University of Pennsylvania Working Papers in Linguistics

Volume 17
Issue 1 *Proceedings of the 34th Annual Penn Linguistics Colloquium*

Article 16

2011

Why Unaccusatives Have it Easy: Reduced Relative Garden Path Effects and Verb Type

Kyle Grove
Cornell University

Why Unaccusatives Have it Easy: Reduced Relative Garden Path Effects and Verb Type

Abstract

This paper provides a new account for why unaccusative verbs are easier to process than unergative verbs in the reduced relative garden path construction, as demonstrated in Stevenson and Merlo [1997]. Reanalysis to the passivized reduced relative clause form requires the verb to be causative. Stevenson and Merlo [1997] argued that unaccusatives are causativized in the lexicon, while unergatives are causativized in the syntax. This account argues instead that an independently attested co-occurrence restriction contributes to greater initial ambiguity in the unergative case; causative unergatives require an argument/directional attachment of prepositional phrase [Hoekstra, 1988, Levin and Rappaport-Hovav, 1995, Folli and Harley, 2006].

We implement the unergative-PP co-occurrence restriction in Minimalist Grammars [Stabler, 1997]. We model the contribution of prepositional phrase ambiguity to unergative reduced relative ambiguity with Entropy Reduction [Hale, 2003]. We obtain greater Entropy Reductions for the unergative condition, modeling that human comprehenders are more taxed by compounded ambiguity.

Why Unaccusatives Have it Easy: Reduced Relative Garden Path Effects and Verb Type*

Kyle Wade Grove

1 Introduction

This paper provides a new account of the asymmetrical effect of argument structure on garden pathing difficulty. (1a) with unaccusative 'melted' is easier to comprehend than (1b) with unergative 'raced'.

(1) a. **Unaccusative:** The butter melted in the oven was lumpy.
b. **Unergative:** The horse raced past the barn fell. (Stevenson and Merlo 1997)

Verbal ambiguity contributes to the difficulty of such reduced relative clause (RRC) garden paths. When the embedded verb is obligatorily transitive (2a), RRC disambiguation is less difficult than in optionally transitive conditions (2b), because it renders the misleading active voice, intransitive analysis of the ambiguous substring untenable. (Pritchett 1992, MacDonald 1994)[1]

(2) a. The ruthless dictator captured in the coup was hated throughout the country.
b. The ruthless dictator fought in the coup was hated throughout the country. (MacDonald 1994)

Stevenson and Merlo (1997) appeal to the lexicon for an explanation of the asymmetry of (1) and (2), positing that causative *v* applies lexically for unaccusatives, but syntactically for unergatives. We argue instead that unaccusatives and unergatives are both made causative "in the syntax". We explain the extra difficulty of unergative causation by appealing to the causative-PP co-occurrence restriction noted in Hoekstra (1988), Levin and Rappaport-Hovav (1995), and Folli and Harley (2006): unergative causation requires an argument-attached prepositional phrase (PP).

(3) a. **The window broke.**
b. Pat **broke the window.**

(4) a. **The soldiers marched** *(to their tents.)*
b. The general **marched the soldiers** **(to the tents)*[2]. (Levin and Rappaport-Hovav 1995)

Hoekstra (1988) and Folli and Harley (2006) show that adjunct PPs do not license unergative causation. Spatial PPs occurring with unergative manner of motion verbs (such as 'float', below) are potentially ambiguous between an directional (argument) and a locative (adjunct) reading.

(5) The boat floated under the bridge. (Zubizarreta and Oh 2007:28)

This ambiguity is indeed syntactic; Hoekstra (1988) and Zubizarreta and Oh (2007) show that Dutch auxiliary selection is sensitive to this ambiguity. With intransitive unergatives in Dutch in perfect aspect, the zijn ('be') auxiliary forces the directional reading of the PP, the hebben ('have') auxiliary the atelic locative reading. Unergatives with argument attachment of PP pattern together with unaccusatives in being easily causativized.

