Syntax
A Journal of Theoretical, Experimental and Interdisciplinary Research

Syntax 2011 DOI: 10.1111/j.1467-9612.2011.00160.x

The Acceptability Cline in VP Ellipsis

Christina S. Kim, Gregory M. Kobele, Jeffrey T. Runner, and 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 1 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 synonymous 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 (ii) 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 (ii) underlines the urgency of problem (i). 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

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).

standpoint of language use, because 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, Shieber & Pereira 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; Kennedy 2003; Merchant 2007, 2008a; Kobele 2009) suggests that the identity condition in ellipsis is sensitive to properly *syntactic* distinctions. Second, work in psycholinguistics (Tanenhaus & Carlson 1990, Arregui et al. 2006, Frazier & Clifton 2006) suggests that comprehenders are indeed sensitive to such fine-grained syntactic distinctions in VPE.[1] 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.

[1] Kertz (2008), in a magnitude estimation study similar to the one presented here, argues that information structure influences acceptability, in sentences both with and without ellipsis. The current study holds information structure constant and therefore cannot comment on potential effects of information structure. However, our proposal is entirely consistent with the existence of additional information structural effects on acceptability. Section 5.3 addresses the problem of integrating multiple information sources in a single parsing mechanism.

Arregui et al. (2006) observe a cline of acceptability values across VPE that match or mismatch in a variety of 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 2011). This approach is consistent with the views of Grodner et al. (2003), who argue that the use of repair rules in syntactic processing is to be dispreferred when alternative accounts of the same phenomena are available, which involve only the standard mechanisms available in "first-pass" parsing. 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 approach 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

Our proposal is based on the observation that the decompositional style of analysis typical of the Minimalist Program allows us to characterize even cases of "mismatching" ellipsis as governed by deletion under identity. In section 2.1 we present our syntactic analysis of the constructions usually taken to involve "inference" of some kind (Hardt 2004). We will work within the broad assumptions of the Minimalist Program. In accord with recent minimalist ideas, linear order will not be directly represented in our syntactic structures (in the sense of the order in which words are pronounced). We assume that the surface ordering

is determined by a linearization algorithm at the relevant interface (Kayne 1994).[2] Chomsky (1995) proposes "bare phrase structure" (BPS), which replaces the familiar X-bar notation with simple copies of the label of the daughter (which is ultimately the lexical item heading the projection). As noticed by Stabler (1997) (see also Collins 2002), labeling can be even more economically represented in terms of ordering sister subtrees, with the projecting daughter coming before the daughter projected over. These notations are presented in Figure 1. We tentatively adopt the Stablerian perspective, if only for the ease it provides in the depiction of syntactic structures.

Following Gärtner (2002), who interprets Chomsky's (1995) "copy theory of movement" in terms of multiple dominance, we take the basic data structure of minimalism to be a special kind of rooted directed acyclic graph—a multiple dominance structure (Kracht 2001). Alternative but equivalent data structures are trace chain structures or copy chain structures, as shown in Figure 2.[3] Although it is more standard to use structures with copies (although see, among others, Starke 2001, Gärtner 2002, Zhang 2004, Citko 2005, and de Vries 2009), the multiple domination representation seems most faithful to Chomsky's admonition against treating copies as distinct and will allow us to state our conditions on deletion most simply.

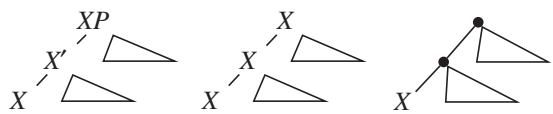


Figure 1. Three notations for headedness

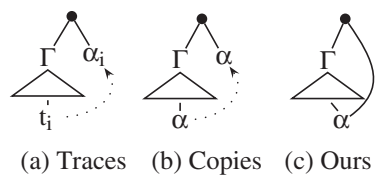


Figure 2. Equivalent representations of syntactic structure

[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, Retoré & Salvati 2007)—this is the standard in presentations that leave move arcs implicit and thus in which expressions are structured as trees (as in Stabler 1997).

[3] Kracht (2001) provides a thorough comparison of these three possibilities. Among other things, he shows that, given natural constraints on movement, all three encode the same information (each representational scheme can be uniquely translated into any of the others), and thus that, from a mathematical perspective, they are notational variants of each other.

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 (2009) Minimalist Grammars (MGs) extended with hypothetical reasoning (Kobele 2010). Minimalist Grammars provide a formal framework for reasoning about work in the Minimalist Program. Hypothetical reasoning, in this technical sense, can be viewed as an implementation of ideas about A- and A′-movement proposed by Manzini & Roussou (2000) (see also Lasnik 1999). These ideas have recently been reformulated by Takahashi & 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 multidominance structures as shown schematically in Figure 3. 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. To see how these operations apply in clause-level syntax, consider the structures in Figure 4. These multidominance structures present a basic analysis of active and passive English sentences, to be further fleshed out later in the paper in Figure 8. The structure for the passive in Figure 4b 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).

Multidominance structures are but one of the notations in which MG-derived expressions can be written. The same expressions can equivalently be described by

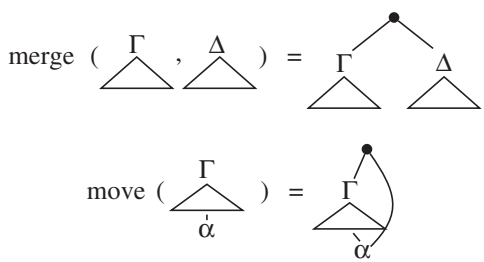


Figure 3. The operations merge and move

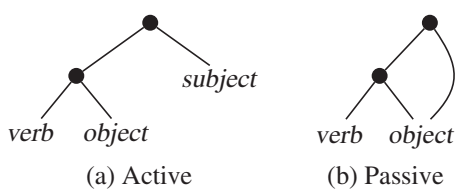


Figure 4. Rough structure of active and passive sentences

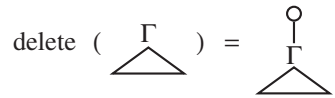


Figure 5. The operation delete

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 5. The phonological consequences of the delete operation are clear. Syntactically, delete extends a multidominance structure upward by one node. In Figure 5 this node is labeled with a white dot.[4] 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 later 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.

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 6.

In Figure 6, both conjuncts fill their subject position by merger with nodes labeled "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 4b. A passive–active VPE mismatch item like (4) thus has the structure shown in Figure 7.

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

[4] 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 does not alter the featural constitution of its complement.

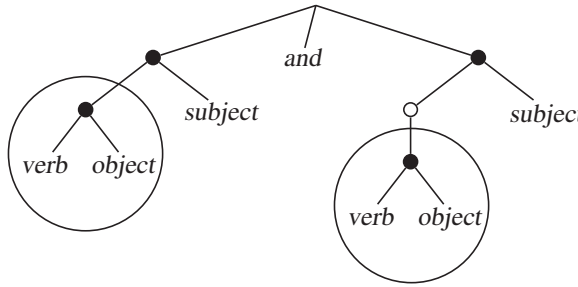


Figure 6. Rough structure of a VPE item like (3)

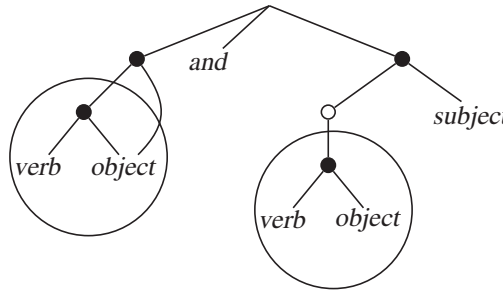


Figure 7. Rough structure of passive–active voice mismatch item like (4)

The strict syntactic identity requirement requires that passives and actives share a common multidominance substructure. This *identical* substructure is circled in Figures 6 and 7. 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 by Kobele (2006), synthesizing the ideas of passivization as Case-checking, the VP-internal subject hypothesis (Koopman & Sportiche 1991), and overt object Case-checking (Koizumi 1995, Runner 1995).

