The acceptability cline in VP ellipsis

Christina S. Kim, Gregory M. Kobele,
Jeffrey T. Runner, John T. Hale*

Abstract. This paper lays the foundations for a processing model of relative acceptability levels in verb phrase ellipsis (VPE). In the proposed model, mismatching VPE examples are grammatical but less acceptable because they violate heuristic parsing strategies. This analysis is presented in a Minimalist Grammar formalism that is compatible with standard parsing techniques. The overall proposal integrates computational assumptions about parsing with a psycholinguistic linking hypothesis. These parts work together with the syntactic analysis to derive novel predictions that are confirmed in a controlled experiment.

1 Introduction

The term verb phrase ellipsis (VPE) refers to a construction in which a verb phrase (VP) that intuitively ought to appear fails to do so. Example (1a) shows a simple case.

(1) a. Jill betrayed Abby, and Matt did, too
b. Jill betrayed Abby, and Matt betrayed Abby too
c. Jill betrayed Abby, and Matt did ~~betray Abby~~, too

VPE sentences pose two main analytical problems: (i) under what conditions can a VP be omitted, and (ii) what do sentences with missing VPs mean? In connection with this second question, it is easy to see that (1a) is syn-

*The authors are grateful for the improvements to this work suggested by audiences at CUNY 2008, the Ohio State University (February 2008), University of Michigan (March 2008), Potsdam University (May 2009), Formal Grammar (July 2009), and the University of Delaware (December 2009).

onymous with (1b). Indeed, a paraphrase such as (1b) can serve as a tool to characterize the missing VP. We indicate this by ~~striking it out~~, as in (1c).

This purely notational convention already suggests an analysis. On this analysis, (1a) and (1b) have the same syntactic structure. There is no missing VP *in the syntax*. Rather, the phonological properties of the missing VP have been 'deleted' in (1a) but not in (1b). This analysis reduces the meaning-problem (i) to the problem of sentence-meaning in general. The meaning of an elliptical sentence could be computed compositionally from its phonologically unexpressed syntactic structure. However, this progress on problem (i) underlines the urgency of problem (ii). Any such analysis must appeal to conditions on the applicability of deletion in explaining the synonymy between (1a) and (1b).

In (1) as in other cases of VPE the deleted material is similar to material elsewhere in the sentence. We will refer to this material as the ANTECEDENT. Transformational grammars of the 1960s typically required deleted expressions to have an *identical* antecedent elsewhere in the clause [Lees, 1960, Chomsky, 1964]. This identity requirement, the condition on recoverability of deletion (CRD), is appealing from the standpoint of language use, since it seems to limit the space of alternatives a parser has to consider when looking for deletion sites. But the extent of any such limitation depends in detail on the precise formulation of the CRD.

The question of what this CRD-like identity criterion should be in an adequate grammar is a vexed one. One view supposes that deletion is licensed under identity of surface structure. However, this hypothesis would not allow one to treat the quite acceptable examples in (2) as grammatical.

(2) a. This information could have been released by Gorbachev, but he chose not to ~~release this information~~. [Hardt, 1993]

b. In March, four fireworks manufacturers asked that the decision be reversed, and on Monday the ICC did ~~reverse the decision~~. [Dalrymple et al., 1991]

In (2a) for instance the passive VP "been released by Gorbachev" in the first conjunct cannot be identical with elided material in the second conjunct; surrounding words are only compatible with the *active* voice. Similarly, in (2b), the elided material must have been active but the only available antecedent, "be reversed", is *passive*. The crossed-out material in these mismatching examples has been deleted but not under any pretheoretical notion

of surface identity.

The relatively high acceptability of sentences like these presents a dilemma. One response is to abandon a syntactic notion of identity. Perhaps the relevant notion of identity is semantic. Another response simply denies that examples like those in (2) are grammatical. This response calls out for some other explanation for their relatively high acceptability; section 2.2 briefly addresses these and other difficulties. The bulk of this paper develops a third kind of response, one that is founded on the contention that a strict syntactic identity criterion indeed can be upheld in a modern deletion analysis of VPE. Three sorts of background considerations motivate this approach. First, work in theoretical syntax [Tomioka, 1997, Chung, 2006, Merchant, 2007, 2008a, Kobele, 2009] suggests that the identity condition in ellipsis is sensitive to properly *syntactic* distinctions. Second, work in psycholinguistics [Tanenhaus and Carlson, 1990, Arregui et al., 2006] suggests that comprehenders are indeed sensitive to such fine-grained syntactic distinctions in VPE. Finally, by treating the identity in question as syntactic, it becomes possible to formalize the analysis in a grammar that is amenable to standard parsing techniques. That is to say, it becomes much easier to incorporate such an analysis into a reasonable model of language use [Chomsky, 1965, 9]. To be "reasonable" a performance model must, of course, line up with available data. Arregui et al. [2006] observe a cline of acceptability values across VPE that match or mismatch in a variety of different ways. These authors suggest that mismatching VPE should be handled by special processor rules that *repair* ungrammatical structural descriptions in comprehension and *blend* grammatically-incompatible representations in production [Frazier, 2008]. By contrast, our analysis, which synthesizes a handful of ideas from recent work in syntax, makes it unnecessary for a performance model to rely on this special class of repair rules in VPE. We confirm the acceptability cline in three studies of our own, and account for it in terms of independently-motivated parser heuristics. These heuristics are not repair rules, but rather ways of prioritizing the search for syntactic structure [Kay, 1986, Pereira, 1985, Hale, 2009]. Such an organization of the overall theory, whereby mismatches are fundamentally grammatical and the acceptability cline is a consequence of greater or lesser searching work that the parser must do anyway, differentiates the present proposal from the foundational work by Frazier and colleagues on which we build.

The paper thus characterizes the role grammar might play in an adequate model of VPE acceptability. Section 2 briefly introduces a derivational ap-

proach to syntax that can be combined with standard notions of parsing. This section states our proposed parser heuristics and relates them to previous work in general linguistics. Section 3 reports an acceptability study that measures the acceptability cline in VPE; this study serves as a partial replication of [Arregui et al., 2006]. The proposed VPE analysis accounts for the observed pattern when applied in a parser that operates in accordance with the proposed heuristics. Section 4 follows up key predictions of our account by testing the same heuristics in new constructions. The results reported here are consistent with pre-hoc predictions derived from the overall theory. Section 5 draws some connections to functional pressures that might serve as explanations for the heuristics. Section 6 concludes with some reflections on other constructions implicated by the proposal.

2 The proposal

2.1 Grammar

This section reviews the VPE analysis presented in Kobele [2009] in preparation for its use deriving the observed patterns to be reported in sections 3 and 4. The analysis is stated in a version of Stabler's [1997] Minimalist Grammars (MGs) extended with hypothetical reasoning [Kobele, 2010]. Minimalist grammars provide a formal framework for reasoning about work in the minimalist program [Chomsky, 1995]. Hypothetical reasoning, in this technical sense, can be viewed as an implementation of ideas about A and A-bar movement proposed by Manzini and Roussou [2000] (see also Lasnik [1999]). These ideas have recently been reformulated by Takahashi and Hulsey [2009] in terms of 'late merger' [Lebeaux, 1988].

In MGs, there is a generalized transformation, MERGE, and a singulary transformation, MOVE. The results of these operations can be described using multi-dominance structures as shown schematically in Figure 1 [Kracht, 2001, Gärtner, 2002]. Nodes created by merge or move are represented with black dots in these and subsequent figures. Move can be viewed as the special case of merge where the second argument is a proper part of the first.[1]

[1]The objects we derive can also be described with trace chain structures or copy chain structures (see Kracht [2001] for a comparison of these three possibilities). The standard presentation in literature on minimalist grammars uses structures with traces (though Kobele [2006] employs copies). However, in the context of the widespread adoption of the

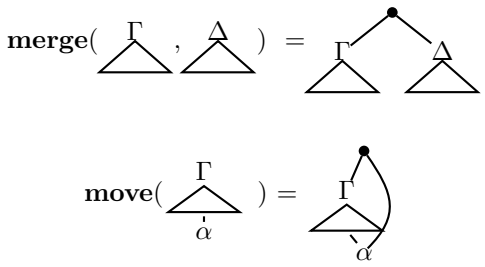


Figure 1: The operations merge and move

In accord with recent minimalist ideas, neither linear order nor category information is directly represented in these structures. There is no linguistic significance to the order in which sister substructures are displayed on the page. Rather, we assume that a surface ordering is determined by a linearization algorithm at the relevant interface [Kayne, 1994].[2] To see how these operations apply in clause-level syntax, consider the structures in Figure 2. These multi-dominance structures present a basic analysis of active and passive English sentences, to be further fleshed out in Figure 7 on page 9. The structure for the passive in Figure 2(b) crucially makes use of the move operation. As a result, the right-hand daughter of the topmost node is a subtree of the left-hand daughter. The two structures reflect the claim that passive is phrasal [Bach, 1980, Keenan, 1980].

(a) Active (b) Passive

Figure 2: Rough structure of active and passive sentences

copy theory of movement, the multiple domination representation seems most faithful to Chomsky's admonition against treating copies as distinct.

[2]One way of obtaining the string yield of the sentence is to first recover a tree by forgetting all but the lowest branches to nodes, and then to apply a finite copying top-down tree-to-string-transducer with regular look-ahead [Kobele et al., 2007] – this is the standard in presentations which leave move arcs implicit and thus in which expressions are structured as trees (as in Stabler [1997]).

Multi-dominance structures are but one of the notations in which MG-derived expressions can be written. The same expressions can equivalently be described by tuples of categorized strings. The intuitive idea is that a constituent still able to move remains in a separate component of the tuple. From this latter perspective, MGs can be seen to be equivalent [Michaelis, 2001] to other mildly context-sensitive formalisms [Joshi, 1985]. This tuple-oriented perspective makes it possible to extend parsing algorithms, such as Earley's algorithm, from context-free grammars to MGs [Harkema, 2001]. We sketch a small fragment covering sentences that can appear as conjuncts in VPE.

To begin, let us implement the deletion-under-identity idea by adding another grammatical operation: DELETE. This operation is depicted in Figure 3. The phonological consequences of the delete operation are clear. Syntac-

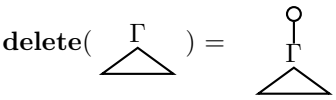


Figure 3: The operation delete

tically, delete extends a multi-dominance structure upwards by one node. In Figure 3 this node is labelled with a white dot.[3] Unlike merge and move, delete is licensed globally in a derivation. In the spirit of the CRD, we condition the application of delete upon the presence of a *syntactically identical* structure elsewhere in the finished derivation. The strictness of this requirement ultimately motivates a more flexible notion of merger, as discussed below in section 2.1.1 and in more detail in Kobele [2009].

To see how the deletion operation plays out in VPE, consider derivations for example (1) repeated here as (3). In this example, a verb phrase in the second conjunct has been deleted.