In parsing unergative RRCs, the verbal ambiguity and the PP-attachment ambiguity both precede the disambiguating token. The co-occurrence restriction between argument-PPs and the causative forms of unergative verbs means that the parser encounters greater uncertainty in the unergative case, which we hypothesize explains the greater difficulty of unergative RRC:

*Many thanks due to John Hale, Julie Balazs, Effi Georgala, Ed Stabler, Michael Putnam, and Paola Merlo for their helpful comments and feedback. All the rest of it is my fault.

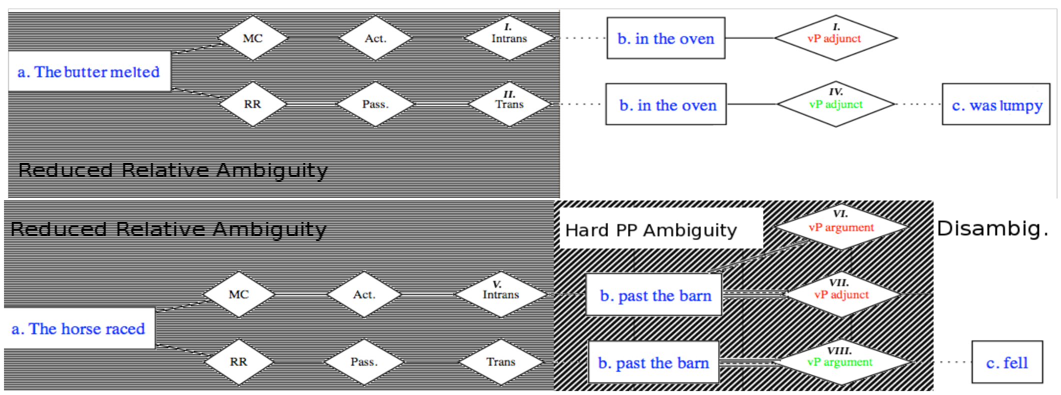
[1] We use the term 'reanalysis' in this paper to simply mean the effort required to successfully integrate the disambiguating token into the parse, and abstract away from whether the parsing architecture is serial, parallel, or somewhere in between.

[2] Bold and italics added for emphasis.

Hypothesis: PP-attachment ambiguity contributes to the greater difficulty of unergative RRC via a co-occurrence restriction between unergative causativization and argument-attachment of PP.

The account of the reduced relative asymmetry herein honors the Miller and Chomsky (1963) methodology in taking the form of an explicit parsing architecture whose components are a memory component, a control structure (parsing module), and the grammatical knowledge that the parser uses to inform decisions. We assume no special role of memory in this account, and abstract away from it henceforth. As a proxy for the control structure, we employ the Entropy Reduction Hypothesis (Hale 2003), which models disambiguation effort as the reduction of conditional entropy over a stochastic grammar. We argue that the grammatical knowledge used by the parser in the reduced relative task to be the knowledge of acceptable causative forms for unergatives and unaccusatives argued for in Levin and Rappaport-Hovav (1995), Hoekstra (1988), and Folli and Harley (2006). We formalize this grammatical co-occurrence restriction using Stabler (1997)'s mildly context sensitive Minimalist Grammars (MG) formalism. In our grammar, we assume a constructional view of the lexicon, adopting the Distributed Morphology (DM) framework for verbal argument structure.

On the co-occurrence account, unergative reduced relatives are more difficult to process because they compound the uncertainty of reduced relatives with the uncertainty of prepositional phrase attachment, resulting in a severe drop in entropy as the disambiguating token is processed. Both ambiguities must be resolved together. Unaccusative reduced relatives are the less complex case; the disambiguation of unaccusative RRC is independent of PP-attachment.