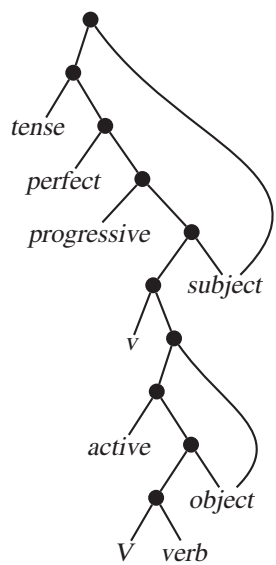
The substance of the analysis resides primarily in the syntactic properties of lexical items. Each lexeme includes a bundle of *features*. We differentiate between *attractor* features, written with an asterisk before their name, $*z$, and *attractee* features, which lack this asterisk, z. Operations 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 noncomplements. We assume with Stabler (1997) that features within a bundle are totally ordered (see also Müller 2010). For example, the feature bundle for the proper name *John* might be associated with a sequence of two attractee features as in Table 1. The first feature `d` means that this lexeme can occupy DP positions in the syntax. It is selected by expressions with the corresponding attractor feature `*d`. The second feature, `k`, encodes a requirement for Case. This requirement could be satisfied by movement, triggered by `*k`, to the specifier of some other head.

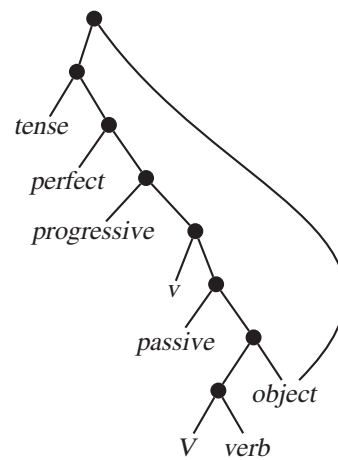
Figure 8 sums up our proposal about the structure of active and passive sentences. These more detailed structures illustrate the role of a voice head that differentiates active (lexical item: ⟨ϵ *V *k act⟩) from passive (lexical item: ⟨−*en*, *V pass⟩). Whereas the passive voice head contributes an overt 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 `*k` attractor feature, the active voice head can check the case of a direct object. Lack of such a

Table 1. Lexical entry for a proper name

⟨*John*,d k⟩



(a) Refined structure for actives (see Figure 4a)



(b) Refined structure for passives (see Figure 4b)

Figure 8. More detailed structures corresponding to Figure 4

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 "v."[5]

In our fragment, v is broken up into two types, defective and nondefective. Defective v (lexical item: ⟨*be*, ∗pass v⟩) is unable to license an external argument, whereas nondefective v can do so in virtue of an extra attractee feature (lexical item: ⟨ϵ, ∗act ∗d v⟩) (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: ⟨*tense*, ∗perf ∗k t⟩).

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 8a and 8b 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 V and its selected verb.[6] 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.[7] This situation is analogous to the behavior of A-moved expressions with respect to reconstruction (Fox 2000). In reconstruction, 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 & 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 2011, Takahashi & Hulsey 2009).

[5] The passive voice head that we employ plays the same role as Agr_O. The key property is that the needed head be situated between V and 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.

[6] 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.

[7] 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 behavior 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.

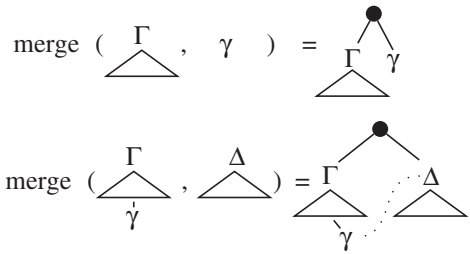


Figure 9. Late merger

In our analysis of VPE, we apply a version of this same mechanism (Manzini & 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.[8] This is shown in Figure 9, 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 10b (cf. Figure 8b). In Figure 10, the attractor feature ∗d of the verb is checked by the merger of the feature bundle ⟨d k⟩. This feature bundle belongs to the object, which is merged directly in its surface position.

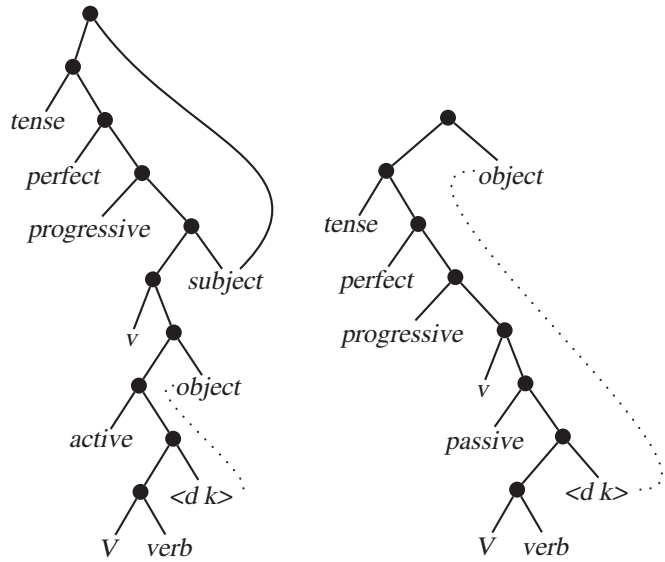
With late merger in mind, consider sentence (5), in which a passive VP serves as the antecedent for a passive ellipsis site.

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

Where the second conjunct is in the passive voice, the deep direct objects in VPE would be nonidentical 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 10a.

[8] This version of the late merger mechanism differs from that advanced by Takahashi & 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 nonidentical Ds are able to head the NPs in passive–passive ellipsis:

(i) John seems to have been kissed, and every ninja does ~~seem to have been kissed~~, too.



(a) Late merged object in active (cf. Figure 8a)

(b) Late merged object in passive (cf. Figure 8b)

Figure 10. Alternative derivations involving late-merged objects

Table 2. Lexical items (I)

⟨*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⟩	

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

The lexical items needed to derive (5) and (6) in this manner are given in Table 2.

