.

Chomsky, N. (1957) *Syntactic Structures*. Mouton, The Hague.

Chomsky, N. (1964) *Current Issues in Linguistic Theory*. Mouton, The Hague.

Chomsky, N. (1965) *Aspects of the Theory of Syntax*. MIT Press, Cambridge, Mass.

Chomsky, N. (1973) Conditions on transformations. In S. Anderson and P. Kiparsky (eds.), *A Festschrift for Morris Halle*, pp. 232–86. Holt, Rinehart & Winston, New York.

Chomsky, N. (1981) *Lectures on Government and Binding*. Foris, Dordrecht.

Chomsky, N. (1995) *The Minimalist Program*. MIT Press, Cambridge, Mass.

Chomsky, N., and H. Lasnik, (1993) The theory of Principles and Parameters. In J. Jacobs, W. Sternefeld, and T. Vennemann (eds.), *Syntax: An International Handbook of Contemporary Research*, pp. 506–69. Berlin: de Gruyter. Repr. in Chomsky (1995: 13–127).

Clahsen, H., and Featherston, S. (1999) Antecedent-priming at trace positions: evidence from German scrambling. *Journal of Psycholinguist Research*, 28: 415–37.

Clifton, C. E., Jr, and Frazier, L. (1989) Comprehending sentences with long-distance dependencies. In MK Tanenhaus and G. N. Carlson (eds.), *Linguistic Structure in Language Processing*, pp. 273–317. Dordrecht: Kluwer Academic.

Crain, S., and Fodor, J. D. (1985) How can grammars help parsers? In D. Dowty, L. Kartunnen, and A. M. Zwicky (eds.), *Natural Language Parsing: Psycholinguistic, Computational, and Theoretical Perspectives*, pp. 94–128. Cambridge University Press, Cambridge.

Crocker, M. W. (1994) On the nature of the principle-based sentence processor. In C. Clifton, Jr, L. Frazier, and K. Rayner (eds.), *Perspective on Sentence Processing*, pp. 245–66. Erlbaum, Hillside, NJ.

Culicover, P. (1997) *Principles and Parameters: An Introduction to Syntactic Theory*. Oxford University Press, Oxford.

Culicover, P. (2001) Parasitic gaps: a history. In P. S. Culicover and P. Postal (eds.), *Parasitic Gaps*, pp. 3–68. MIT Press, Cambridge, Mass.

Deane, P. (1991) Limits to attention: a cognitive theory of island constraints. *Cognitive Linguistics*, 2: 1–63.

de Vincenzi, M. (1991) *Syntactic Parsing Strategies in Italian*. Kluwer Academic, Dordrecht.

Engdahl, E. (1983) Parasitic gaps. *Linguistics and Philosophy*, 5: 5–34.

Ferreira, F. (2005) Psycholinguistics, formal grammars, and cognitive science. *Linguistic Review*, 22: 365–80.

Fiebach, C. M., Schlesewsky, M., and Friederici, A. D. (2002) Separating syntactic memory costs and syntactic integration costs during parsing: the processing of German wh-questions. *Journal of Memory and Language*, 47: 250–72.

Fiengo, R. (1977) On trace theory. *Linguistic Inquiry*, 8: 35–62.

Fodor, J. D. (1978) Parsing strategies and constraints on transformations. *Linguistic Inquiry*, 9: 427–73.

Fodor, J. D. (1983) Phrase structure parsing and the island constraints. *Linguistics and Philosophy*, 6: 163–223.

Frank, R. (2002) *Phrase Structure Composition and Syntactic Dependencies.* MIT Press, Cambridge, Mass.

Frazier, L. (1987) Syntactic processing: Evidence from Dutch. *Natural Language and Linguistic Theory*, 5: 519–60.

Frazier, L., and Flores D"Arcais, G. B. (1989) Filler-driven parsing: a study of gap filling in Dutch. *Journal of Memory of Language*, 28: 331–44.

Garnsey, S. M., Tanenhaus, M. K., and Chapman, R. M. (1989) Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research*, 18: 51–60.

Gazdar, G., Klein, E., Pullum, G., and Sag, I. (1985) *Generalized Phrase Structure Grammar*. Harvard University Press, Cambridge, Mass.

Gibson, E., and Hickok, G. (1993) Sentence processing with empty categories. *Language and Cognitive Processes*, 8: 147–61.