Diamonds:*Parser States* **Rectangles:***Sentence Segments (a). Reduced Relative Ambiguity (b). PP Attachment Ambiguity (c). Disambiguation Point*

Green: Possible Analysis Red: Rejected Analysis

Figure 1: Relationships Between Reduced Relative and PP-Ambiguity

2 A Parsing Metric: Entropy Reduction Measures Incremental Disambiguation Work

Psycholinguists are often interested in the temporary ambiguity found in temporally ordered language stimuli, including garden path sentences. The speech stream is inherently linearly ordered in real time. Psycholinguists exploit this with visual experiments which simulate the speech stream by imposing incremental visual presentation, either in self paced reading or eye-tracking experiments. In the garden path experiment, sentence comprehenders are prompted to abandon a preferred structural analysis because of its incongruity with an incoming word. Garden path sentences thus represent a segue from a multistable perceptual state to a disambiguated perceptual state as the linguistic signal is incrementally encountered.

Ambiguity is a type of uncertainty that language users have about sentences. The linguistic notion of ambiguity, however, does not possess sufficient grain size for employ in real-time human sentence processing; we would like a measure of uncertainty with finer grain than the number of trees congruent with the sentence prefix. We would also like to render the notion of preferred and dispreferred analyses in a formal metric. These preferences could be heuristic, but may also be frequency effects encountered in the acquisition process of the human sentence processor, including recency effects. Stochastic models of sentence processing possess fine grain size, accommodate preference, and are sensitive to experience. Thus, the measure of uncertainty in this paper is the Shannon (1948) entropy, which operationalizes the notion of ambiguity.

Hale (2003) argued that the cognitive load humans experience when comprehending difficult sentences can be modeled as greater decrease in entropy on the incremental structural analysis of the sentence, given some probabilistic grammar known to the sentence comprehender. The incremental difficulty in human sentence processing is modeled by decrease in the conditional entropy of the parse forest conditioned on the string prefix at each point in the sentence. Segues from highly ambiguous states to less ambiguous states mean that he human comprehender has done work, and gained information about the underlying structure of the incoming sentence. A particularly rapid segue of this kind means that the human sentence processor has been particularly taxed by the comprehension task.

3 Formal Background

3.1 Entropy

This paper suggests that the argument structure of a verb can play a role in reduced relative processing. We present a Derivational Theory of Lexical Complexity which uses independently attested linguistic evidence as the basis for a complexity metric. This complexity metric requires explicit formalization of both the grammar and the parser. As the focus of our investigation concerns the effect of argument structure on reduced relative processing, we require not only a formalization of the syntactic module, but also an independent formalization of the lexicon.

Entropy represents the amount of uncertainty about a random variable. For a discrete random variable X with possible outcomes $x_1, x_2, \ldots$, with probabilities of outcomes $p_{x_1}, p_{x_2}, \ldots$, the entropy $H(X)$ is equal to $-\sum_{x \in X} p_x \log_2 p_x$. A fair coin, for example, has an entropy of $-((0.5 log_2 0.5) + (0.5 log_2 0.5))$, i.e., 1.0 bits of entropy, since each of the two possible outcomes (Heads, Tails) has a 0.5 probability of occurring.

For the purposes of modeling linguistic difficulty, it is helpful to conceptualize three separate aspects of difficulty (ambiguity) which the entropy models. First, entropy models the fact that, *ceteris paribus*, variables with equiprobable outcomes are more difficult than cases where probability is skewed towards certain outcomes. This is seen in Figure 2, where the fair coin exhibits greater entropy (1.0 bits) than a loaded coin with a 0.75 probability of Heads (0.81 bits). The crooked gambler who employs this biased coin with this knowledge has an advantage of 0.19 bits of information over a naive mark who assumes the coin to be equiprobable. Second, entropy models the greater relative difficulty of decisions with more outcomes, when the outcomes are equiprobable. Figure 3 depicts a probabilistic grammar modeling a fair coin versus a probabilistic grammar modeling a fair die. Being fair, both the coin and the die have purely equiprobable outcomes, but the die by virtue of

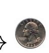
	Biased Coin	Fair Coin
PCFG	0.75 S →  0.25 S → 	0.5 S →  0.5 S → 
H	0.81 bits	1.0 bits