(3) Jill betrayed Abby, and Matt did ~~betray Abby~~, too

[3]Various alternative implementations of deletion are compatible with the fragment developed here. For example, Merchant [2001] proposes that deletion is triggered by an E feature at the PF interface. The delete operation can be viewed as a way of introducing a feature into the derivation without violating the no-tampering condition. Another option is to look at the operation of delete as abbreviating the merger of a silent head which itself contains an E feature, but which otherwise doesn't alter the featural constitution of its complement.

This application of delete is licensed under strict syntactic identity with the VP in the first conjunct, which is the node immediately dominating the leaves labeled verb and object. The relevant syntactic identity is schematically indicated by circling in Figure 4.

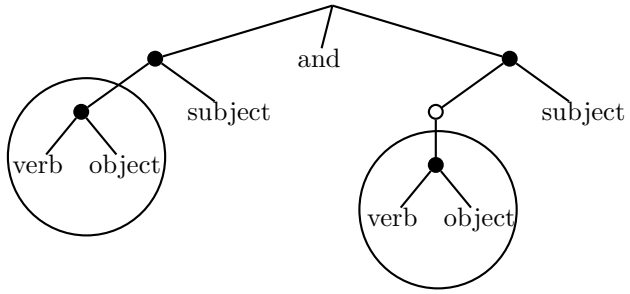


Figure 4: Rough structure of a VPE item like (3)

In Figure 4 both conjuncts fill their subject position by merger with nodes labelled subject. This is appropriate in a case where both conjoined sentences are in the active voice. Passives, on the other hand, are derived by promoting the direct object to surface subject — as suggested by Figure 2(b). A Passive-Active VPE mismatch item like (4) thus has the structure shown in Figure 5.

(4) Abby was betrayed, and Matt did ~~betray Abby~~, too.

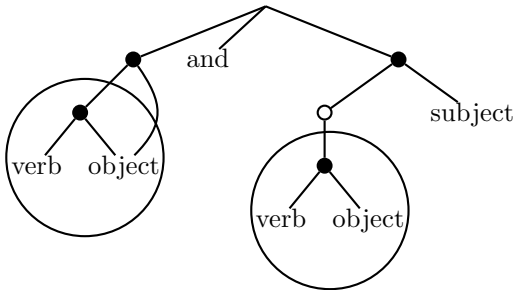


Figure 5: Rough structure of Passive-Active voice mismatch item like (4)

The strict syntactic identity requirement requires that passives and actives share a common multi-dominance substructure. This *identical* substructure

is circled in Figures 4 and 5. The next section lays out the details of a larger fragment founded on the assumption that a syntactically-identical match is truly present in VPE.

2.1.1 Analysis

Whereas the previous section developed the intuition behind a syntactic treatment of voice mismatches in VPE, the present section provides a concrete analysis of active and passive sentences in English. This analysis obviates the need for special repair rules in the processor à la Arregui et al. [2006]. It conservatively extends the analysis presented in Kobele [2006], synthesizing the ideas of passivization as case-checking, the VP-internal subject hypothesis [Koopman and Sportiche, 1991], and overt object case-checking [Koizumi, 1995, Runner, 1995].

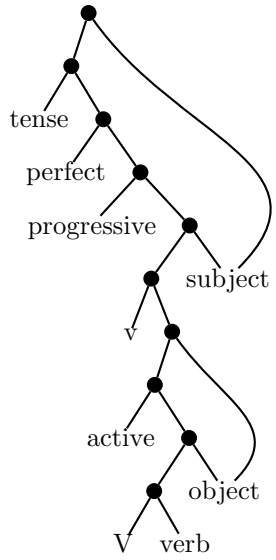
The substance of the analysis resides primarily in the syntactic properties of lexical items. Each lexeme includes a bundle of FEATURES. We differentiate between ATTRACTOR features written with an asterisk before their name, $*\mathtt{z}$ and ATTRACTEE features which lack this asterisk, $\mathtt{z}$. Operations such as merge or move only apply if the head of their first argument has an unchecked attractor feature and the head of the second argument has a corresponding unchecked attractee feature of the same name. Both of these features are said to be 'checked' in the resulting structure.

Following Chomsky [1995], the difference between complements and non-complements reduces to the order of merger: the first-merged item is a complement, all others are non-complements. We assume with Stabler [1997] that features within a bundle are totally-ordered (see also Müller [2008]). For example, the feature bundle for the proper name John might be associated with a sequence of two attractee features as in Figure 6.

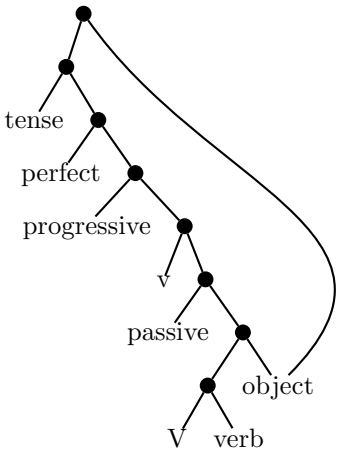
$$\langle \mathsf{John}, \mathtt{d}\ \mathtt{k} \rangle$$

Figure 6: Lexical entry for a proper name

The first feature $\mathtt{d}$ means that this lexeme can occupy DP positions in the syntax. It is selected by expressions with the corresponding attractor feature $*\mathtt{d}$. The second feature, $\mathtt{k}$, encodes a requirement for case. This requirement could be satisfied by movement, triggered by $*\mathtt{k}$, to the specifier of some other head.



(a) Refined structure for actives (see Figure 2(a))



(b) Refined structure for passives (see Figure 2(b))

Figure 7: More detailed structures corresponding to Figure 2

Figure 7 sums up our proposal about the structure of active and passive sentences. These more detailed structures illustrate the role of a voice head that differentiates active (lexical item: $\langle \epsilon, *\texttt{V}\ *\texttt{k}\ \texttt{act} \rangle$) from passive (lexical item: $\langle \texttt{-en}, *\texttt{V}\ \texttt{pass} \rangle$). Whereas the passive voice head contributes an overt morpheme "-en", the active voice head is phonologically null. This is notated with the empty-string symbol ϵ. These two heads also differ in their combinatory potential. Because it has a $*\texttt{k}$ attractor feature, the active voice head can check the case of a direct object. Lack of such a feature means that the corresponding passive voice head cannot. Both heads combine with a verb phrase to form a configurationally-higher verbal category known as 'little v.'[4]

[4]The passive voice head that we employ plays the same role as AgrO. The key property is that the needed head be situated between big V and little v [Koizumi, 1995, Runner, 1995]. Runner [2006] argues against the 'split VP' analysis in the context of an argument against treating raising to object as overt movement. The present split VP analysis, presented in more detail in Kobele [2006], does not suffer from the conceptual problems discussed there.

In our fragment, little v is broken up into two types, defective and non-defective. Defective little v (lexical item:$\langle$`be`, $*$`pass v`$\rangle$) is unable to license an external argument whereas non-defective little v can do so in virtue of an extra attractee feature (lexical item:$\langle\epsilon$, $*$`act` $*$`d v`$\rangle$) [Chomsky, 2000]. The heads above vP belong to the T-domain. They introduce tense and aspect morphemes. We adopt an affix-hopping analysis [Lasnik, 1995]. We treat finite T as assigning nominative case; the $*$`k` feature on the tense head triggers movement of the closest DP with unchecked case features (lexical item: $\langle$*tense*, $*$`perf` $*$`k t`$\rangle$).

These lexical assumptions, in combination with the hypothesis that deletion is licensed by strict syntactic identity, entail that the voice head is not included as part of the matching structure in Passive-Active VPE items such as (4). The voice heads cannot be involved because the conjuncts do not match in voice. Only the smaller subtree containing the big V head, the verb and the object in Figures 7(a) and 7(b) can be deleted. By contrast, an Active-Active item such as (3) can delete in the second conjunct above the voice head. Section 2.3 discusses this uncertainty about the configurational height of the ellipsis site from the perspective of parsing preferences.

In Passive-Passive examples such as (5), the only matching substructure consists of big V and its selected verb.[5] Descriptively speaking, objects in these cases seem to count for identity only in their 'surface' positions, whereas for the other cases of ellipsis looked at thus far, they count as being in their 'deep' positions.[6] This situation is analogous to the behavior of A-moved expressions with respect to reconstruction [Fox, 2000]. In reconstruction,

[5]Although one might treat example (5) as involving, not VP deletion, but V deletion, this approach does not scale up to deal with superficially similar phenomena, which are clearly to be classified with VP ellipsis:

i. Abby seems to have been betrayed, and Max does ~~seem to have been betrayed~~ too.

[6]These very constructions motivated Sag [1976] to weaken the identity condition from *trees* to *contexts* (which he implemented by checking identity at LF, which he viewed as λ-terms over trees). Fox [2002], working in the present framework, adopts a similar idea, 'trace-conversion', which (functionally) allows multiply dominated nodes to only 'count' as being in a higher position. Late-merger is an implementation of this common idea, that of conditioning identity (and deletion) on contexts, instead of trees (or graphs, in our case). It differs from these previous ideas both in connecting the behaviour of arguments in ellipsis to the superficially distinct phenomenon of reconstruction, and in that it does not force us to adopt multiple levels of representation – there is just one structure, but a richer notion of derivation.

A-moved elements can count, for the purposes of determining condition C effects, as being either in their surface positions or in their underlying positions. By contrast, A-bar moved expressions always count as being in their underlying positions. We follow Kobele [2009] and Takahashi and Hulsey [2009] in supposing that this analogous behavior reflects a common mechanism, late merger [Lebeaux, 1988]. This mechanism has proved useful in accounting for reconstruction asymmetries [Lasnik, 1999, Kobele, to appear, Takahashi and Hulsey, 2009].

In our analysis of VPE, we apply a version of this same mechanism [Manzini and Roussou, 2000]. We adopt the specific formalization from Kobele [2010]. Whereas in Stabler [1997], feature bundles were taken to simply be parts of expressions, forcing move and merge to apply to check them, Kobele [2010] treats feature bundles as objects in their own right. These feature bundles can be temporarily dissociated from their expressions and manipulated by derivational operations. The innovation is that attractor features may be satisfied by feature bundles as well as expressions. Such a feature bundle then plays the same role that an expression would play. However, in order to be interpretable at the PF and LF interfaces, a feature bundle must be reunited with its expression.[7] This is shown in Figure 8 where the feature bundle is written with the symbol γ and its reassociation with its expression during late merger is indicated by a dotted line. The flexibility provided by late merger allows for a second derivation of passives, shown in Figure 9(b) (cf. Figure 7(b)). In Figure 9, the attractor feature $*\mathtt{d}$ of the verb is checked by the merger of the feature bundle $\langle \mathtt{d}\ \mathtt{k} \rangle$. This feature bundle belongs to the object, which is merged directly in its surface position.

With late merger in mind, consider sentence (5), below, in which a passive VP serves as the antecedent for a passive ellipsis site.

(5) Abby was betrayed, and Matt was ~~betrayed~~, too.