Gorrell, P. (1993) Evaluating the direct association hypothesis: a reply to Pickering and Barry (1991), *Language and Cognitive Processes*, 8: 129–46.

Gouvea, A., Phillips, C., Kazanina, N., and Poeppel, D. (2005) The syntactic processes underlying the P600. MS submitted for publication.

Hawkins, J. (1999) Processing complexity and filler-gap dependencies across languages. *Language*, 75: 224–85.

Jackendoff, R. (2002) *Foundations of Language*. Oxford University Press, New York.

Kaan, E. (1997) Processing subject–object ambiguities in Dutch. Doctoral dissertation, University of Groningen.

Kaan, E., Harris, A., Gibson, E., and Holcomb, P. (2000) The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, 15: 159–201.

Kamide, Y., and Mitchell, D. C. (1999) Incremental pre-head attachment in Japanese parsing. *Language and Cognitive Processes*, 14: 631–62.

Kayne, R. (1984) *Connectedness and Binary Branching*. Foris, Dordrecht.

Kazanina, N., Lau, E., Lieberman, M., Yoshida, M., and Phillips, C. (2006) Effects of syntactic constraints on the processing of backward anaphora. MS, submitted for publication.

Kempson, R., Meyer-Viol, W., and Gabbay, D. (2001) *Dynamic Syntax*. Blackwell, Oxford.

Kluender, R., and Kutas, M. (1993) Subjacency as a processing phenomenon. *Language and Cognitive Processes*, 8: 573–633.

Koster, J. (1978) *Locality principles in syntax*. Foris, Dordrecht.

Kroch, A., and Joshi, A. K. (1985) The linguistic relevance of Tree Adjoining Grammar. MS-CIS-85–16: University of Pennsylvania.

Kurtzman, H. S., and Crawford, L. S. (1991) Processing parasitic gaps. In T Sherer (ed.), *Proceedings of the 21st Annual Meeting of the North East Linguistics Society*, pp. 217–31. GLSA, Amherst, Mass.

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

Love, T., and Swinney, D. (1996) Coreference processing and levels of analysis in object-relative constructions: demonstration of antecedent reactivation with the cross-modal priming paradigm. *Journal of Psycholinguistic Research*, 25: 5–24.

Manzini, M. R. (1992) *Locality*. MIT Press, Cambridge, Mass.

McElree, B., and Bever, T. G. (1989) The psychological reality of linguistically defined gaps. *Journal of Psycholinguistic Research*, 18: 21–36.

McElree, B., and 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: 432–60.

McKinnon, R., and Osterhout, L. (1996) Constraints on movement phenomena in sentence processing: Evidence from event-related potentials. *Language and Cognitive Processes*, 11: 495–523.

McKoon, G., Allbritton, D., and Ratcliff, R. (1996) Sentential context effects on lexical decisions with a cross-modal instead of all-visual procedure. *Journal of Experimental Psychology, Language, Memory and Cognition*, 22: 1494–97.

McKoon, G., and Ratcliff, R. (1994) Sentential context and on-line lexical decision tasks. *Journal of Experimental Psychology, Language, Memory and Cognition*, 20: 1239–43.

McKoon, G., Ratcliff, R., and Ward, G. (1994) Testing theories of language processing: an empirical investigation on the on-line lexical decision task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20: 1219–28.

Miller, G. A. (1990) Linguists, psychologists, and the cognitive sciences. *Language*, 66: 317–22.

Miyara, S. (1982) Reordering in Japanese. *Linguistic Analysis*, 9: 307–40.

AU: Is it ok?

Nakano, Y., Felser, C., and Clahsen, H. (2002) Antecedent reactivation in the processing of scrambling in Japanese. *MIT Working Papers in Linguistics*, 43: 127–42.

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

Nicol, J. L., Fodor, J. D., and Swinney, D. (1994) Using cross-modal lexical decision tasks to investigate sentence processing. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 20: 1229–38.

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

Nicol, J., Swinney, D., Love, T., and Hald, L. A. (1997) *Examination of Sentence Processing with Continuous vs. Interrupted Presentation Paradigms*. Center for Human Information Processing Technical Report 97–3. University of California, San Diego.

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

Phillips, C. (2004) Linguistics and linking problems. In M. Rice and S. Warren (eds.), *Developmental Language Disorders: From Phenotypes to Etiologies*, pp. 241–87. Erlbaum, Mahwah, NJ.