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

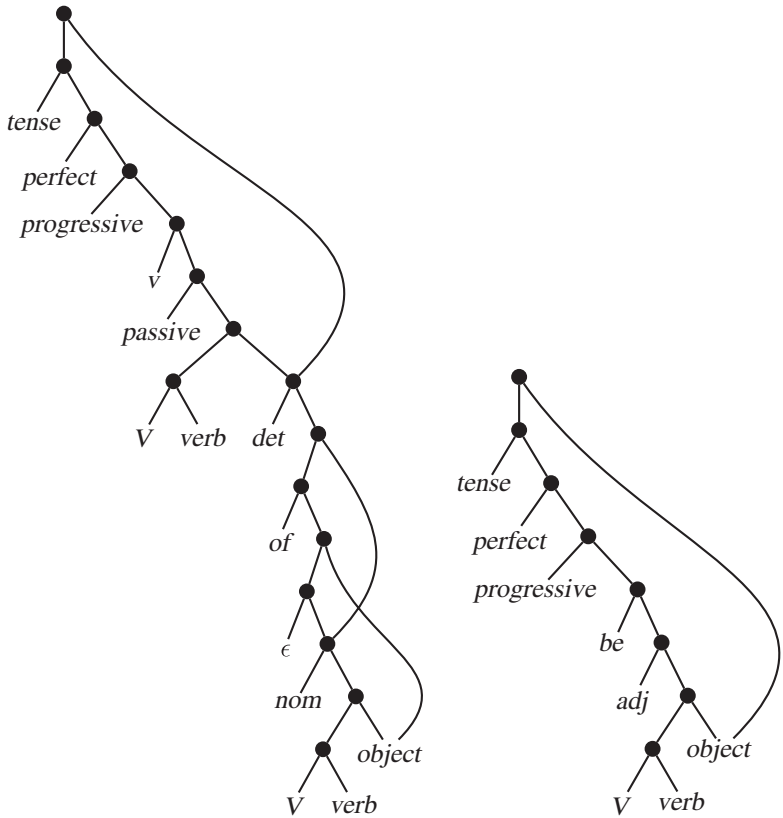


Figure 11. Structures underlying examples (7a) and (7b)

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⟩).[9] Derivations for these two classes of examples 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~~.

[9] 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 statable as such in the present system, we have opted for a work-around involving remnant movement in the spirit of Kayne 1994. The informal sketch here is captured precisely via the lexical items in Figure 3.

(i) a. [nom [V Obj]] (merge nominalizing head with the VP, which introduces a feature p)
b. [ϵ [nom [V Obj]]] (merge another empty head with the VP, which checks the object's Case)
c. [Obj [ϵ [nom [V *t*]]]] (object moves for Case)
d. [of [Obj [ϵ [nom [V *t*]]]]] (merge *of*, triggers remnant movement of the VP)
e. [[nom [V *t*]] [of [Obj [ϵ *t*]]]] (remnant VP moves to check feature p)

Table 3. Lexical items (II)

⟨*suspect*, n⟩	⟨*guilt*, d k⟩
⟨*director*, n⟩	⟨*the*, ∗n d k⟩
⟨*mistake*, n⟩	
⟨*admit*, V⟩	⟨*excuse*, V⟩
⟨*nom*, ∗V n p⟩	⟨*adj*, ∗V a⟩
⟨ϵ, ∗n ∗k N⟩	⟨*be*, ∗a v⟩
⟨*of*, ∗N ∗p n⟩	

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 8.

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 merge, move and delete operations introduced in section 2.1. Table 3 shows the additional lexical items needed to derive examples like these.

2.2 Conceptual Burden

The conceptual burden imposed by the theory presented here 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:sect. 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 (Grodner et al. 2003). 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 & 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, Santorini & Marcinkiewicz 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, Goldwater & Johnson (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 & Simon 1972; Gigerenzer, Todd & the ABC Research Group 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:sect. 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.[10] 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:

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

(9) Canonical Realization
Surface subjects preferentially are underlying subjects as well.

These parsing preferences reflect certain leading ideas in the literature. MaxElide was first proposed by Merchant (2008b) (first circulated in 2001) and has been expanded on by Takahashi & 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 & Bever 2001:sect. 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 echoes Stevenson & Smolensky's (2006:315) application of Case hierarchies in the parser.

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 in (10).

(10) MaxElide elision hierarchy: v > voice > V

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,

[10] 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 (2011) takes up these issues in greater depth.

Canonical Realization favors surface structures 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 (6) and (7), we conducted an acceptability study using magnitude estimation (Bard, Robertson & Sorace 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 4a shows how the design embeds these two ways of mismatching into both the experimental ("Ellipsis") and control conditions ("No ellipsis"), exemplified in (11)–(14). 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, Grosz & Gilliom 1993). The No Ellipsis control conditions served as a baseline for their corresponding Ellipsis conditions, enabling us to more accurately measure the change in acceptability due to ellipsis.[11]

Table 4. Experiment 1

(a) Design of experiment

		Match	**Mismatch**
Voice	No ellipsis	(11a) (11b)	(12a) (12b)
	Ellipsis	(11a) (11b)	(12a) (12b)
Category	No ellipsis	(13)	(14a) (14b)
	Ellipsis	(13)	(14a) (14b)

(b) Observed acceptability cline, Ellipsis conditions

Condition	**Mean log acceptability**	**Example**
Active–Active	0.235	(11a)
Passive–Passive	−0.285	(11b)
Passive–Active	−0.616	(12b)
Noun–VP	−0.690	(14a)
Active–Passive	−0.697	(12a)
Adjective–VP	−0.981	(14b)

[11] Note that the active–active Voice match and the Category match conditions ((11a) and (13) in Table 4a) are identical in structure. Separate items were constructed for each of these conditions in order to create minimal contrasts with the corresponding items in the mismatching conditions.

(11) Voice match
  a. Jill betrayed Abby, and Matt did ~~betray Abby~~, too. active–active
  b. Abby was betrayed by Jill, and Matt was ~~betrayed by Jill~~, too. passive–passive

(12) Voice mismatch
  a. Jill betrayed Abby, and Matt was ~~betrayed by Jill~~, too. active–passive
  b. Abby was betrayed by Jill, and Matt did ~~betray Abby~~, too. passive–active

(13) Category match
  a. Everyone criticized Roy, but Kate didn't ~~criticize Roy~~. VP–VP
  b. The report criticized Roy, but Kate didn't ~~criticize Roy~~. VP–VP

(14) Category mismatch
  a. The criticism of Roy was harsh, but Kate didn't ~~criticize Roy~~. noun–VP
  b. The report was critical of Roy, but Kate didn't ~~criticize Roy~~. adjective–VP

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. The Voice and Category mismatches, and their corresponding matched conditions, were treated essentially as two separate experiments. We present them together here because they represent instances of similar phenomena for our proposal; however, pretheoretically, Voice and Category mismatches may not underlyingly be the same phenomenon.[12] To avoid the risk of building features of our proposal into the experimental conditions, we did not assume a priori that the Voice and Category data sets would show the same contrasts, and instead treated the two mismatch types as separate subexperiments. 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.

The proposed parser heuristics predict certain contrasts in the data. First, MaxElide favors elliptical sentences where configurationally-higher projections are elided; in our items, therefore, MaxElide predicts that Mismatch conditions should have degraded acceptability relative to their corresponding Match conditions. Furthermore, given that MaxElide is, by hypothesis, about ellipsis, this degradation should be limited to sentences containing ellipsis. Second, Canonical Realization predicts that sentences with certain word orders—the "canonical" ones, perhaps ultimately the frequent ones—will be more acceptable than corresponding sentences with

[12] For example, on certain analyses, either Voice or Category alternations could be lexical, not syntactic alternations.

noncanonical order. Nothing about how Canonical Realization is formulated is specific to ellipsis, and as such, any effect of canonical order should be observed in both Ellipsis and No Ellipsis conditions.[13]

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.[14] 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, and a sentence 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, Robertson & Sorace 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, whereas a negative value means an item was rated lower than the modulus. These values are meaningless as absolute acceptability values. However, 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

[13] In the current experiment, only the Voice mismatch items vary with respect to canonical word order (Active versus Passive). In the Category mismatch items, both Mismatch and Match conditions have canonical order; experiments 2 and 3 directly address effects of Canonical Realization in sentences containing category mismatches.

[14] We presented stimuli to human participants using PsyScope X (Bonatti 2008).

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.