Figure 2: Entropy of Variables with Equiprobable vs. Biased Outcomes

	Fair Coin	Fair Die
PCFG	0.5 S →  0.5 S → 	$0.1\bar{6}$ S → ⚀ $0.1\bar{6}$ S → ⚁ $0.1\bar{6}$ S → ⚂ $0.1\bar{6}$ S → ⚃ $0.1\bar{6}$ S → ⚄ $0.1\bar{6}$ S → ⚅
H	1.0 bits	2.58 bits

Figure 3: Entropy of Variables with Few vs. Many Outcomes

	Fair Coin	Coin Game	
PCFG	0.5 S →  0.5 S → 	0.5 S → FFF 0.5 S → FF	0.5 F →  0.5 F → 
H	1.0 bits	2.5 bits	

Figure 4: Entropy of Simple Variables versus Hierarchical Random Processes

having more possible outcomes has greater entropy (2.58 bits). Finally, decisions with outcomes dependent on other decisions are more uncertain than simple random variables. Figure 4 compares the simple fair coin case to a more convoluted case, where, depending on the outcome of the first coin flip (designated below with start symbol S), the game player must flip (designated below with F) the coin either two or three times. This hierarchical random process has 2.5 bits of entropy, reflecting that the uncertainty of each coin flip is propagating up into the global ambiguity. Hierarchical uncertainty is of particular importance to psycholinguistics because it measures the 'depth' that broadly informs structural metrics in sentence processing.

The Entropy Reduction Hypothesis of Hale (2003) holds that difficult sentence processing tasks involve the rapid reduction of a large, entropic parse forest to a less ambiguous parse forest. A human incrementally reading through a garden path sentence maintains several parses of the sentence in parallel, through the locally ambiguous region; when the human encounters the disambiguating token, the human sentence processor must reject analyses which are suddenly no longer congruent with the sentence. This reduction in entropy is hypothesized to correspond to the disambiguation work performed.

3.2 Minimalist Grammars

The Entropy Reduction Hypothesis requires a formalized grammar. The Minimalist Grammars framework of Stabler (1997) allows the grammars written by linguists working within the Minimalist Program to be formalized for use in computational applications. The lexical items which compose Minimalist Grammars are triples of phonetic, syntactic and semantic features. The phonetic feature for any Minimalist lexical item is simply the string yield of that item, and can be null, as in the case of covert functional heads. The operations of Merge and Move manipulate phonetic features and are driven by syntactic features. The features governing Merge are of two kinds, Selected Categories and Selector Features. A lexical item whose leftmost Selector Feature matches the Selected Category of another lexical item Merges to create a derived lexical item whose phonetic feature is the concatenation of those of the children, and whose syntactic head is the Selected Category of the selector item.

The features which govern Move similarly consist of Licensor and Licensee features. An item with a licensor feature selects an item with the corresponding Licensee feature, to create a derived

Roots & PP-Attachment	**Passive**	**Reduced Relative Clause**
::=P =D SC	::=T C	the::⇒ agrD D
::=SC =$\checkmark_{do}$ V_{delta}	::=voice T	::=Crel +nom agrD
::=D =$\checkmark_{delta}$ V_{delta}	::=V_{do} T	::=T +arel Crel
::=$\checkmark_{do}$ =D V_{do}	the :: =N D -k;	::=pass +prel Crel
melted::$\checkmark_{delta}$	the :: =N D;	::=Opass T
walked::$\checkmark_{do}$	::=V_{delta} +k T	who::=N D -k -arel
fell::$\checkmark_{do}$	::=V_{delta} =D voice	which::=N D -k -arel
past::=D P	was::=V_{delta} +k Opass	::=N D -k -prel
in::=D P	were::=V_{delta} +k Opass	who::=N D -arel
D << P	::=V_{delta} +k pass	which::=N D -arel
$\checkmark_{do}$ << P	horse::N -nom	::=N D -prel
$\checkmark_{delta}$ << P	butter::N -nom	

Figure 5: MG of reduced relative clauses, argument structure

item whose phonetic feature is some concatenation of the childrens', and whose head is the category of the Licensor.