Phillips, C. (2006) The realtime status of island phenomena. *Language*, 82: 795–823.

Phillips, C., Kazanina, N., and Abada, S. H. (2005) ERP effects of the processing of syntactic long-distance dependencies. *Cognitive Brain Research*, 22: 407–28.

Pickering, M. (1993) Direction association and sentence processing: a reply to Gorrell and to Gibson and Hickok. *Language and Cognitive Processes*, 8: 163–96.

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

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

Pollard, C., and Sag, I. A. (1994) *Head-Driven Phrase Structure Grammar*. University of Chicago Press, Chicago.

Pritchett, B. L. (1991) Subjacency in a principle-based parser. In R. C. Berwick (ed.), *Principle-Based Parsing: Computation and Psycholinguistics*, pp. 301–45. Kluwer Academic, Dordrecht.

Radó, J. (1999) Some effects of discourse salience on gap-filling. Poster presented at the 12th Annual CUNY Conference on Human Sentence Processing.

Rizzi, L. (1986) On chain formation. In H. Borer (ed.), *The Syntax of Pronominal Clitics*, pp. 65–76, Academic Press, New York.

Ross, J. R. (1967) Constraints on variables in syntax. Ph.D. dissertation, MIT.

Sag, I. A., and Fodor, J. D. (1994) Extraction without traces. In R. Aranovich, W. Byrne, S. Preuss, and M. Senturia (eds.), *Proceedings of the 13th Annual Meeting of the West Coast Conference on Formal Linguistics*, pp. 365–84. CSLI, Stanford, Calif.

Sag, I. A., Wasow, T., and Bender, E. M. (2003) *Syntactic Theory: A Formal Introduction*. CSLI, Stanford, Calif.

Schlesewsky, M., Fanselow, G., Kliegl, R., and Krems, J. (2000) The subject preference in the processing of locally ambiguous wh-questions in German. In B. Hemforth and L. Konieczny (eds.), *German Sentence Processing*, pp. 65–93. Kluwer Academic, Dordrecht.

Sekerina, I. A. (2003) Scrambling and processing: dependencies, complexity, and constraints. In S. Karimi (ed.), *Word Order and Scrambling*, pp. 301–24. Blackwell, Malden, Mass.

Steedman, M. (2000) *The Syntactic Process*. MIT Press, Cambridge, Mass.

Steedman, M., and Baldridge, J. (2003) Combinatory Categorial Grammar. Unpublished tutorial paper, online: http://groups.inf.ed.ac.uk/ccg/publications.html.

Stowe, L. A. (1986) Evidence for on-line gap-location. *Language and Cognitive Processes*, 1: 227–45.

Sturt, P. (2003) The time course of the application of binding constraints in reference resolution. *Journal of Memory and Language*, 48: 542–62.

Sussman, R. S., and Sedivy, J. C. (2003) The time-course of processing syntactic dependencies: evidence from eye-movements during spoken wh-questions. *Language and Cognitive Processes*, 18: 143–63.

Swinney, D., Ford, M., Frauenfelder, U., and Bresnan, J. (1988) On the temporal course of gap filling and antecedent assignment during sentence comprehension. MS.

Swinney, D. A., Onifer, W., Prather, P., and Hirshkowitz, M. (1979) Semantic facilitation across sensory modalities in the processing of individual words and sentences. *Memory and Cognition*, 7: 159–65.

Tanenhaus, M., Stowe, L., and Carlson., G. (1985) The interaction of lexical expectation and pragmatics in parsing filler-gap constructions. In *Proceedings of the Seventh Annual Cognitive Science Society Meeting*, pp. 361–5, Irvine, Calif.

Townsend, D. J., and Bever, T. G. (2001) *Sentence Comprehension: The Integration of Habits and Rules*. MIT Press, Cambridge, Mass.

Traxler, M. J., and Pickering, M. J. (1996) Plausibility and the processing of unbounded dependencies: an eye-tracking study. *Journal of Memory and Language*, 35: 454–75.

Wagers, M., and Phillips, C. (2006) (Re)active filling. Talk presented at the 19th Annual CUNY Conference on Human Sentence Processing, New York.

Wanner, E., and Maratsos, M. (1978) An ATN approach to comprehension. In M. Halle, J. Bresnan, and G. A. Miller (eds.), *Linguistic Theory and Psychological Reality*, pp. 119–61, MIT Press, Cambridge, Mass.

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