3.3 Results

Log scores were fit to a linear mixed-effects regression model (Gelman & Hill 2007, Baayen 2008) with Subject, Item, and Trial as random effects.[15] The data was divided into two subsets, representing the Voice mismatches and the Category mismatches, and a separate model was fit for each subset.[16] 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 Ellipsis by Match interaction was also included in the model. Tables 5 and 6 show the estimates of the coefficients for these fixed effects.

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

Table 5. 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

Table 6. 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

[15] We fitted the mixed-effects models using the `lme 4` library of the "R" system for statistical computing (Bates & Sarkar 2007). All predictor variables were contrast coded.

[16] 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.

less acceptable than their counterparts without ellipsis. There was, however, a significant Mismatch by Ellipsis interaction ($p < 0.001$): mismatch had a negative impact on acceptability judgments when there was ellipsis in the second conjunct ($t = -9.32$, $p < 0.001$), but not when there was no ellipsis ($t = -0.55$, $p > 0.5$), confirming that the condition on matching conjuncts is specific to cases of ellipsis. Condition means are plotted in Figure 12. The conditions represented in these graphs collapse over subcategories of Voice Match (see (11)), Voice Mismatch (see (12)), and Category Mismatch (see (14)).

Within Category mismatches, adjectival antecedents (14b) were rated worse than nominal antecedents (14a) ($t = 3.49$, $p < .001$). Table 4b summarizes the acceptability cline observed in this study in the Ellipsis conditions.

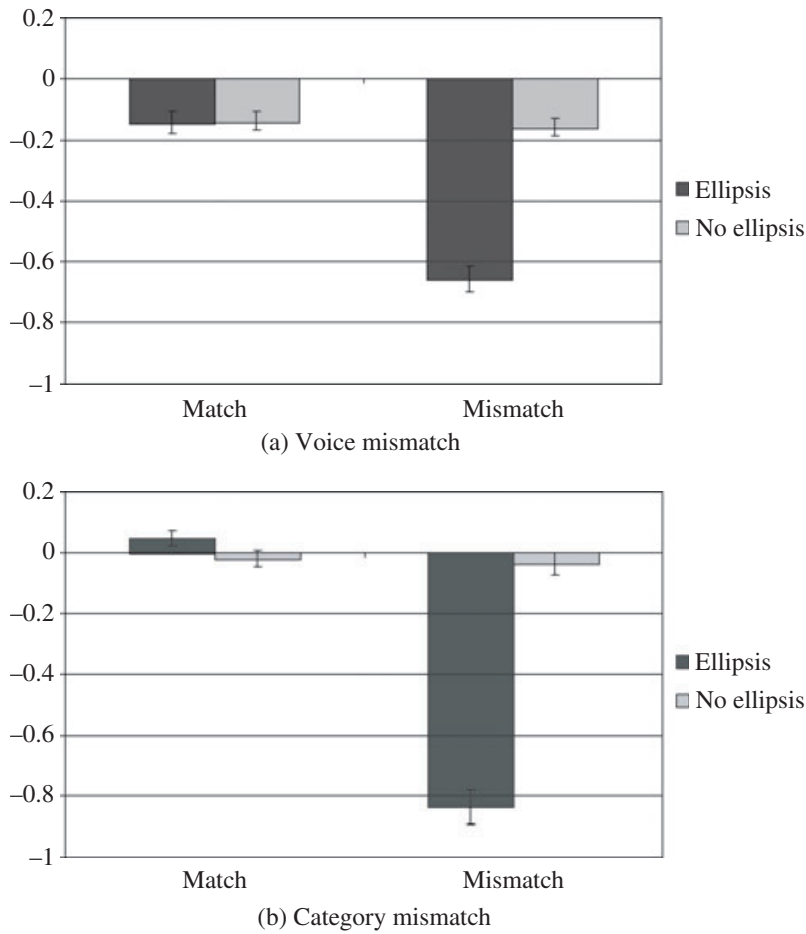


Figure 12. Experiment 1 condition means (error bars represent standard error for each condition)

Table 7. Experiment 1, Voice mismatches: Estimates of fixed effects, model including Antecedent type predictor

	Coefficient	Standard error	*t*	*p*
Intercept	−0.27	0.036	−7.40	<0.001
Ellipsis	−0.12	0.016	−7.30	<0.001
Mismatch	−0.13	0.016	−8.17	<0.001
ActiveAntecedent	0.046	0.016	2.83	<0.005
Ellipsis:Mismatch	−0.12	0.016	−7.48	<0.001
Ellipsis:ActiveAntecedent	0.005	0.016	0.29	*n.s.*
Mismatch:ActiveAntecedent	−0.05	0.016	−3.20	<0.005
Ellipsis:Mismatch:ActiveAnt	−0.02	0.016	−1.34	*n.s.*

For Voice mismatches, we also fit a regression model that included Antecedent type (Active or Passive) as a predictor in addition to Ellipsis and Mismatch; all two-way and three-way interaction terms were included. Table 7 shows the estimates of the model coefficients. In addition to the effects observed in the smaller model in Table 5, there is a main effect of Antecedent type, such that sentences with Active antecedents are judged more acceptable than those with Passive antecedents ($t = 2.38$, $p < 0.05$). This effect is carried by the Match conditions, as revealed by an Antecedent type by Mismatch interaction: when there is Match between the two clauses in a sentence, Active antecedents are judged more acceptable than Passive antecedents ($t = 5.17$, $p < 0.001$), but Mismatch sentences with Active and Passive antecedents did not differ ($t = -0.51$, $p > 0.5$). In other words, active–active matches are better than passive–passive matches.

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:

(15) {AA,PP} > {PA,AP}
Mismatch imposes an acceptability penalty in VPE.

(16) 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 earlier 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 V in violation of MaxElide. By contrast,

matching controls can delete at the configurationally higher 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 & Clifton 2005). If there is such a preference, it is not necessary that violating it result in unacceptability; a violation of MaxElide could, 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.[17]

3.4.2 *Category mismatches*

The lexical decomposition analysis presented in section 2.1.2 exposes a distributional difference between the adjective–VP stimuli and the noun–VP stimuli. Table 8 shows that, in the noun–VP stimuli, the argument of the predicate {criticize, expect, approve, distrust, intend, contribute,…} is often sequestered in a prepositional phrase. By contrast, the majority of the arguments in Table 8b are in a precopular position. Only in the subset of transitive-*of* stimuli does the element in the subjectlike 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 (17),

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

the argument of the root (e.g., √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 noncanonically realized example (18),

(18) 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., √excuse). The vast majority of the adjective–VP items in experiment 1 attested the noncanonical order. Generating one order or the other depends on specific syntactic steps summarized in Table 9. In this table, the ⇝ symbol indicates rewriting from child subderivations 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

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

Table 8. Canonical ordering in category mismatch stimuli

(a) Majority canonical noun–VP stimuli

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.

(b) Minority canonical adjective–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.

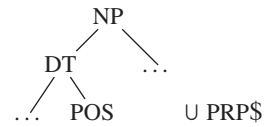
Table 9. The specific syntactic steps that determine compliance with Canonical Realization

	adjectival	**nominal**
Canonical	[critical of Matt]+[report]⇝AP	[an]+[admission of guilt]⇝DP
Noncanonical	[-able]+[excuse the mistake]⇝AP	movement⇝[$_{DP}$guilt's admission]

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 9. However, the effect of Canonical Realization can be seen in corpus attestation frequencies. In the Brown corpus (Kučera & Francis 1967; Marcus, Santorini & Marcinkiewicz 1993), for

Table 10. Attestation rates of genitive vs. simple NPs in the Brown corpus

Schema	Attestations in Brown
	41,032
	2004 + 8802 = 10,806

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 10 shows these counts. These Canonical Realization asymmetries would derive the observed processing difficulty differences in a frequency-sensitive comprehension model (Jurafsky 1996; Crocker & Brants 2000; Hale 2006, 2011).