Where the second conjunct is in the passive voice, the deep direct objects

[7] This version of the late merger mechanism differs from that advanced by Takahashi and Hulsey [2009] in that here we assume that the entire DP is late merged, whereas they assume that just the NP complement to D is late merged. Our version is able to account for the fact that non-identical Ds are able to head the NPs in passive-passive ellipsis:

ii. John seems to have been kissed, and every ninja does ~~seem to have been kissed~~, too.

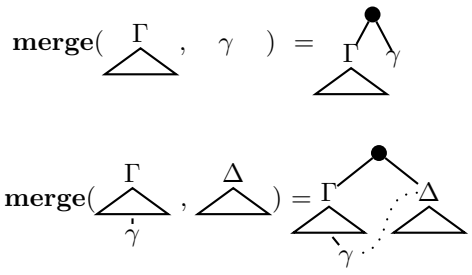


Figure 8: Late merger

in VPE would be non-identical if they were introduced into the derivation at the same point as in earlier examples. However, by late merging both objects, the VPs become identical. Both deep object positions contain an identical feature bundle, ⟨d k⟩. In the Active-Passive case, the active object is late merged in its case position, *above* the voice head, as depicted in Figure 9(a).

(6) Jill betrayed Abby, and Matt was ~~betrayed~~, too.

The lexical items needed to derive (5) and (6) in this manner are given in Figure 10.

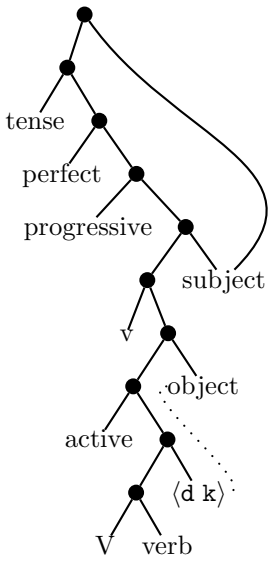
2.1.2 Category

The fragment developed so far extends to VPE examples that mismatch on grammatical category. We consider nominalization and adjectivization as in example (7). We adopt a lexical decomposition analysis that is broadly consistent with work in the principles and parameters tradition (e.g. Pesetsky [1995]). According to this sort of analysis, a nominal like "admission of guilt" is derived from the VP "admit guilt" by the addition of a nominalizing head (lexical item: ⟨nom, ∗V N p⟩).[8] Derivations for these two classes of examples

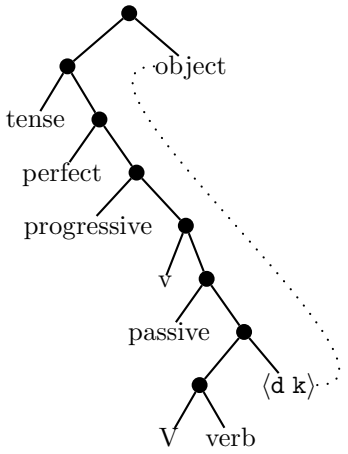
[8]We have in mind a traditional *of-insertion* analysis, whereby the NP which would otherwise lack case is 'rescued' by the last-resort insertion of the dummy preposition *of*. As this story is not stateable as such in the present system, we have opted for the following work-around involving remnant movement in the spirit of Kayne [1994]:

i. [nom [V Obj]] (merge nominalizing head with the VP, which introduces a feature 'p')

ii. [ϵ [nom [V Obj]]] (merge another empty head with the VP, which checks the



(a) Late merged object in active (cf. Figure 7(a))



(b) Late merged object in passive (cf. Figure 7(b))

Figure 9: Alternative derivations involving late-merged objects

⟨Abby, d k⟩	⟨Matt, d k⟩
⟨betray, V⟩	
⟨ϵ, ∗V ∗d V⟩	
⟨ϵ, ∗V ∗k act⟩	⟨-en, ∗V pass⟩
⟨ϵ, ∗act ∗d v⟩	⟨be, ∗pass v⟩
⟨ϵ, ∗v prog⟩	⟨-ing, ∗v y⟩
	⟨be, ∗y prog⟩
⟨ϵ, ∗prog perf⟩	⟨-en, ∗prog z⟩
	⟨have, ∗z perf⟩
⟨-ed, ∗perf ∗k t⟩	

Figure 10: Lexical items (I)

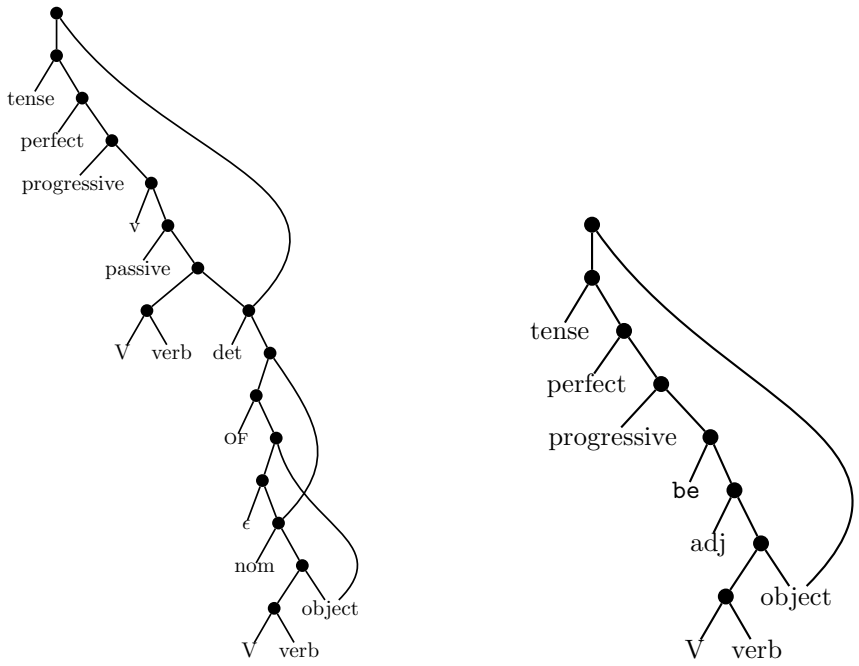


Figure 11: Structures underlying examples (7a) and (7b)

are given in Figure 11.

(7) a. An admission of guilt was needed, but the suspect wouldn't ~~admit guilt~~

b. The mistake was excusable, and the director did ~~excuse the mistake~~

In both cases a VP-like constituent, consisting of a verb and its deep object, is available to serve as an antecedent for 'regular' VPs as in Figure 7.

We assume the existence of morphological rules that spell out admit+nom as "admission" and excuse+adj as "excusable". No additional rule types are required to handle these sorts of examples beyond the the merge, move and

---

object's case)

iii. [Obj [ϵ [nom [V t]]]] (object moves for case)

iv. [of [Obj [ϵ [nom [V t]]]]] (merge of, triggers remnant movement of the VP)

v. [[nom [V t]] [of [Obj [ϵ t]]]] (remnant VP moves to check feature p)

$\langle \mathsf{suspect}, \mathtt{n} \rangle$	$\langle \mathsf{guilt}, \mathtt{d\ k} \rangle$
$\langle \mathsf{director}, \mathtt{n} \rangle$	$\langle \mathsf{the}, *\mathtt{n\ d\ k} \rangle$
$\langle \mathsf{mistake}, \mathtt{n} \rangle$	
$\langle \mathsf{admit}, \mathtt{V} \rangle$	$\langle \mathsf{excuse}, \mathtt{V} \rangle$
$\langle \mathtt{nom}, *\mathtt{V\ n\ p} \rangle$	$\langle \mathtt{adj}, *\mathtt{V\ a} \rangle$
$\langle \epsilon, *\mathtt{n}\ *\mathtt{k\ N} \rangle$	$\langle \mathsf{be}, *\mathtt{a\ v} \rangle$
$\langle \mathsf{of}, *\mathtt{N}\ *\mathtt{p\ n} \rangle$	

Figure 12: Lexical items (II)

delete operations introduced in section 2.1. Figure 12 shows the additional lexical items needed to derive examples like these.

2.2 Conceptual Burden

The conceptual burden imposed by the theory presented above is comparatively light. The key assumptions are that the relevant notion of identity in VPE is exact syntactic identity, and that passive clauses are derivationally related to active ones. In the context of minimalism, these assumptions lead inexorably to a grammar that admits mismatched VPE. This motivates a reconsideration of the boundary between acceptability and grammaticality [Joshi, 2004, §4.3]. An alternative positioning of this boundary classifies all mismatched VPE as ungrammatical. This move raises the opposite question: why are some cases of mismatched VPE so acceptable? Any repair mechanism brings up questions of scope and generality. When does this mechanism apply? Does it duplicate mechanisms already present in the grammar? The next section develops an account for the acceptability cline in VPE that sidesteps these difficult questions.

2.3 Processor

We assume with Berwick and Weinberg [1984] that operations of the human sentence processing mechanism can be put into correspondence (perhaps many-to-one) with grammar rules. At the same time, any realistic grammar that can be deployed in parsing will be highly ambiguous. For instance, the context-free grammar implicit in the Penn Treebank [Marcus et al., 1993] leads to an average of 1.2 million individual parser state updates when sentences of less than 40 words are analyzed using the methods of Charniak et al.

[1998]. We take this as an indication that exhaustive search is implausible for any system, human or machine. Instead, some additional considerations, or heuristics, must guide the search for syntactic structure just as in other domains of cognition [Newell and Simon, 1972, Gigerenzer et al., 1999]. These heuristics are not part of grammar, but are rather claims about how grammar is used. We propose that the acceptability cline in VPE falls out from the ordering on parser states induced by the heuristics. It falls out because, as Schütze [1996, §5.3.3] emphasizes, parsability is a task-related factor that influences acceptability. We contend that highly acceptable items — those consistent with the proposed parsing heuristics — have syntactic analyses that are found comparatively early in the sentence comprehension mechanism's self-terminating search process.[9] By contrast the less acceptable items require more states to be explored because they are inconsistent with "hints" that usually help. The two heuristics are:

MaxElide VP ellipsis preferentially targets configurationally higher rather than lower nodes.

Canonical Realization Surface subjects preferentially are underlying subjects as well.

These parsing preferences reflect certain leading ideas in the literature. MaxElide seems to have first crystallized in Merchant [2008b] (first circulated in 2001), and has been expanded upon in Takahashi and Fox [2005]. Our specific interpretation of it entails that deletion above the voice head is to be expected, even though deletion at lower points is also possible. Canonical Realization synthesizes various strands of thinking, the oldest being Bever's [1970] perceptual strategies [Townsend and Bever, 2001, §2.4.2]. Kaplan [1972] formalizes perceptual strategies in a way that penalizes relabeling of, say, a subject as some other grammatical function such as direct object in light of evidence from later words. This is exactly the spirit of our proposed heuristic, which may itself follow from the markedness of certain word-orders. Canonical Realization also echos Stevenson and Smolensky's [2006, 315] application of case hierarchies in the parser.