While Minimalist Grammars are not context free grammars, they share a useful property with the mildly context sensitive formalisms they are equivalent to: they possess a context-free backbone which can be weighted and estimated as a PCFG can. Parsers which compute the conditional entropy of mildly context sensitive formalism do exactly this, and treat the rules of Minimalist Grammars and other mildly context sensitive formalisms as context-free rules with more complex string yield functions than simple concatenation.

3.3 Distributed Morphology

Distributed Morphology (henceforth, DM; Halle and Marantz 1993) eliminates the lexicon as a realm of special operations. DM holds that the processes that underlie lexical composition are essentially syntactic processes, and that lexical entries are themselves derived. Although DM uses only the Minimalist derivational operations of Merge and Move, the DM ontology of representational units requires motivation. First, DM assumes that lexical items do not inherently possess category, but begin as prelexical derivational items known as Roots. DM eliminates appeal to lexical modularity by restricting idiosyncratic meaning to Roots which are otherwise syntactically impoverished. Roots are denoted with a check mark symbol, for example, $\checkmark BREAK$. While roots possess the sound-meaning pair element of lexical knowledge, verb valency and other morphosyntactic elements of lexical meaning are moderated by an economy of functional heads. As Roots do not inherently possess category, subcategorization and causation are implemented by a verbalizing element, *v* (or a nominalizing element *n*, etc.). By employing Distributed Morphology, we avoid special appeal to the Lexicon as a zone of causative application.

4 Methodology

We developed a Stablerian Minimalist Grammar to operationalize our hypothesis that co-occurrence restrictions on causation were responsible for the more severe garden pathing of unergatives reported in Stevenson and Merlo (1997). This Minimalist Grammar formalizes the co-occurrence restriction on the causative alternation of unergative verbs reported in Hoekstra (1988), Levin and Rappaport-Hovav (1995), Folli and Harley (2006) and Zubizarreta and Oh (2007). The correspondence between directed motion unergatives and unaccusatives was rendered with a Small Clause category (Hoekstra 1988, Folli and Harley 2006) which selects a Distributed Morphology-style root, while standard motion unergatives are treated as a simple root structure. The causative *v*-head selects the Small Clause, so that passives and causatives of unaccusatives and directed-motion unergatives are possible, but causatives and passives of unergatives with no small clause are not possible. We also formalized

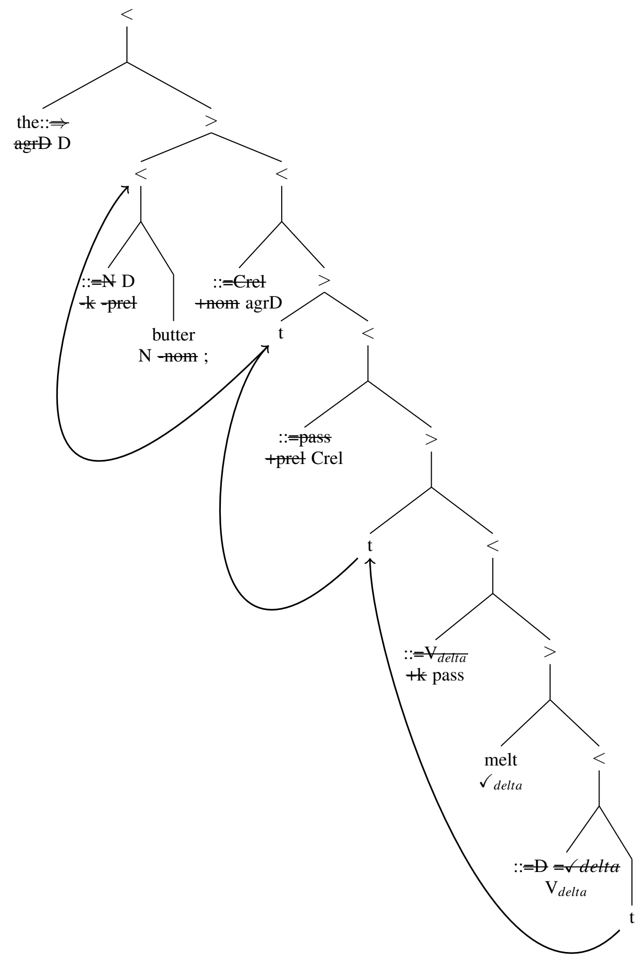


Figure 6: Derived Fragment for Unaccusative RRC 'The butter melted in the oven'

the promotion analysis (Kayne 1994) of relative clauses, with the reduced-relative relative clause co-occurrence with passivization. RRC are generated in our grammar via a covert relative pronoun which selects only a covert passive morpheme. This captures the distribution in English of the RRC construction.