Given that 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 measuring its acceptability in these new constructions. Note that, unlike MaxElide, Canonical Realization is a general heuristic that is expected to apply to the processing of any clause, whether or not it contains ellipsis. As such, Canonical Realization predicts that sentences with noncanonical argument order will be degraded across the board relative to their canonical order counterparts. Specifically, this penalty for noncanonical order should appear in experiments 2 and 3 as a main effect of canonicality. Canonicality is predicted to contribute to sentence acceptability over and above the main effects of mismatch and ellipsis, and the mismatch-ellipsis interaction observed in experiment 1.

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 11, 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 (19)–(22)), and half represented noncanonical order (see the (b) examples in (19)–(22)). Both experiments contained a total of 32 experimental items.

Table 11. Experiments 2 and 3

Experiment 2—Adjectival			
		Match	**Mismatch**
No ellipsis	Canonical Noncanonical	As below but with overt second conjunct	
Ellipsis	Canonical	(20a)	(19a)
	Noncanonical	(20b)	(19b)
Experiment 3—Nominal			
		Match	**Mismatch**
No ellipsis	Canonical Noncanonical	As below but with overt second conjunct	
Ellipsis	Canonical	(22a)	(21b)
	Noncanonical	(22b)	(21b)

(19) Mismatching adjectival antecedents
   a. The parents were critical of the uniforms, but the cheerleaders didn't ~~criticize the uniforms~~. — of Theme
   b. The boy's exhaustion was understandable, and the coach did ~~understand the boy's exhaustion~~. — -able

(20) Matching VP antecedents
   a. Paul deciphered the riddle, but Kevin didn't ~~decipher the riddle~~. — Active
   b. The wedding party was accommodated by the hotel, but the guests' pets weren't ~~accommodated by the hotel~~. — Passive

(21) Mismatching nominal antecedents
   a. The landing of the plane was unplanned, but the pilot did ~~land the plane~~. — of Theme
   b. The plane's landing was unplanned, but the pilot did ~~land the plane~~. — ThemePoss

(22) Matching VP antecedents
   a. The experienced pilot landed his plane during the storm, but the novice didn't. — Active
   b. The plane was landed by an experienced pilot during the storm, but a nearby jet wasn't. — Passive

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.[18]

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 Noncanonical). 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 ($p < 0.05$): noncanonical 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 ($t = -2.74$, $p < 0.01$), but not for sentences without ellipsis ($t = 1.01$, $p > 0.3$).

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 12 gives the estimates of the coefficients in the Adjectivization regression. Condition means are plotted in Figure 13.

[18] 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 and contrasts on the other end being reduced.

Table 12. 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.*

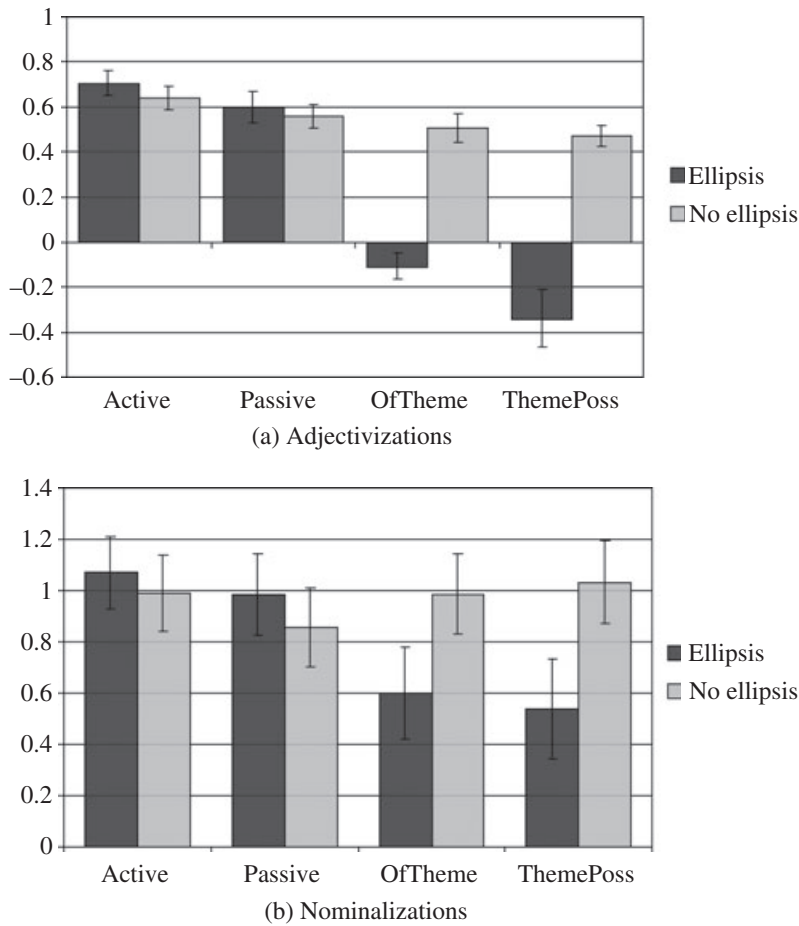


Figure 13. Condition means for experiments 2 and 3 (error bars represent standard error for each condition)

Table 13. 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.*

4.2.2 *Experiment 3: Nominalizations*

The Nominalization regression revealed the same pattern of fixed effects as the Adjectival regression. There were main effects of 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 ($t = -8.85$, $p < 0.001$), but not for sentences without ellipsis ($t = 0.93$, $p > 0.3$). 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 13 gives the estimates of the coefficients for the Nominalization regression. The condition means are plotted in Figure 13b.

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.

Additionally, the coefficient on the Canonicality predictor, in the fitted regression models for experiments 2 and 3, is significantly different from zero (Tables 12 and 13). 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 12 and 13). 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 (Frazier & Clifton 2006, Carlson 2001, Chambers & Smyth 1998); in particular, we offer a partial answer to the question of what linguistic environments induce the strongest parallelism effects.[19]

5. General Discussion

The results of these three experiments confirm a cline of acceptability levels across different sorts of VPE. This cline would follow naturally under the proposals made in section 2. The derivation assumes that (a) parsability affects the acceptability judgment task, (b) parsing preferences such as MaxElide and Canonical Realization guide the search for syntactic structure within the space of grammatical derivations, and (c) violation of such parsing preferences degrades sentence parsability. This section develops point (b) by explaining one formal notion of sentence parsability. The explanation highlights the compatibility of the present proposal with a standard conceptualization of the parsing process.

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*.[20] 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 reexamined a second time. It is complemented by another data structure, the *agenda*, that keeps track of as-yet-unexplored possibilities. Kay (1986:68) 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.

In natural language processing, the priorities associated with agenda entries are typically probabilities (Stolcke 1995, Roark 2001, Klein & Manning 2004), although this need not be the case (Caraballo & Charniak 1998, Dzikovska et al. 2005). One can equally apply heuristic preferences like MaxElide and Canonical Realization to order the agenda such that subtrees consistent with them are explored before subtrees

[19] A related body of work investigates effects of discourse coherence relations on a number of linguistic phenomena, including ellipsis resolution (Kehler 2000, 2002; Kertz 2008; Kim & Runner 2009). The experiments presented here do not address the role of discourse coherence—all items used connectives that indicated a Resemblance relation, in Kehler's terms. However, our proposal can be straightforwardly extended to account for cases involving different coherence relations, which have been shown to modulate how strictly the structural identity constraint on VP ellipsis is enforced (Kim & Runner 2009, to appear).