[9]Section 5.1 details how chart parsing formalizes (one notion of) parallel syntactic analysis. This sort of parallel processing supports a graded notion of search error that is qualitatively different from the classic notion of garden pathing that assumes serial processing [Frazier, 1979]. Hale [2009] takes up these issues in greater depth.

little v > voice > big V

Figure 13: MaxElide elision hierarchy

MaxElide predicts that elliptical constructions involving configurationally-higher verbal projections should be judged more acceptable than those involving configurationally-lower, smaller ones. The hierarchy that MaxElide sets up, in combination with the grammar fragment presented in section 2.1 is shown below in Figure 13.

Canonical Realization predicts a word order effect. The word order it favors is one in which surface subjects, in specifier of TP, are also underlying subjects, introduced in specifier of vP. These constructions should be judged more acceptable than those whose surface subjects originate in other positions. In more theory-neutral terms, Canonical Realization favors surface structure that assign the Agent thematic role to a pre-predicate position.

3 Experiment 1: acceptability of voice and category mismatches

3.1 Experimental design and Materials

To measure the acceptability of mismatched VPE items such as (2), we conducted an acceptability study using magnitude estimation [Bard et al., 1996]. This study examines both the grammatical voice of the (mis)matching conjuncts, as well as the grammatical category, either nominal or adjectival, of the antecedent. Table 1(a) shows how the design embeds these two ways of mismatching into both the experimental ("Ellipsis") and control conditions ("No ellipsis"). In the latter conditions, participants did see the crossed-out material; the proper name in the second clause was sometimes replaced by the corresponding pronoun, to minimize 'repeated name' effects [Gordon et al., 1993]. The No Ellipsis control conditions served as a baseline for their corresponding Ellipsis conditions, allowing us to more accurately measure the change in acceptability due to ellipsis.

(8) Voice match

a. Active-Active: Jill betrayed Abby, and Matt did ~~betray Abby~~,

too.

b. Passive-Passive: Abby was betrayed by Jill, and Matt was ~~betrayed by Jill~~, too.

(9) Voice mismatch

a. Active-Passive: Jill betrayed Abby, and Matt was ~~betrayed by Jill~~, too.

b. Passive-Active: Abby was betrayed by Jill, and Matt did ~~betray Abby~~, too.

(10) Category mismatch

a. Noun-VP: The criticism of Roy was harsh, but Kate didn't ~~criticize Roy~~

b. Adjective-VP: The report was critical of Roy, but Kate didn't ~~criticize Roy~~

Within each half of the materials (Voice or Category mismatch), items were counterbalanced such that each item appeared once per list, and appeared an equal number of times in each experimental condition across the entire experiment. There was a total of 48 experimental items, with equal numbers of Voice and Category mismatches, and corresponding controls. Each participant was assigned to one of four lists, and saw a different randomization of list items.

3.2 Method

3.2.1 Procedure

All experiments used the magnitude estimation paradigm. As in typical acceptability judgment studies, participants are asked to judge the acceptability of a series of sentences.[10] Unlike in a fixed-scale rating study, a participant assigns a score to each sentence relative to a standard score that she has herself chosen at the beginning of the experimental session. Thus each participant rates the same standard sentence — the MODULUS — at the beginning of the session, establishing her own anchor for judging the sentences in the experiment. Participants are instructed to make proportional estimates relative to the modulus value. For example, if the modulus value chosen was 100, a sentence that sounds twice as acceptable should be scored 200, while a sentence

[10]We presented stimuli to human participants using PsyScope X [Bonatti, 2008].

Table 1: Experiment 1

(a) Design

		Match	Mismatch
No ellipsis	Voice	(8a) (8b)	(9a) (9b)
	Category	like (8a)	(10a) (10b)
Ellipsis	Voice	(8a) (8b)	(9a) (9b)
	Category	like (8a)	(10a) (10b)

(b) Observed acceptability cline

condition	mean log acceptability	example
Active-Active	0.235	(8a)
Passive-Passive	-0.285	(8b)
Passive-Active	-0.616	(9b)
Noun-VP	-0.690	(10a)
Active-Passive	-0.697	(9a)
Adjective-VP	-0.981	(10b)

that sounds half as acceptable should be scored 50. Participants practiced this method of estimation before beginning the experiment, first using line lengths, then with sentences, as in Bard et al. [1996].

Participants read whole sentences on a computer screen. Experiment 1 used the modulus: "The kids were amused by the cartoon, but their parents weren't." On each trial, a sentence appeared on the screen along with the modulus sentence. Participants entered their score for the sentence in a text box.

3.2.2 Data analysis

All analyses were performed on log-transformed normalized scores, calculated by dividing a subject's raw scores by the modulus value assigned by that subject. In the transformed scores, therefore, a positive value means a particular item was rated higher than the modulus value, while a negative value means an item was rated lower than the modulus. These values are meaningless as absolute acceptability values. But because all sentences are estimated relative to the same standard, we can meaningfully compare the direction and magnitude of the differences among conditions from the same experiment.

The observed acceptability estimates can be affected by the other items in a set of stimuli. A person judging a set of disproportionately complex items may score a single simple item as sounding better, compared to how they might rate the same sentence when it appears among other simple sentences. Therefore, in interpreting our results, we draw conclusions from the relative magnitudes of differences only within an experiment. Across experiments, we compare only the overall patterns of results.

3.2.3 Participants

Twenty University of Rochester undergraduates, who were native speakers of English, were paid $7.50 to participate in the experiment.

Table 2: Experiment 1, Voice mismatches: Estimates of fixed effects

	Coefficient	*Standard error*	t	p
Intercept	-0.26	0.038	-6.90	< 0.001
Ellipsis	-0.12	0.020	-5.95	< 0.001
Mismatch	-0.13	0.020	-6.44	< 0.001
Ellipsis:Mismatch	-0.12	0.021	-5.72	< 0.001

3.3 Results

Log scores were fit to a linear mixed-effects regression model [Gelman and Hill, 2007, Baayen, 2008] with Subject, Item, and Trial as random effects.[11] The data was divided into two subsets, representing the Voice mismatches and the Category mismatches, and a separate model was fit for each subset.[12] These regression models predict a numerical acceptability level on the basis of the following independent variables: whether the sentence contains ellipsis (Ellipsis or No Ellipsis), and whether the structure of the second conjunct matches that of the first conjunct (Match or Mismatch). The two subsets of the data and within the Mismatch conditions, the type of mismatch i.e. Voice or Category. The Ellipsis by Match interaction was also included in the model. Tables 2 and 3 show the estimates of the coefficients for these fixed effects. The absolute values of the coefficients are plotted in Figure 14. The extent to which the absolute values differ from zero in these graphs provide a visual representation of effect size.

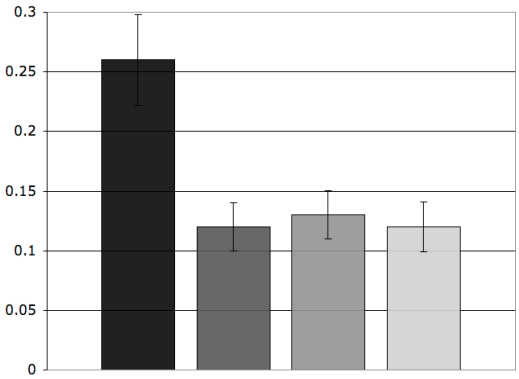
The direction and significance of the fixed effects in the Voice and Category models were identical. In both subsets of the data, there was a main effect of Mismatch ($p < 0.001$), such that sentences with structurally mismatched conjuncts were judged less acceptable than their matched counterparts. There was also a main effect of Ellipsis ($p < 0.001$), such that sentences containing ellipsis were judged less acceptable than their counterparts without ellipsis. There was, however, a significant Mismatch by Ellipsis

[11] We fitted the mixed-effects models using the `lme4` library of the 'R' system for statistical computing [Bates and Sarkar, 2007]. All predictor variables were contrast coded.

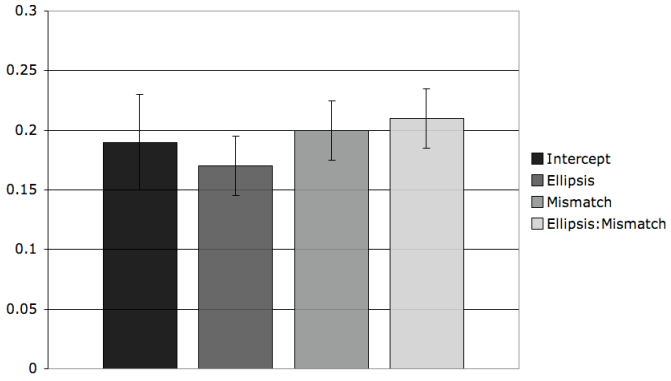
[12] In all analyses presented in this paper, extreme outliers that fell more than four standard deviations from the mean were excluded from the analysis, representing less than 1% of the data.

Table 3: Experiment 1, Category mismatches: Estimates of fixed effects

	Coefficient	*Standard error*	*t*	*p*
Intercept	-0.19	0.040	-4.74	< 0.001
Ellipsis	-0.17	0.025	-6.74	< 0.001
Mismatch	-0.20	0.025	-7.99	< 0.001
Ellipsis:Mismatch	-0.21	0.025	-8.59	< 0.001

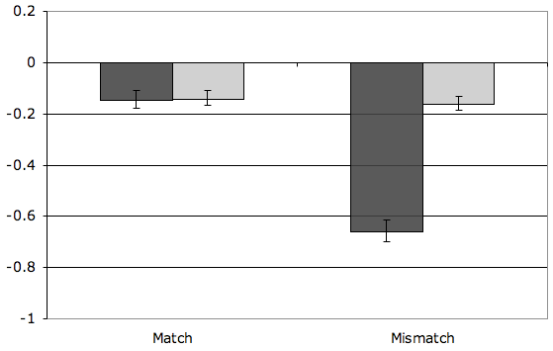


(a) Voice mismatch

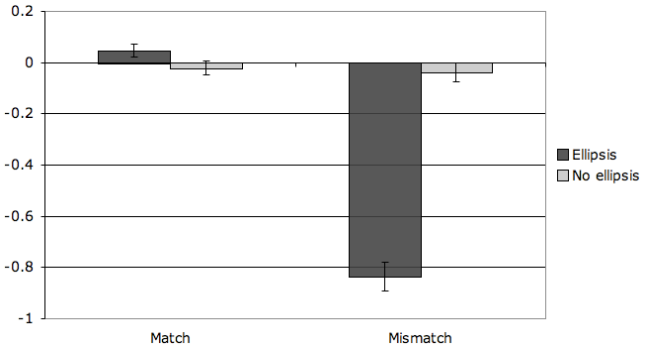


(b) Category mismatch

Figure 14: Absolute values of Experiment 1 model coefficients (error bars represent standard error).



(a) Voice mismatch



(b) Category mismatch

Figure 15: Experiment 1 condition means (error bars represent standard error).

interaction ($p < 0.001$): mismatch had a greater negative impact on acceptability judgments when there was ellipsis in the second conjunct, confirming that the condition on matching conjuncts is specific to cases of ellipsis. Condition means are plotted in Figure 15. The conditions represented in these graphs collapse over sub-categories of Voice Match ((8a) and (8b)), Voice Mismatch ((9a) and (9b)), and Category Mismatch ((10b) and (10a)).