We employed an MG parsing system which uses the Guillaumin (2004) compiler as a front end into an intermediate (MCFG) formalism. This system generates a probabilistic model by using a training and a testing phase. At training time, the parser parses a mini-corpus of full sentences, and maintains counts of how many times a particular rule was used. The parser uses the Weighted Relative Frequency estimation of Chi (1999) to estimate a backbone PCFG from these counts. The sentences in the training mini-corpus are representative of sentence forms which are possible disambiguated forms for the conditions the experimenter is interested in. For the task at hand, the training corpus presented each verb form in both a reduced relative frame and a main clause frame.

In this methodology, the testing mini-corpus contains the sentences whose conditional entropies the experimenter is interested in. The PCFG which was estimated at training is renormalized to fit the testing sample (rules which are used in training but not in testing are factored out, and other probabilities are adjusted similarly), and the parser computes from the training sample the right-

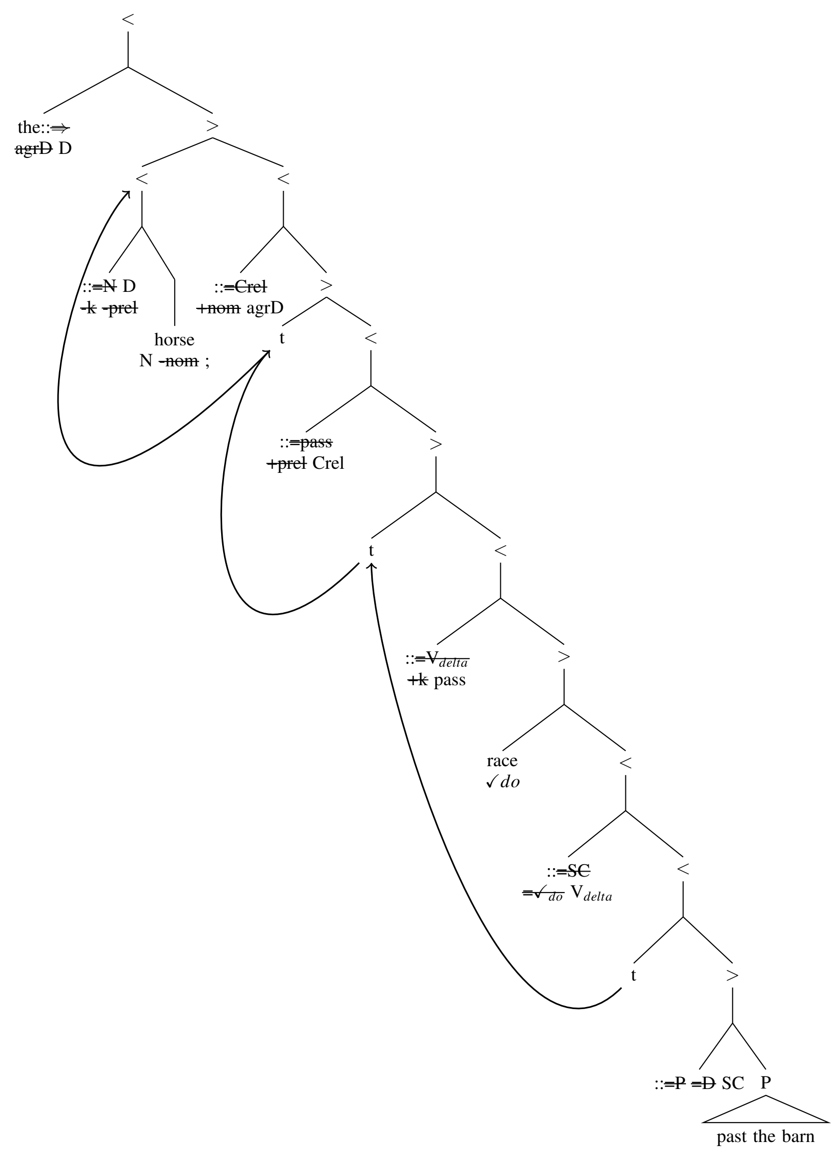


Figure 7: Derived Tree for Unergative RRC 'The horse raced past the barn'

hand side vector of local entropies and the fertility matrix ala Grenander (1967) for each sentence prefix.

A sample of 24 reduced relative garden path sentences (12 unaccusative; 12 unergative) from the Stevenson and Merlo (1997) study was obtained. From this sample, training and testing corpora were constructed. Sentences were controlled for length. The locally ambiguous segments of sentences, which consisted in each case of the sample