[20] Chart parsing is described in all computational linguistics textbooks. See, for instance, Gazdar & Mellish 1989:chap. 5, Allen 1995:chap. 3, or Jurafsky & Martin 2009:sect. 13.4.3.

that are inconsistent with them. Ordering an agenda in this way encourages search in a certain direction—a "soft" prioritization. As section 2.3 points out, realistic grammars set up particularly difficult parsing problems because there are so many combinations to sift through. Even with just the handful of lexical entries shown in Figures 2 and 3, the six words of a conjunct like *an admission of guilt was needed* lead to 295 chart entries on an exhaustive, bottom-up parsing strategy unaided by any heuristics. These charts encode attempts to analyze subsequences of the given morphemes in ways that are locally consistent with the grammar. The granularity is at the level of individual applications of *merge* and *move*. Winnowing of these structural combinations must proceed largely unconsciously in unambiguous sentences. We join Kay (1986) in supposing that this search goes on longer in sentences that thwart our heuristic expectations. It is this heightened amount of computational work 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:

1. Determining whether something has been elided
2. Locating the antecedent
3. 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, which 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 & Stock (1990) would immediately account for the experimental results obtained by Frazier, Munn & Clifton (2000). In virtue of using two charts, this method is applicable to both intra sentential and cross-sentential ellipsis. Indeed, Kim & Runner (2009) have demonstrated a similar acceptability cline in cases of cross-sentence VPE.

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 if it 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:chap. 2).

5.3 Implications as Regards the Organization of Grammar

Much work in psycholinguistics has been devoted to discovering significant factors in language processing. Although it may at first glance seem to be in competition

with these, the syntactic account sketched in section 2 does not in fact preclude the influence of other factors. Rather it invites the combination of syntactic and nonsyntactic factors in detailed theories of the comprehension process. Conceiving, with Kay (1986), of parsing as a search problem leads to a straightforward view of intermediate parser states as rated by goodness or badness. Lexical idiosyncrasy, end-weight, discourse relation, definiteness, event typicality, prefix probability—all of these factors might well be traded off against syntactic factors in a parser that does informed search. These factors need not be part of grammar per se, but rather can be viewed as guiding *use* of the grammar. They may contribute numerical penalties without any entailed commitment to grammar itself as "gradient," "numerical," or "probabilistic." Attributing a preference or dispreference in this way to the processing mechanism can be contrasted with proposing a repair rule. Whereas a repair rule is a recipe for destructively changing a representation, a parsing preference simply judges which representation seems more likely to lead to the correct parse.

5.4 Functional Motivation

5.4.1 *Origins of these heuristics*

By way of interpretation, we suggest that the 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 2011).

5.4.2 *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 echoes Fiengo and May's (1994:192) 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.

This sort of pressure to avoid redundancy is also attested in the "repeated name penalty" (Gordon, Grosz & Gilliom 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 & 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 (2011) details, an appropriate set of heuristics may direct a parser to impose an information-theoretic prior during inference—that is, not to expect redundancy.

5.5 Extensions

The lexical decomposition approach of section 2.1.2 extends naturally to other constructions. One such is the verbal gerunds studied by Arregui et al. (2006) and exemplified by (23).[21]

(23) 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 Table 14, 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 (23) as a morphological variant of the V + *nom* construction discussed in that section. In other words, *avoid* + *nom* could be understood in greater detail as the structure $[_{N^0}[_{V^0}\ avoid][_{N^0}\ ing]]$.

Table 14. Abney's analysis of gerundival NPs

Type	Attachment point
Acc-*ing*	IP
Poss-*ing*	VP
ing + *of*	V^0

[21] 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.

Considering the other two cases leads to novel predictions. An updated version (Schueler 2004) of Abney's analysis attaches *-ing* at vP in the so-called Poss-*ing* construction, and at the perfective aspect head in Acc-*ing* constructions.[22] This predicts that both possessive and accusative *-ing* gerunds may occur grammatically in voice mismatch VPE. Two such examples are presented in (24a) and (24b).

(24) 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 22, 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 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 that 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 on 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 extragrammaticality 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. Although current examples of voice mismatches in sluicing are quite unacceptable, it may prove to be that these are unacceptable for independent reasons. Indeed, it took careful work to reveal the existence of quite acceptable voice mismatches in VPE (Hardt 1993). Adjudication between these alternatives awaits future work.

[22] Note that the *-ing* in a Poss-*ing* construction really does attach higher than the voice head.

(i) Your asking Bill out is extremely unlikely.
(ii) His being asked out at all seems wildly implausible.

References

Abney, S. P. 1987. The English noun phrase in its sentential aspect. Ph.D. dissertation. MIT, Cambridge, MA.

Allen, J. 1995. *Natural language understanding*. Redwood City, CA: Benjamin/Cummings.

Arregui, A., C. Clifton, Jr., L. Frazier & K. Moulton. 2006. Processing elided verb phrases with flawed antecedents: The recycling hypothesis. *Journal of Memory and Language* 55:232–246.

Baayen, R. H. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

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

Bard, E. G., D. Robertson & A. Sorace. 1996. Magnitude estimation of linguistic acceptability. *Language* 72:32–68.

Bates, D. & D. Sarkar. 2007. *lme4: Linear mixed-effects models using S4 classes*. R package version 0.99875.9.

Berwick, R. C. & A. S. Weinberg. 1984. *The grammatical basis of linguistic performance*. Cambridge, MA: MIT Press.

Bever, T. G. 1970. The cognitive basis for linguistic structures. In *Cognition and the development of language*, ed. J. Hayes, 279–362. New York: Wiley.

Bonatti, L. 2008. PsyScope X, Build 53. PsyScope X Project. SISSA Language, Cognition, and Development Lab.

Caraballo, S. & E. Charniak. 1998. New figures of merit for best-first probabilistic chart parsing. *Computational Linguistics* 24:275–298.

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

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

Charniak, E., S. Goldwater & M. Johnson. 1998. Edge-based best-first chart parsing. In *Proceedings of the sixth Workshop on Very Large Corpora*, ed. E. Charniak, 127–33. New Brunswick, NJ: Association for Computational Linguistics.

Chomsky, N. 1964. *Current issues in linguistic theory*. The Hague: Mouton.

Chomsky, N. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. R. Martin, D. Michaels & J. Uriagereka, 89–155. Cambridge, MA: MIT Press.

Chung, S. 2006. Sluicing and the lexicon: The point of no return. In *BLS 31: General session on parasession on prosodic variation and change*, ed. R. T. Corver & Y. Kim, 73–91. Berkeley, CA: Berkeley Linguistics Society.

Citko, B. 2005. On the nature of Merge: External Merge, Internal Merge, and Parallel Merge. *Linguistic Inquiry* 36:475–497.

Collins, C. 2002. Eliminating labels. In *Derivation and explanation in the Minimalist Program*, ed. S. D. Epstein & T. D. Seely, 42–64. Oxford: Blackwell.

Crocker, M. W. & T. Brants. 2000. Wide-coverage probabilistic sentence processing. *Journal of Psycholinguistic Research* 29:647–669.

Dalrymple, M., S. M. Shieber & F. C. Pereira. 1991. Ellipsis and higher-order unification. *Linguistics and Philosophy* 14:399–452.

de Vries, M. 2009. On multidominance and linearization. *Biolinguistics* 3:344–403.