Within Category mismatches, adjectival antecedents (10b) were rated worse than nominal antecedents (10a) ($t = 3.49, p < .001$). Table 1(b) summarizes the acceptability cline observed in this study.

3.4 Discussion

The results of experiment 1 are consistent with a comprehension mechanism that searches for syntactic analyses and takes more steps when these analyses fail to fit the expectations set up by the heuristics MaxElide and Canonical Realization. Apart from the superior acceptability of active (AA) over passive (PP) the two key results are

{ AA,PP } > { PA,AP } Mismatch imposes an acceptability penalty in VPE

NV>AdjV Within category mismatches, Noun-VP is the more acceptable form of VPE

We discuss each of the results separately in light of the general theory presented above in section 2.

3.4.1 Voice mismatches

Because their conjuncts involve different voice heads, the Active-Passive and Passive-Active stimuli must delete at big V in violation of MaxElide. By contrast, matching controls can delete at the configurationally higher little v node. This derives the observed relationship {AA,PP} > {PA,AP}. Canonical Realization derives AA>PP exactly as in Kaplan [1972].

The intuition behind MaxElide, that there is a preference for large antecedents, appears elsewhere in the literature [Frazier and Clifton, 2005]. If there is such a preference, it is not necessary that violating it result in unacceptability; a violation of MaxElidecould, for instance, result in greater processing time, but not greater unacceptability. In other words, processing difficulty is not necessarily the same as degraded acceptability. It is part of our claim that variations in processing difficulty is the source of the graded acceptability we observe.[13]

3.4.2 Category mismatches

The lexical decomposition analysis presented in 2.1.2 exposes a distributional difference between the Adjective-VP stimuli and the Noun-VP stimuli. Figure 16 shows that, in the Noun-VP stimuli, the argument of the predicate

[13]In fact, the experimental results in Frazier and Clifton [2005] and Arregui et al. [2006] also show a correspondence between processing time, as measured by reading times, and acceptability judgments.

canonical?	
yes	Pete's criticism of the book was harsh, but Jill didn't
yes	Emma's expectation of an easy win was obvious, but her teammates didn't
yes	Pam's approval of the plan was crucial, but she wouldn't
yes	Amy's distrust of the administration was clear, but Tom didn't
yes	Andy's intention to run for class president was apparent, but Lisa didn't
yes	Everyone's contribution to the effort was important, but Paul didn't
yes	The spy's exposure of the plot was risky, but he did
yes	The retrieval of the supplies from the warehouse was important, but the assistant forgot to
yes	An admission of guilt was needed, but the suspect wouldn't
yes	Evaluation of the earthquake survivors was recommended, but the doctors didn't
yes	Suggestions of changes for the draft would've been helpful, but no one did
yes	Assistance with the deliveries would've been useful, but few people did

(a) majority canonical Noun-VP stimuli

canonical?	
yes	The report was critical of Matt, but Frank didn't
yes	The letters were supportive of Kate's application, but the committee didn't
	The article was hardly praiseworthy, but everyone did
	The window was open, but Frank hadn't
	The cookies were burnt, but Lucy didn't
	The mistake was hardly excusable, but the director did
	The dishes were clean, but the girls hadn't
	The lecture was barely understandable, but Ben did
	The food was barely edible, but Meghan did
	Lauren wasn't very trustworthy, but Mark did
	The story was well-known, but Kelly didn't
	The librarian was well-liked on campus, but Jane didn't like her

(b) minority canonical Adjective-VP stimuli

Figure 16: Canonical ordering in category mismatch stimuli

{ criticise,expect,approve,distrust,intend,contribute . . . } is often sequestered in a prepositional phrase. By contrast, the majority of the arguments in 16(b) are in a pre-copular position. Only in the subset of transitive-"of" stimuli does the element in the subject-like position { report,letter } express a predication over an argument that comes linearly later in a prepositional phrase.

Canonical Realization imposes an expectation for precisely this word order. In a typical Noun-VP item such as

(11) An admission of guilt was needed, but the suspect wouldn't ~~admit guilt~~ (7a)

the argument of the root (e.g. $\sqrt{}$admit) appears linearly to its right in conformity with Canonical Realization. This is just the ordering expected in a Bever-type [1970] NVN template such as suspect-admit-guilt. By contrast, in a non-canonically-realized example,

(12) The mistake was hardly excusable, but the director did ~~excuse the mistake~~ (7b)

the argument (e.g. "mistake") appears linearly before the root (e.g. $\sqrt{}$excuse). The vast majority of the Adjective-VP items in Experiment 1 attested the

Table 4: The specific syntactic steps that determine compliance with Canonical Realization

	adjectival
Canonical	[critical of Matt] + [report] $\rightsquigarrow$ AP
non-Canonical	[-able] + [excuse the mistake] $\rightsquigarrow$ AP

	nominal
Canonical	[an] + [admission of guilt] $\rightsquigarrow$ DP
non-Canonical	movement $\rightsquigarrow$ [$_{\text{DP}}$ guilt's admission]

non-canonical order. Generating one order or the other depends on specific syntactic steps summarized in Table 4. In this table, the $\rightsquigarrow$ symbol indicates rewriting from child sub-derivations on the left to derived categories on the right. Each rewriting corresponds to a particular branch of the structures in Figure 11. To implement Canonical Realization, these branches would receive contrasting weights in a stochastic branching-process model of MG derivations [Hale, 2006]. The category mismatch conditions in Experiment 1 sampled almost exclusively from the gray-colored cells, and obtained a pattern that reflects exactly the expectations imposed by Canonical Realization. This pattern implicates Canonical Realization as a possible explanation for the observed pattern. However it also motivates inquiry into the other cells.

Experiment 1 did not examine stimuli of the form "guilt's admission...". These items would occupy the bottom-right cell of Table 4. However the effect of Canonical Realization can be seen in corpus attestation frequencies. In the Brown corpus [Kučera and Francis, 1967, Marcus et al., 1993] for instance, noun phrases starting with a simple determiner are attested about four times more frequently than an alternative class of possessive structures consisting of those that either begin with a pronoun (such as "hers", "his", "my", "yours") or contain the Saxon genitive. Table 5 shows these counts. These Canonical Realization asymmetries would derive the observed processing difficulty differences in a frequency-sensitive comprehension model [Jurafsky, 1996, Crocker and Brants, 2000, Hale, 2006, 2009].

Since Canonical Realization seems to be consistent both with the observed acceptability pattern in Category mismatch stimuli and with corpus attestation rates in related constructions, we sought to confirm its role by

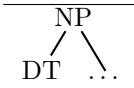
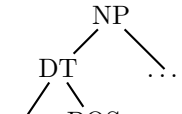
schema	attestations in Brown
	41,032
 $\cup$ PRP\$	2004 + 8802 = 10,806

Table 5: Attestation rates of genitive vs simple NPs in the Brown corpus

measuring its acceptability in these new constructions. The prediction is that there is an additional penalty for violating Canonical Realization in VPE.

4 Experiments 2 and 3: adjectivizations and nominalizations

To confirm the role of Canonical Realization in VPE acceptability, we conducted two further acceptability studies. As shown in Table 6, Experiments 2 and 3 crossed Ellipsis, Match, and Canonicality. In Experiment 2, category mismatches resulted from having an adjectival antecedent in the first conjunct. In Experiment 3, mismatching antecedents were nominal. Across both studies, the antecedents varied in Canonicality: half of the items in each experiment represented canonical argument realizations (*agent—verb—theme*) (see the (a) examples in (13)-(16)), and half represented non-canonical order (see the (b) examples in (13)-(16)). Both experiments contained a total of 32 experimental items.

(13) mismatching Adjectival antecedents

  a. ofTheme: The parents were critical of the uniforms, but the cheerleaders didn't ~~criticize the uniforms~~

  b. -able: The boy's exhaustion was understandable, and the coach did ~~understand the boy's exhaustion~~

(14) matching VP antecedents

(a) Experiment 2 – Adjectival

		Match	Mismatch
No ellipsis	canonical	As below but with overt second conjunct	
	¬canonical		
Ellipsis	canonical	(14a)	(13a)
	¬canonical	(14b)	(13b)

(b) Experiment 3 – Nominal

		Match	Mismatch
No ellipsis	canonical	As below but with overt second conjunct	
	¬canonical		
Ellipsis	canonical	(16a)	(15a)
	¬canonical	(16b)	(15b)

Table 6: Experiments 2 and 3

a. Active: Paul deciphered the riddle, but Kevin didn't ~~decipher the riddle~~
b. Passive: The wedding party was accommodated by the hotel, but the guests' pets weren't ~~accommodated by the hotel~~

(15) mismatching Nominal antecedents

a. ofTheme: The landing of the plane was unplanned, but the pilot did ~~land the plane~~
b. ThemePoss: The plane's landing was unplanned, but the pilot did ~~land the plane~~

(16) matching VP antecedents

a. Active: The experienced pilot landed his plane during the storm, but the novice didn't.
b. Passive: The plane was landed by an experienced pilot during the storm, but a nearby jet wasn't.

Experimental trials were intermixed with an equal number of filler sentences, which varied in length, complexity, and acceptability. A different modulus sentence was also used ("The kids were amused by the cartoon, but their parents didn't"), to better approximate a middle point for the distribution

of acceptability represented in the items.[14]

4.1 Method

The procedure for Experiments 2 and 3 was identical to that of Experiment 1.

Twenty-six University of Rochester undergraduates, who were native speakers of English and had not participated in Experiment 1, participated in Experiment 2; a separate twenty-six participated in Experiment 3.

4.2 Results

As before, transformed scores were fit to a linear mixed-effects model. The models included the following fixed effects: whether the sentence contained ellipsis (Ellipsis or No Ellipsis); whether the structure of the second clause matched that of the first clause (Match or Mismatch); and whether the order of arguments in the second clause was in canonical order (Canonical or Non-canonical). All two-way and three-way interactions among these fixed effects were also included in the model.

4.2.1 Experiment 2: Adjectivizations

There was a main effect of Canonical order ($p < 0.05$): non-canonical order in the antecedent corresponded to lower acceptability estimates than canonical order, confirming our prediction. As in the previous experiment, there was a main effect of Match, such that sentences with structurally mismatching antecedents were judged less acceptable than their matched counterparts ($p < 0.001$). Also as in Experiment 1, there was a main effect of Ellipsis ($p < 0.001$), such that sentences with ellipsis were judged less acceptable than their counterparts without ellipsis, and the Ellipsis by Match interaction was significant ($p < 0.001$): mismatch corresponded to lower acceptability estimates than match for sentences with ellipsis, but not for sentences without ellipsis.