as a reduced relative sentence fragment consisting of the sequence D NP V P NP, was only edited by simplifying, equiprobably balancing, and length-controlling NPs. The verb tokens of interest were not altered. The disambiguating segments were edited for length and to control for passivization, with 6 active voice and 6 passive voice disambiguations for each condition. We weighted the active voice disambiguations relative to reduced relative disambiguations using a parameter estimated from BNC (2007), reflecting the concept that the parser is uncertain between reduced relative disambiguations and main clause disambiguations but has knowledge regarding the much greater likelihood of main clauses. From this manipulated sample, a 48 sentence training corpus was created, where each sentence was represented twice: once with just the locally ambiguous segment as a main clause, and once with the complete garden path sentence. The 24 sentence sample of complete garden path sentences in the manipulated sample served as the testing mini-corpus.

5 Results

Table 2	*Entropy*$_{b.}$	*Entropy*$_{c.}$	*Entropy Reduction* $(E_{b.}\text{-}E_{c.})$
Unaccusatives	0.317	0.111	0.206
Unergatives	1.157	0.000	1.157

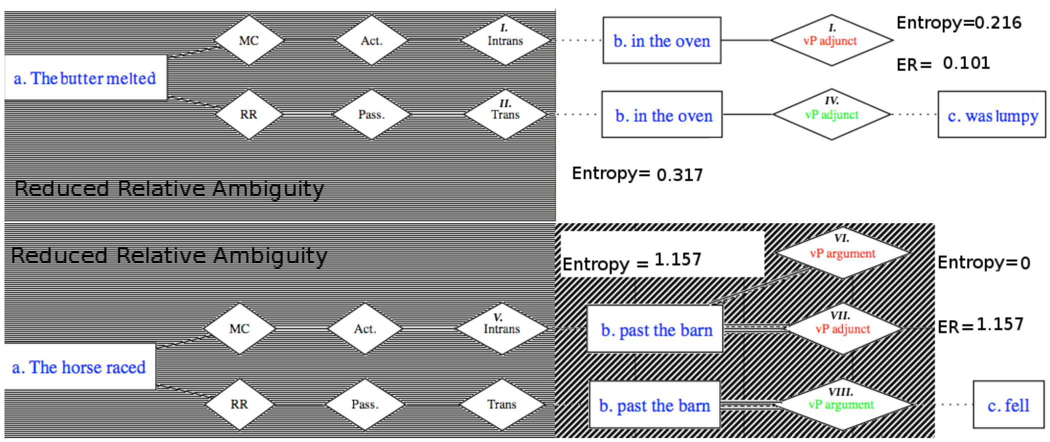


Figure 8: **Diamonds:***Parser States* **Rectangles:***Sentence Segments (a). Reduced Relative Ambiguity (b). PP Attachment Ambiguity (c). Disambiguation Point* Green: Possible Analysis Red: Rejected Analysis

The unergative condition elicits significantly greater reduction in entropy from the locally ambiguous segment to the disambiguated terminal, correctly deriving the result that unergative reduced relative processing is more difficult for human subjects.