Dzikovska, M. O., M. D. Swift, J. F. Allen & W. de Beaumont. 2005. Generic parsing for multi-domain semantic interpretation. In *Proceedings of the International Workshop on Parsing Technologies*, 196–197. New Brunswick, NJ: Association for Computational Linguistics.

Fiengo, R. & R. May. 1994. *Indices and identity*. Cambridge, MA: MIT Press.

Fox, D. 2000. *Economy and semantic interpretation*. Cambridge, MA: MIT Press.

Fox, D. 2002. Antecedent-contained deletion and the copy-theory of movement. *Linguistic Inquiry* 33:63–96.
Frazier, L. 1979. On comprehending sentences: syntactic parsing strategies. Ph.D. dissertation, University of Connecticut, Storrs.
Frazier, L. 2008. Processing ellipsis: A processing solution to the undergeneration problem? In *Proceedings of the 26th West Coast Conference on Formal Linguistics*, ed. C. B. Chang & H. J. Haynie, 21–32. Somerville, MA: Cascadilla Proceedings Project.
Frazier, L. & C. Clifton, Jr. 2005. The syntax-discourse divide: Processing ellipsis. *Syntax* 8:421–430.
Frazier, L. & C. Clifton, Jr. 2006. Ellipsis and discourse coherence. *Linguistics and Philosophy* 29:315–346.
Frazier, L. & J. D. Fodor. 1978. The sausage machine: A new two-stage parsing model. *Cognition* 6:291–325.
Frazier, L., A. Munn & C. Clifton. 2000. Processing coordinate structures. *Journal of Psycholinguistic Research* 29:421–430.
Gärtner, H.-M. 2002. *Generalized transformations and beyond: Reflections on minimalist syntax*. Berlin: Akademie Verlag.
Gazdar, G. & C. Mellish. 1989. *Natural language processing in Lisp: An introduction to computational linguistics*. Workingham England: Addison Wesley.
Gelman, A. & J. Hill. 2007. *Data analysis using regression and multilevel/hierarchical models*. Cambridge: Cambridge University Press.
Gigerenzer, G., P. M. Todd & the ABC Research Group. 1999. *Simple heuristics that make us smart*. Oxford: Oxford University Press.
Gordon, P. C., B. J. Grosz & L.A. Gilliom. 1993. Pronouns, names, and the centering of attention in discourse. *Cognitive Science* 17:311–347.
Grodner, D., E. Gibson, V. Argaman & M. Babyonyshev. 2003. Against repair-based reanalysis in sentence comprehension. *Journal of Psycholinguistic Research* 32:141–166.
Hale, J. 2001. 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*, 159–166. New Brunswick, NJ: Association for Computational Linguistics.
Hale, J. 2006. Uncertainty about the rest of the sentence. *Cognitive Science* 30:609–642.
Hale, J. 2011. What a rational parser would do. *Cognitive Science* 35:399–443.
Hardt, D. 1993. Verb phrase ellipsis: Form, meaning, and processing. Ph.D. dissertation, University of Pennsylvania, Philadelphia.
Hardt, D. 1997. An empirical approach to VP Ellipsis. *Computational Linguistics* 23:525–541.
Hardt, D. 2004. Ellipsis resolution and inference. *Acta Linguistica Hafniensia* 36:65–77.
Harkema, H. 2001. Parsing minimalist languages. Ph.D. dissertation, University of California, Los Angeles.
Hay, J. 2003. *Causes and consequences of word structure*. New York: Routledge.
Joshi, A. K. 1985. Tree adjoining grammars: How much context-sensitivity is required to provide reasonable structural descriptions? In *Natural language parsing: Psychological, computational, and theoretical perspectives*, ed. D. Dowty, L. Karttunen & A. M. Zwicky, 206–250. Cambridge: Cambridge University Press.
Joshi, A. K. 2004. Starting with complex primitives pays off: Complicate locally, simplify globally. *Cognitive Science* 28:637–668. (2003 Rumelhart Prize special issue honoring Aravind K. Joshi.)
Jurafsky, D. 1996. A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science* 20:137–194.
Jurafsky, D. & J. H. Martin. 2009. *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*, 2nd ed. Upper Saddle River, NJ: Pearson Education.
Kaplan, R. M. 1972. Augmented transition networks as psychological models of sentence comprehension. *Artificial Intelligence* 3:77–100.

Kaplan, R. M. 1973. A general syntactic processor. In *Natural language processing*, ed. R. Rustin, 195–241. New York: Algorithmics Press.
Kay, M. 1986. Algorithm schemata and data structures in syntactic processing. In *Readings in natural language processing*, ed. B. J. Grosz, K. S. Jones & B. L. Webber, 35–70. Los Altos, CA: Morgan Kaufmann.
Kayne, R. S. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
Keenan, E. 1980. Passive is phrasal (not sentential or lexical). In *Lexical grammar (Publications in Language Sciences 3)*, ed. T. Hoekstra, H. van der Hulst & M. Moortgat, pp. 181–214. Dordrecht: Foris.
Kehler, A. 2000. Coherence and the resolution of ellipsis. *Linguistics and Philosophy* 23:533–575.
Kehler, A. 2002. *Coherence, reference, and the theory of grammar*. Stanford, CA: CSLI Publications.
Kennedy, C. 2003. Ellipsis and syntactic representation. In *The interfaces: Deriving and interpreting omitted structures (Linguistik Aktuell/Linguistics Today 61)*, ed. K. Schwabe & S. Winkler, 29–53. Amsterdam: John Benjamins.
Kertz, L. 2008. Focus structure and acceptability in verb phrase ellipsis. In *Proceedings of the 27th West Coast Conference on Formal Linguistics*, ed. N. Abner & J. Bishop, 283–291. Somerville, MA: Cascadilla Proceedings Project.
Kim, C. S. & J. T. Runner. 2009. Strict identity, coherence, and parallelism in VP ellipsis. In *Proceedings of Semantics and Linguistic Theory 19*, ed. S. Ito & E. Cormany, 275–287. Columbus, OH: CLC Publishing.
Kim, C. S. & J. T. Runner. To appear. Discourse parallelism and VP ellipsis. In *Recent work on ellipsis (UMass Occasional Papers in Linguistics 38)*, ed. M. Grant & J. A. Harris. Amherst, MA: GLSA Publications.
Klein, D. & C. D. Manning. 2004. Parsing and hypergraphs. In *New developments in parsing technology*, ed. H. Bunt, J. Carroll & G. Satta, pp. 351–372. Dordrecht: Kluwer.
Kobele, G. M. 2006. *Generating copies: An investigation into structural identity in language and grammar*. Ph.D. dissertation, University of California, Los Angeles.
Kobele, G. M. 2009. Syntactic identity in survive minimalism: Ellipsis and the derivational identity hypothesis. In *Towards a purely derivational syntax: Survive-minimalism*, ed. M. T. Putnam, 195–230. Amsterdam: John Benjamins.
Kobele, G. M. 2010. A formal foundation for A and A-bar movement in the Minimalist Program. In *MOL 10/11 (Lecture Notes in Computer Science 6149)*, ed. C. Ebert, G. Jäger & J. Michaelis, 145–159. Berlin: Springer.
Kobele, G. M. 2011. Deriving reconstruction asymmetries. Ms., University of Chicago.
Kobele, G. M., C. Retoré & S. Salvati. 2007. An automata-theoretic approach to minimalism. In *Proceedings of the workshop Model-Theoretic Syntax at 10*, ed. J. Rogers & S. Kepser, pp. 73–82. http://cs.earlham.edu/esslli07mts/.
Koizumi, M. 1995. *Phrase structure in minimalist syntax*. Ph.D. dissertation, MIT, Cambridge, MA.
Koopman, H. & D. Sportiche. 1991. The position of subjects. *Lingua* 85:211–258.
Kracht, M. 2001. Syntax in chains. *Linguistics and Philosophy* 24:467–529.
Kučera, H. & W. N. Francis. 1967. *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
Lappin, S. 1999. An HPSG account of antecedent-contained ellipsis. In *Fragments: Studies in ellipsis and gapping*, ed. S. Lappin & E. Benmamoun, 68–97. Oxford: Oxford University Press.
Lappin, S. & H. Gregory. 1997. A computational model of ellipsis resolution. Talk presented at Formal Grammar 1997, Aix-en-Provence, France (August).
Lappin, S. & M. McCord. 1990. Anaphora resolution in slot grammar. *Computational Linguistics* 16:197–212.