[14] The modulus sentence used for Experiments 2 and 3 is an instance of Voice mismatched ellipsis (*... but their parents ~~didn't amuse the kids~~*). This sentence type is expected to be less acceptable than some of the better sentences, and more acceptable than some of the worse ones. Using a modulus that is roughly in the middle of the acceptability distribution of the experimental items is necessary in order to avoid ceiling or floor effects, which could introduce systematic warping in the data, with contrasts on one end of the acceptability scale being amplified, while contrasts on the other end being reduced.

Table 7: Experiment 2: Estimates of fixed effects

	Coefficient	*Standard error*	t	p
Intercept	0.81	0.30	2.68	< 0.05
Ellipsis	-0.098	0.016	-5.96	< 0.001
Canonical	-0.035	0.016	-2.16	< 0.05
Mismatch	-0.10	0.016	-6.14	< 0.001
Ellipsis:Canonical	-0.016	0.07	-0.97	*n.s.*
Ellipsis:Mismatch	-0.12	0.016	-7.60	< 0.001
Canonical:Mismatch	0.025	0.016	1.53	*n.s.*
Ellipsis:Canonical:Mismatch	-0.016	0.016	-0.80	*n.s.*

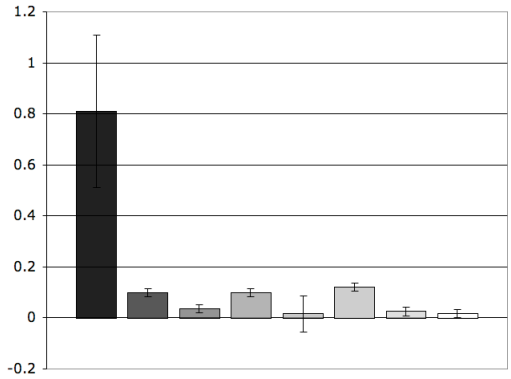
The Ellipsis by Canonical order, Canonical order by Mismatch, and Ellipsis by Canonical order by Mismatch interactions were not significant (all $p > 0.1$).

Table 7 gives the estimates of the coefficients in the Adjectivization regression. The absolute values of the coefficients are plotted in Figure 17, and condition means are plotted in Figure 18.

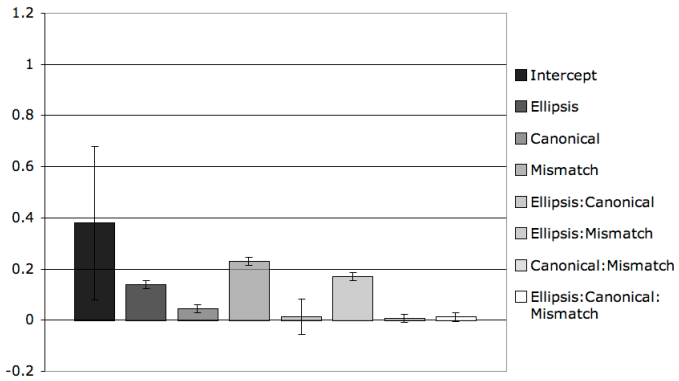
4.2.2 Experiment 3: Nominalizations

The Nominalization regression revealed the same pattern of fixed effects as the Adjectival regression. There were main effects of both Canonical order ($p < 0.01$), Mismatch ($p < 0.001$), and Ellipsis ($p < 0.001$). There was a significant Ellipsis by Match interaction ($p < 0.001$), such that mismatched sentences were degraded relative to their matched counterparts for sentences with ellipsis, but not for sentences without ellipsis. Again as in the Adjectival model, the Ellipsis by Canonical order, Canonical order by Mismatch, and Ellipsis by Canonical order by Mismatch interactions were not significant (all $p > 0.1$).

Table 8 gives the estimates of the coefficients for the Nominalization regression. The absolute values of the coefficients are plotted in Figure 17(b), and condition means are plotted in Figure 18(b).

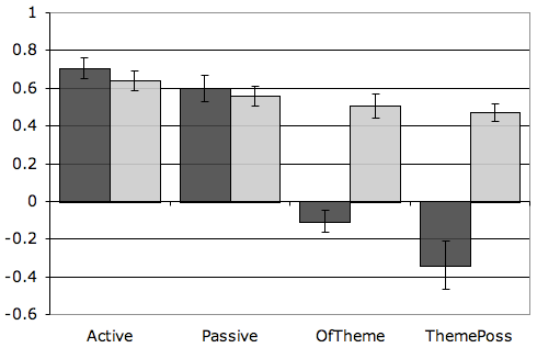


(a) Adjectivizations

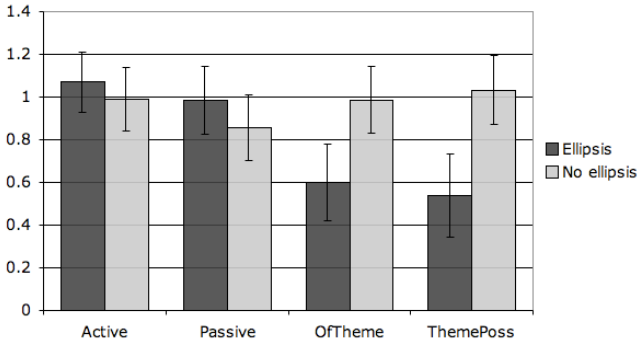


(b) Nominalizations

Figure 17: Absolute values of model coefficients for Experiments 2 and 3 (error bars represent standard error).



(a) Adjectivizations



(b) Nominalizations

Figure 18: Condition means for Experiments 2 and 3 (error bars represent standard error).

Table 8: Experiment 3: Estimates of fixed effects

	Coefficient	*Standard error*	t	p
Intercept	0.38	0.078	4.85	< 0.001
Ellipsis	-0.14	0.014	-10.08	< 0.001
Canonical	-0.046	0.014	-3.23	< 0.01
Mismatch	-0.23	0.014	-15.99	< 0.001
Ellipsis:Canonical	-0.014	0.014	-0.96	*n.s.*
Ellipsis:Mismatch	-0.17	0.014	-12.17	< 0.001
Canonical:Mismatch	0.0075	0.014	0.53	*n.s.*
Ellipsis:Canonical:Mismatch	-0.012	0.014	-0.82	*n.s.*

4.3 Discussion

In Experiments 2 and 3, whether the two clauses matched structurally had a strong influence on acceptability. This effect interacted strongly with whether the second clause contained ellipsis, replicating our findings from Experiment 1.

In addition, the coefficient on the Canonicality predictor, in the fitted regression models for Experiments 2 and 3, is significantly different from zero (Figure 17). This means that across both nominalization and adjectivizations there is a reliable effect of Canonical Realization in the measured acceptability ratings, over and above the preference for matching in VPE. This effect would be accounted for if Canonical Realization were true of the human sentence comprehension mechanism, under the assumptions detailed in section 2.3.

Comparing the effects of Canonicality and Mismatch in the regression models reveals an interesting difference. The fact that we see significant main effects of both factors indicates that both Canonicality and Mismatch influence acceptability in a general way, regardless of whether a sentence contains ellipsis. This appears to be the *only* way that Canonicality affects acceptability—it is a general processing pressure. This can be seen in the nonsignificance of any of the interactions involving Canonicality in the regression models (Tables 7 and 8). In contrast, the preference for syntactically matching clauses appears to be particularly sensitive to the presence

of ellipsis: when the second clause contains ellipsis, it becomes much more important for the clauses to match syntactically than when there is no ellipsis. These findings can be seen as contributing to the accumulating body of literature on effects of parallelism in processing [Carlson, 2001, Chambers and Smyth, 1998]; in particular, we offer a partial answer to the question of what linguistic environments induce the strongest parallelism effects.

5 General Discussion

The results of Experiments 1,2 and 3 support the proposed parsing preferences MaxElide and Canonical Realization. But what sort of parsing architecture would actually use these preferences? Since one key advantage of the current proposal is its compatibility with standard notions of parsing, this section reviews some of those notions and explains how ellipsis resolution could work.

5.1 Parsing preferences

Since at least Kaplan [1973], efficient parsing has relied on a 'cache' or store of information about the sentence being analyzed. Such a cache is known as a CHART.[15] A chart parser applies the 'fundamental rule' over and over again to systematically explore the space of parser states relevant to the word string it has been given. The chart itself is a data structure that stores parser states so that they are not re-examined a second time. It is complemented by another data structure, the AGENDA, that keeps track of as-yet-unexplored possibilities. Kay [1986] notes that an agenda ordering can be construed as a psycholinguistic hypothesis:

> A model based on the agenda can associate priorities with tasks in more or less complex ways and can thus ascribe the variation observed in experimental results to a variety of sources
>
> page 68

[15] Chart parsing is described in all computational linguistics textbooks. See for instance Gazdar and Mellish [1989, chapter 5], Allen [1995, chapter 3] or Jurafsky and Martin [2009, §13.4.3].

In natural language processing, the priorities associated with agenda entries are typically probabilities [Stolcke, 1995, Roark, 2001, Klein and Manning, 2004] although this need not be the case [Caraballo and Charniak, 1998, Dzikovska et al., 2005]. One can equally apply heuristic constraints like MaxElide and Canonical Realization to order the agenda such that subtrees consistent with the constraints are explored before subtrees that are inconsistent with them. To be more specific, if a parser has a choice to do a unit of work that either will build a derivation that respects MaxElide or fails it, it will do the unit of work that respects the constraints. Ties may be broken arbitrarily or by applying some theory of constraint interaction such as Optimality Theory [Prince and Smolensky, 2004]. The important consequence of this agenda-ordering is that sentences that violate the constraints require more applications of the fundamental rule because all of the analyses consistent with the constraints had to be explored first. It is this heightened amount of computational effort that we offer in this paper as an explanation for the less acceptable cases of VPE described in sections 3 and 4.

5.2 Ellipsis resolution

One can break up the problem of parsing VPE examples into three steps

i. determining whether something has been elided

ii. locating the antecedent

iii. substituting the antecedent into the ellipsis site

As regards the first step, we follow Lappin [1990, 1996, 1997, 1999] in assuming that surface cues can trigger a search for VPE antecedents. This assumption is buttressed by empirical studies like Hardt [1997] and Nielsen [2004] who find that auxiliary verbs in the Penn Treebank can be diagnosed as part of VPE or not at rates in excess of 70%.

As regards the second step, a chart parser may be straightforwardly modified to allow derivations in an initial conjunct to do double-duty in a second conjunct. A method like that of Lavelli and Stock [1990] would immediately account for the experimental results obtained in [Frazier et al., 2000]. In virtue of using two charts, this method is applicable to both intra-sentential and cross-sentential ellipsis. Indeed, Kim and Runner [2009] have shown experimentally that the same gradient pattern of acceptability reported in this paper extends almost identically to cases of cross-sentential VP ellipsis.

The third and final step is rendered trivial within the deletion-under-identity approach. On this approach, once a VPE site is located and an appropriate antecedent is found, the derivation of the antecedent becomes available to the parser, just as it if were located at the elision point in the input string. No further operations are required above and beyond the usual compositional semantics (e.g. [Kobele, 2006, chapter 2]).

5.3 Functional motivation

5.3.1 Origins of these heuristics

By way of interpretation, we suggest that acceptability cline in VPE reflects comprehension difficulty. Comprehenders have more difficulty understanding structures that can be generated by the grammar just in cases where they violate the parser heuristics MaxElide and Canonical Realization. This pattern of results, however, does not pinpoint the reason why MaxElide and Canonical Realization should hold in the first place. There may be some aspect of cognition in general or the language faculty in particular that leads to these preferences. If these preferences reflect the distribution of structures in a comprehender's language community, then the heuristics would be explained as rational reflexes of an adaptive cognitive system that seeks to capitalize on accurate estimates of what people are likely to say [Hale, 2009].

5.3.2 Failure to elide

The results of Experiment 3 in particular suggest that a failure to elide can result in lower acceptability compared to VPE stimuli. This echos Fiengo and May's [1994] invocation of the functional pressure to elide when possible.

> an antecedent is apparently necessary for an ellipsis. But this stems, we believe, from the role of ellipsis in the theory of use as a device for the reduction of redundancy. For such reduction to be effective, there must be some token expression with respect to which the elided material would have been redundant. The latter expression is the antecedent of the ellipsis.
>
> page 192

This sort of pressure to avoid redundancy is also attested in the 'repeated name penalty' [Gordon et al., 1993]. However the question of where best

to account for it remains. The Condition on Recoverability of Deletion represents an attempt to account for this observed tendency within the competence grammar. The repair rules of Arregui et al. [2006] seek to account for the parallelism preference by broadening the computational repertoire of the processing mechanism. The parsing heuristics proposed here represent a third way. They encode knowledge not about *what to do* but rather knowledge about *what to try first*. As a kind of 'control' information [Lewis, 2000] in the processor, they bridge the gap between the step-counting complexity metrics of the 1970s [Kaplan, 1972, Frazier and Fodor, 1978] and the violated-probabilistic-expectations approaches of the early 2000s [Hale, 2001, 2006, Levy, 2008] without having to postulate an enlarged set of processing mechanisms. As Hale [2009] details, an appropriate set of heuristics may direct a parser to impose an information-theoretic prior during inference — that is, not to expect redundancy.

5.4 Extensions

The lexical decomposition approach of section 2.1.2 extends naturally to other constructions. One such is the verbal gerunds studied in Arregui et al. [2006] and exemplified below as (17).[16]

(17) The candidate was dogged with charges of avoiding the draft, or at least trying to ~~avoid the draft~~. [Hardt, 1993]

This case is one of the three, shown in Figure 19, that Abney [1987] analyzes. Abney's proposal treats these gerunds as combinations of a nominalizing head ing with a verbal category at one of three different positions. The natural extension of section 2.1.2's account views the ing+*of* in (17) as a morphological variant of the V+nom construction discussed in that section.

[16] Agentive nominalizations are also attested in elliptical contexts [Hardt, 1993]:

i. Harry used to be a great speaker, but he can't ~~speak~~ anymore, because he has lost his voice.

To capture the possibility that agentive nominalizations may serve as antecedents for the verbs they are derived from, we would syntactically decompose them into a verb plus a category changing head -er. Our intuitive judgments suggest that there is a cline in acceptability here as well. For instance, we find that speaker serves as a better antecedent for speak than does computer for compute. This acceptability cline reflects attestation rates. A model of lexical access along the lines of Hay [2003] would relate these two asymmetries.

type	attachment point
Poss-ing	IP
Acc-ing	VP
ing+*of*	V^0

Figure 19: Abney's analysis of gerundival NPs

In other words, admit+nom could be understood in greater detail as the structure $[_{N^0}\ [_{V^0}\ admit]\ [_{N^0}\ ing]]$.

Considering the other two cases leads to novel predictions. An updated version [Schueler, 2004] of Abney's analysis attaches ing at little vP in the so-called Acc-ing construction, and at the perfective aspect head in Poss-ing constructions.[17] This predicts that both possessive and accusative ing-gerunds may occur grammatically in voice mismatch VPE. Two such examples are presented in (18a) and (18b).

(18) a. These documents being released would be good, but we refuse to.

b. Us releasing these documents would be dangerous, and so they won't be.

Arregui et al. [2006] observe lower acceptability with ing-*of* gerunds as antecedents for VP ellipsis than with other kinds of gerunds. The proposal in section 2 is consistent with such a finding, as ing-*of* gerunds only allow for elision below the voice head, in violation of MaxElide. As discussed in footnote 17, Acc-ing and Poss-ing gerunds contain VPs specified for voice. Our analysis allows for gerunds to elide above the voice head. They can thus serve as antecedents in VPE without violating MaxElide. It remains mysterious under both accounts why Arregui et al. [2006] find that these latter more verbal gerunds are less acceptable in ellipsis contexts than normal VPs. If this result is indeed systematic, then it suggests a role for other

[17]Note that the ing in an Acc-ing construction really does attach higher than the voice head.

i. Him asking Mary out is extremely unlikely.

ii. Her being asked out is wildly implausible.

parsing heuristics beyond the ones considered in this paper.

6 Conclusion and Future Work

Adding deletion under a simple and strict notion of identity to a well-defined grammar formalism yields an analysis which can directly derive mismatching items. Their acceptability properties, we have argued, follow from heuristic preferences that guide normal sentence comprehension. They reflect generalizations that have been previously recognized in various corners of the cognitive science of language. This approach leads to a simple conception of the acceptability cline in VPE, and brings experimental data directly to bear upon theoretical proposals.

The grammatical analysis in section 2 has implications for sluicing ('IP-ellipsis'), where voice mismatches are thought to be ungrammatical [Merchant, 2001]. Because IPs are specified for voice, sluicing will be grammatical only if the antecedent and deleted IPs match along this dimension. By contrast, Arregui et al. [2006] assume that mismatches are uniformly ungrammatical. The very extra-grammaticality of their repair-rule approach suggests that all cases of mismatched voice ellipsis, no matter what their syntactic type, should be repairable. This view predicts that voice mismatches in VPE should pattern identically with voice mismatches in sluicing. Adjudication between these alternatives awaits future work.

References

Steven P. Abney. *The English Noun Phrase in its Sentential Aspect*. PhD thesis, Massachusetts Institute of Technology, 1987.

James Allen. *Natural Language Understanding*. Benjamin/Cummings, 1995.

Ana Arregui, Charles Clifton, Jr., Lyn Frazier, and Keir Moulton. Processing elided verb phrases with flawed antecedents: The recycling hypothesis. *Journal of Memory and Language*, 55:232–246, 2006.

R. Harald Baayen. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge University Press, 2008.

Emmon W. Bach. In defense of passive. *Linguistics and Philosophy*, 2:297–341, 1980.

Ellen Gurman Bard, Dan Robertson, and Antonella Sorace. Magnitude estimation of linguistic acceptability. *Language*, 72(1):32–68, 1996.

Douglas Bates and Deepayan Sarkar. *lme4: Linear mixed-effects models using S4 classes*, 2007. R package version 0.99875-9.

Robert C. Berwick and Amy S. Weinberg. *The Grammatical Basis of Linguistic Performance*. MIT Press, 1984.

Thomas G. Bever. The cognitive basis for linguistic structures. In J.R. Hayes, editor, *Cognition and the Development of Language*, pages 279–362. Wiley, New York, 1970.

Luca Bonatti. PsyScope X, Build 53. PsyScope X Project. SISSA Language, Cognition and Development Lab, 2008. URL `http://psy.ck.sissa.it/`.

Sharon Caraballo and Eugene Charniak. New figures of merit for best-first probabilistic chart parsing. *Computational Linguistics*, 24(2):275–298, 1998.

Katy Carlson. The effects of parallelism and processing in the processing of gapping structures. *Language and Speech*, 44(1):1–26, 2001.

Craig G. Chambers and Ron Smyth. Structural parallelism and discourse coherence: A test of centering theory. *Journal of Memory and Language*, 39:593–608, 1998.

Eugene Charniak, Sharon Goldwater, and Mark Johnson. Edge-based best-first chart parsing. In *Proceedings of the Sixth Workshop on Very Large Corpora*, 1998.

Noam Chomsky. *Current Issues in Linguistic Theory*. Mouton, 1964.

Noam Chomsky. *Aspects of the Theory of Syntax*. MIT Press, 1965.

Noam Chomsky. *The Minimalist Program*. MIT Press, Cambridge, Massachusetts, 1995.

Noam Chomsky. Minimalist inquiries: The framework. In Roger Martin, David Michaels, and Juan Uriagereka, editors, *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, pages 89–155. MIT Press, Cambridge, Massachusetts, 2000.

Sandra Chung. Sluicing and the lexicon: The point of no return. In Rebecca T. Corver and Yuni Kim, editors, *BLS 31: General Session on Parasession on Prosodic Variation and Change*, pages 73–91, Berkeley, CA, 2006. Berkeley Linguistics Society.

Matthew W. Crocker and Thorsten Brants. Wide-coverage probabilistic sentence processing. *Journal of Psycholinguistic Research*, 29(6):647–669, 2000.

Mary Dalrymple, Stuart M. Shieber, and Fernando C.N. Pereira. Ellipsis and higher-order unification. *Linguistics and Philosophy*, 14:399–452, 1991.

Myroslava O. Dzikovska, Mary D. Swift, James F. Allen, and William de Beaumont. Generic parsing for multi-domain semantic interpretation. In *Proceedings of the International Workshop on Parsing Technologies*, 2005.

Robert Fiengo and Robert May. *Indices and Identity*. MIT Press, 1994.

Danny Fox. *Economy and Semantic Interpretation*. MIT Press, Cambridge, Massachusetts, 2000.

Danny Fox. Antecedent-contained deletion and the copy-theory of movement. *Linguistic Inquiry*, 33(1):63–96, 2002.

Lyn Frazier. *On Comprehending Sentences: syntactic parsing strategies*. PhD thesis, University of Connecticut, 1979.

Lyn Frazier. Processing ellipsis: A processing solution to the undergeneration problem? In Charles B. Chang and Hannah J. Haynie, editors, *Proceedings of the 26th West Coast Conference on Formal Linguistics*, pages 21–32, 2008.

Lyn Frazier and Charles Clifton. The syntax-discourse divide: Processing ellipsis. *Syntax*, 8(2):421–430, 2005.

Lyn Frazier and Janet Dean Fodor. The sausage machine: a new two-stage parsing model. *Cognition*, 6:291–325, 1978.

Lyn Frazier, Alan Munn, and Charles Clifton. Processing coordinate structures. *Journal of Psycholinguistic Research*, 29(4):421–430, 2000.

Hans-Martin Gärtner. *Generalized Transformations and Beyond: Reflections on Minimalist Syntax*. Akademie Verlag, Berlin, 2002.

Gerald Gazdar and Chris Mellish. *Natural Language Processing in Lisp: an introduction to computational linguistics*. Addison Wesley, 1989.

Andrew Gelman and Jennifer Hill. *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press, New York, 2007.