Lappin, S. & H.-H. Shih. 1996. A generalized reconstruction algorithm for ellipsis resolution. In *Proceedings of COLING-96*, 687–692. Copenhagen: Center for Sprogteknologi.
Lasnik, H. 1995. Verbal morphology: Syntactic Structures meets the Minimalist Program. In *Evolution and revolution in linguistic theory: Essays in honor of Carlos Otero*, ed. P. Kempchinsky & H. Campos, 251–275. Washington, DC: Georgetown University Press.
Lasnik, H. 1999. Chains of arguments. In *Working minimalism (Current Studies in Linguistics 32)*, ed. S. D. Epstein & N. Hornstein, 189–215. Cambridge, MA: MIT Press.
Lavelli, A. & O. Stock. 1990. When something is missing: Ellipsis, coordination, and the chart. In *Proceedings of COLING-90*, ed. H. Karlgren, 184–189. Helsinki: University of Helsinki.
Lebeaux, D. 1988. *Language acquisition and the form of grammar*. Ph.D. dissertation. University of Massachusetts, Amherst.
Lees, R. B. 1960. *The grammar of English nominalizations*. Bloomington, IN: Indiana Research Center in Anthropology, Folklore, and Linguistics.
Levy, R. 2008. Expectation-based syntactic comprehension. *Cognition* 106:1126–1177.
Lewis, R. L. 2000. Specifying architectures for language processing: Process, control, and memory in parsing and interpretation. In *Architectures and mechanisms for language processing*, ed. M. W. Crocker, M. Pickering & C. Clifton, Jr., 56–89. Cambridge: Cambridge University Press.
Manzini, M. R. & A. Roussou. 2000. A minimalist theory of A-movement and control. *Lingua* 110:409–447.
Marcus, M. P., B. Santorini & M. A. Marcinkiewicz. 1993. Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics* 19:313–330.
Merchant, J. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis (Oxford Studies in Theoretical Linguistics 1)*. Oxford: Oxford University Press.
Merchant, J. 2007. Voice and ellipsis. Ms., University of Chicago.
Merchant, J. 2008a. An asymmetry in voice mismatches in VP-ellipsis and pseudogapping. *Linguistic Inquiry* 39:169–179.
Merchant, J. 2008b. Variable island repair under ellipsis. In *Topics in ellipsis*, ed. K. Johnson, 132–153. Oxford: Oxford University Press.
Michaelis, J. 2001. On formal properties of minimalist grammars. Ph.D. dissertation, Universität Potsdam.
Müller, G. 2010. On deriving CED effects from the PIC. *Linguistic Inquiry* 41:35–82.
Newell, A. & H. A. Simon. 1972. *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
Nielsen, L. A. 2004. Verb Phrase Ellipsis detection using automatically parsed text. In *Proceedings of COLING-04*, 1093–1099. New Brunswick, NJ: Association for Computational Linguistics.
Pereira, F. 1985. A new characterization of attachment preference. In *Natural language parsing: Psychological, computational, and theoretical perspectives (ACL Studies in Natural Language Processing)*, ed. D. Dowty, L. Karttunen & A. Zwicky, 307–319. Cambridge: Cambridge University Press.
Pesetsky, D. 1995. *Zero syntax: Experiencers and cascades*. Cambridge, MA: MIT Press.
Roark, B. 2001. Probabilistic top-down parsing and language modeling. *Computational Linguistics* 27:249–276.
Runner, J. T. 1995. Noun phrase licensing and interpretation. Ph.D. dissertation, University of Massachusetts, Amherst.
Runner, J. T. 2006. Lingering challenges to the raising-to-object and object-control constructions. *Syntax* 9:193–213.
Sag, I. A. 1976. Deletion and logical form. Ph.D. dissertation, MIT, Cambridge, MA.
Schueler, D. 2004. Extended-projection, categorization, and the English morpheme *-ing*. In *Proceedings of WECOL 33*, ed. M. T. Martínez, A. Alcázar & R. M. Hernández, 342–356. Fresno, CA: Department of Linguistics, California State University.

Schütze, C. T. 1996. *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Chicago: University of Chicago Press.
Stabler, E. 1997. Derivational minimalism. In *Logical aspects of computational linguistics*, ed. C. Retoré, 68–95. Berlin: Springer.
Starke, M. 2001. Move dissolves into Merge: A theory of locality. Ph.D. dissertation, University of Geneva.
Stevenson, S. & P. Smolensky. 2006. Optimality in sentence processing. In *The harmonic mind: From neural computation to optimality-theoretic grammar: Vol. 2, Linguistic and philosophical implications*, ed. P. Smolensky & G. Legendre, 307–338. Cambridge, MA: MIT Press.
Stolcke, A. 1995. An efficient probabilistic context-free parsing algorithm that computes prefix probabilities. *Computational Linguistics* 21:165–201.
Takahashi, S. & D. Fox 2005. MaxElide and the re-binding broblem. In *Proceedings of SALT 15*, ed. E. Georgala & J. Howell, 223–240. Ithaca, NY: CLC Publications.
Takahashi, S. & S. Hulsey. 2009. Wholesale late merger: Beyond the A/A′ distinction. *Linguistic Inquiry* 40:387–426.
Tanenhaus, M. K. & G. N. Carlson. 1990. Comprehension of deep and surface verb phrase anaphors. *Language and Cognitive Processes* 5:257–280.
Tomioka, S. 1997. Focusing effects and NP interpretation in VP ellipsis. Ph.D. dissertation, University of Massachusetts, Amherst.
Townsend, D. J. & T. G. Bever. 2001. *Sentence comprehension: The integration of habits and rules*. Cambridge, MA: MIT Press.
Zhang, N. 2004. Move is Remerge. *Language and Linguistics* 5:189–209.

Christina S. Kim
University of Rochester
Department of Brain and Cognitive Sciences
Department of Linguistics
503 Lattimore Hall
Rochester
NY 14627-0096
USA

ckim@bcs.rochester.edu

Gregory M. Kobele
University of Chicago
Computation Institute and
Department of Linguistics
5735 S. Ellis Ave.
Chicago
IL 60637
USA

kobele@uchicago.edu

Jeffrey T. Runner
University of Rochester
Department of Linguistics
503 Lattimore Hall
Rochester
NY 14627-0096
USA

runner@ling.rochester.edu

John T. Hale
Cornell University
Department of Linguistics
Morrill Hall, Room 203
Ithaca
NY 14853-4701
USA

jthale@cornell.edu