Gerd Gigerenzer, Peter M. Todd, and the ABC research group. *Simple Heuristics that Make us Smart*. Oxford University Press, 1999.

Peter C. Gordon, Barbara J. Grosz, and Laura A. Gilliom. Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, 17:311–347, 1993.

Howard Gregory and Shalom Lappin. A computational model of ellipsis resolution. In Geert-Jan Kruijff, Glyn V. Morrill, and Richard T. Oehrle, editors, *Formal Grammar: proceedings of the conference*, 1997.

John Hale. A Probabilistic Earley Parser as a Psycholinguistic Model. In *Proceedings of the Second Meeting of the North American Chapter of the Association for Computational Linguistics*, 2001.

John Hale. Uncertainty about the rest of the sentence. *Cognitive Science*, 30(4): 609–642, 2006.

John Hale. What a rational parser would do. in revision, 2009.

Daniel Hardt. *Verb phrase ellipsis: Form, meaning, and processing*. PhD thesis, University of Pennsylvania, 1993.

Daniel Hardt. An empirical approach to VP Ellipsis. *Computational Linguistics*, 23(4):525–541, 1997.

Henk Harkema. *Parsing Minimalist Languages*. PhD thesis, University of California, Los Angeles, 2001.

Jennifer Hay. *Causes and Consequences of Word Structure*. Routledge, New York, 2003.

Aravind K. Joshi. Tree adjoining grammars: How much context-sensitivity is required to provide reasonable structural descriptions? In D.R. Dowty, L. Karttunen, and A. M. Zwicky, editors, *Natural Language Parsing: Psychological, Computational and Theoretical Perspectives*, pages 206–250. Cambridge University Press, New York, 1985.

Aravind K. Joshi. Starting with complex primitives pays off: complicate locally, simplify globally. *Cognitive Science*, 28(5):637–668, 2004. 2003 Rumelhart Prize Special Issue Honoring Aravind K. Joshi.

Daniel Jurafsky. A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, 20:137–194, 1996.

Daniel Jurafsky and James H. Martin. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition*. Prentice-Hall, 2nd edition, 2009.

Ronald M. Kaplan. Augmented transition networks as psychological models of sentence comprehension. *Artificial Intelligence*, 3:77–100, 1972.

Ronald M. Kaplan. A general syntactic processor. In Randall Rustin, editor, *Natural Language Processing*. Algorithmics Press, 1973.

Martin Kay. Algorithm schemata and data structures in syntactic processing. In Barbara J. Grosz, Karen Sparck Jones, and Bonnie Lynn Webber, editors, *Readings in Natural Language Processing*, pages 35–70. Morgan Kaufmann, 1986. Also Xerox technical report CSL-80-12.

Richard S. Kayne. *The Antisymmetry of Syntax*. MIT Press, 1994.

Edward Keenan. Passive is phrasal (not sentential or lexical). In Teun Hoekstra, Harry van der Hulst, and Michael Moortgat, editors, *Lexical Grammar*, volume 3 of *Publications in Language Sciences*, chapter 7, pages 181–214. Foris Publications, Dordrecht, 1980.

Christina S. Kim and Jeffrey T. Runner. Strict identity, coherence, and parallelism in VP ellipsis. In *Proceedings of Semantics and Linguistic Theory XIX*, Columbus, OH, 2009.

Dan Klein and Christopher D. Manning. Parsing and hypergraphs. In Harry Bunt, John Carroll, and Giorgio Satta, editors, *New Developments in Parsing Technology*, pages 351–372. Kluwer Academic Publishers, 2004.

Gregory M. Kobele. *Generating Copies: An investigation into structural identity in language and grammar*. PhD thesis, University of California, Los Angeles, 2006.

Gregory M. Kobele. Syntactic identity in survive minimalism: Ellipsis and the derivational identity hypothesis. In Michael T. Putnam, editor, *Towards a purely derivational syntax: Survive-minimalism*. John Benjamins, 2009.

Gregory M. Kobele. A formal foundation for A and A-bar movement in the minimalist program. In Christian Ebert, Gerhard Jäger, and Jens Michaelis, editors,

MOL 10/11, volume 6149 of *Lecture Notes in Computer Science*, pages 145–159. Springer, 2010.

Gregory M. Kobele. Deriving reconstruction asymmetries, to appear.

Gregory M. Kobele, Christian Retoré, and Sylvain Salvati. An automata theoretic approach to minimalism. In James Rogers and Stephan Kepser, editors, *Proceedings of the Workshop Model-Theoretic Syntax at 10; ESSLLI '07*, Dublin, 2007.

Masatoshi Koizumi. *Phrase Structure In Minimalist Syntax*. PhD thesis, Massachusetts Institute of Technology, 1995.

Hilda Koopman and Dominique Sportiche. The position of subjects. *Lingua*, 85: 211–258, 1991.

Marcus Kracht. Syntax in chains. *Linguistics and Philosophy*, 24(4):467–529, 2001.

Henry Kučera and W. Nelson Francis. *Computational Analysis of Present-day American English*. Brown University Press, 1967.

Shalom Lappin. An HPSG account of antecedent-contained ellipsis. In *Fragments: studies in Ellipsis and Gapping*, pages 68–97. Oxford University Press, 1999.

Shalom Lappin and Michael McCord. Anaphora resolution in slot grammar. *Computational Linguistics*, 16(4):197–212, 1990.

Shalom Lappin and Hsue-Hueh Shih. A generalized reconstruction algorithm for ellipsis resolution. In *Proceedings of COLING-96*, pages 687–692, Copenhagen, 1996.

Howard Lasnik. Verbal morphology: *Syntactic Structures* meets the minimalist program. In Paula Kempchinsky and Héctor Campos, editors, *Evolution and Revolution in Linguistic Theory: Essays in Honor of Carlos Otero*. Georgetown University Press, Georgetown, 1995.

Howard Lasnik. Chains of arguments. In Samuel David Epstein and Norbert Hornstein, editors, *Working Minimalism*, number 32 in Current Studies in Linguistics, chapter 8, pages 189–215. MIT Press, Cambridge, Massachusetts, 1999.

Alberto Lavelli and Oliviero Stock. When something is missing: ellipsis, coordination and the chart. In *Proceedings of COLING-90*, pages 184–189, 1990.

David Lebeaux. *Language Acquisition and the Form of Grammar*. PhD thesis, University of Massachusetts, 1988.

Robert B. Lees. *The grammar of English nominalizations*. Indiana Research Center in Anthropology, Folklore and Linguistics, 1960.

Roger Levy. Expectation-based syntactic comprehension. *Cognition*, 106(3):1126–1177, 2008.

Richard L. Lewis. Specifying architectures for language processing: Process, control, and memory in parsing and interpretation. In Matthew W. Crocker, Martin Pickering, and Charles Clifton, Jr., editors, *Architectures and mechanisms for language processing*. Cambridge University Press, 2000.

Maria Rita Manzini and Anna Roussou. A minimalist theory of A-movement and control. *Lingua*, 110(6):409–447, 2000.

Mitchell P. Marcus, Beatrice Santorini, and Mary Ann Marcinkiewicz. Building a large annotated corpus of English: the Penn Treebank. *Computational Linguistics*, 19, 1993.

Jason Merchant. *The Syntax of Silence: Sluicing, Islands, and the Theory of Ellipsis*, volume 1 of *Oxford Studies in Theoretical Linguistics*. Oxford University Press, New York, 2001.

Jason Merchant. Voice and ellipsis. ms., University of Chicago, 2007.

Jason Merchant. An asymmetry in voice mismatches in VP-ellipsis and pseudo-gapping. *Linguistic Inquiry*, 39(1):169–179, 2008a.

Jason Merchant. Variable island repair under ellipsis. In Kyle Johnson, editor, *Topics in Ellipsis*, chapter 6, pages 132–153. Oxford University Press, Oxford, 2008b.

Jens Michaelis. *On Formal Properties of Minimalist Grammars*. PhD thesis, Universität Potsdam, 2001.

Gereon Müller. On deriving CED effects from the PIC. ms., Universität Leipzig, 2008.

Allen Newell and Herbert A. Simon. *Human Problem Solving*. Prentice-Hall, Englewood Cliffs, New Jersey, 1972.

Leif Arda Nielsen. Verb Phrase Ellipsis detection using automatically parsed text. In *Proceedings of COLING-04*, pages 1093–1099, 2004.

Fernando Pereira. A new characterization of attachment preference. In David Dowty, Lauri Karttunen, and Arnold Zwicky, editors, *Natural Language Parsing: Psychological, Computational and Theoretical Perspectives*, ACL Studies in Natural Language Processing, pages 307–319. Cambridge University Press, 1985.

David Pesetsky. *Zero Syntax: experiencers and cascades*. MIT Press, 1995.

Alan Prince and Paul Smolensky. *Optimality theory: constraint interaction in generative grammar*. Blackwell, 2004.

Brian Roark. Probabilistic top-down parsing and language modeling. *Computational Linguistics*, 27(2):249–276, 2001.

Jeffrey T. Runner. *Noun Phrase Licensing and Interpretation*. PhD thesis, University of Massachusetts, Amherst, 1995.

Jeffrey T. Runner. Lingering challenges to the raising-to-object and object-control constructions. *Syntax*, 9(2):193–213, 2006.

Ivan A. Sag. *Deletion and Logical Form*. PhD thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1976.

David Schueler. Extended-projection, categorization, and the English morpheme *-Ing*. In *Proceedings of WECOL-33*, pages 342–356, 2004.

Carson T. Schütze. *The empirical base of linguistics: grammaticality judgments and linguistic methodology*. University of Chicago Press, 1996.

Edward Stabler. Derivational minimalism. In Christian Retoré, editor, *Logical Aspects of Computational Linguistics*, pages 68–95. Springer, 1997.

Suzanne Stevenson and Paul Smolensky. *The Harmonic Mind: from neural computation to optimality-theoretic grammar*, volume 2, chapter 19, pages 307–338. MIT Press, 2006.

Andreas Stolcke. An efficient probabilistic context-free parsing algorithm that computes prefix probabilities. *Computational Linguistics*, 21(2), 1995.

Shoichi Takahashi and Danny Fox. MaxElide and the Re-binding Problem. In Effi Georgala and Jonathan Howell, editors, *Proceedings of SALT 15*, pages 223–240, Ithaca, NY, 2005. CLC Publications.

Shoichi Takahashi and Sarah Hulsey. Wholesale late merger: Beyond the A/$\overline{\text{A}}$ distinction. *Linguistic Inquiry*, 40(3):387–426, 2009.

Michael K. Tanenhaus and Greg N. Carlson. Comprehension of deep and surface verb phrase anaphors. *Language and Cognitive Processes*, 5(4):257–280, 1990.

Satoshi Tomioka. *Focusing Effects and NP Interpretation in VP Ellipsis*. PhD thesis, University of Massachusetts, Amherst, 1997.

David J. Townsend and Thomas G. Bever. *Sentence comprehension : the integration of habits and rules*. MIT Press, 2001.