5.1 Conclusion

Our approach attempts to be maximally parsimonious while according with several widely-held methodological assumptions regarding the modularity of the human sentence processor and the human language faculty. First, following Miller and Chomsky (1963), we employ a processing model where the set of parsable sentences is a subset of the set of grammatical sentences. This commits us to the view that, barring independent evidence to the contrary, both the unaccusative and unergative reduced relatives are grammatical. Second, we respect the Competence Hypothesis (Miller and Chomsky 1963), which argues that the grammar used by the human sentence processor is the same grammar that linguists study. Our performance hypothesis is simply our MG grammar and the co-occurrence restriction it encodes. Third, we adopt the Distributed Morphology (Halle and Marantz 1993) framework for encoding argument structure. Distributed Morphology takes the view that the Lexicon is not modularly encapsulated, but is itself derivations. Finally, we take the view that ambiguity resolution is the main end towards which the human sentence processor works. To this end, we adopt as a parsing model the Entropy Reduction Hypothesis, which models parsing difficulty as the reduction in analytic entropy from one sentence token to the next.

Our account argues that reduced relatives are grammatical but difficult to process. Uncontroversially, reduced relative garden path sentences require that a noun-verb-prepositional phrase sequence which initially appears as an active voice, intransitive main clauses must ultimately be analyzed as passivized, transitive, reduced relative constructions. The co-occurrence restriction observed in the literature is that unaccusatives freely participate in the causative alternation, but motion unergatives require a directional, argument-attached PP. We have proposed that the processing asymmetry for unergative and unaccusative reduced relatives is situated precisely on the production asymmetry for unergative and unaccusative verbs presented in Hoekstra (1988), Levin and Rappaport-Hovav (1995), and Folli and Harley (2006). In the unaccusative reduced relative condition, the reduced relative ambiguity can be resolved independently of the prepositional phrase attachment ambiguity, as unaccusatives can be reanalyzed as causative independent of prepositional phrase attachment. However, in the unergative case, the reduced relative ambiguity and the prepositional phrase attachment ambiguity must be resolved together, since causative reanalysis of unergatives requires a particular prepositional phrase attachment. Transitivity reanalysis of unergatives requires an argument attachment of the prepositional phrase in order to realize the verb as a passive, causative, reduced relative form.

The study confirms that the co-occurrence restriction between prepositional phrase attachment and causation of unergatives is responsible for the increased processing difficulty of unergative reduced relatives. The unergative condition exhibits a later, more severe reduction in entropy than the unaccusative condition. Notably, these results obtained even though we abstracted away from the relative frequency of the reduced relative and main clause constructions. A replication of this experiment where these constructions are weighted realistically would undoubtedly exacerbate the processing asymmetry predicted in our model. It would also be desirable in future work to estimate the parameter for adjunct versus argument attachment of prepositional phrases in general from a corpus; some techniques for this estimation are discussed in Merlo and Ferrer (2006).

We have attempted to analyze the contribution of verb type to reduced relative processing with maximum parsimony. Though the subject matter is inherently concerned with the interaction of syntactic factors with lexical factors, we wanted to explore this interaction while stipulating as little as possible about the nature of the lexicon. Thus, we used the Distributed Morphology framework to model a well reported co-occurrence restriction (Hoekstra 1988, Levin and Rappaport-Hovav 1995, Folli and Harley 2006) in the literature. Our Minimalist Grammar implements this co-occurrence restriction, but does not render a stance on what is done "in" or "out" of the lexicon. Empirically, this account of the effect in Stevenson and Merlo (1997) dovetails neatly with Pritchett (1992); while both unaccusatives and unergatives are optionally transitive, the transitivity of unergatives is conditional on prepositional phrase attachment. The unergative condition is biased against transitive reanalysis; we would predict unaccusatives to be as difficult as Pritchett's optionally transitive verbs, but more difficult than obligatorily transitive verbs.

The account avoids special appeal to lexical modularity, and assumes a strongly incremental sentence processing mechanism common to both unergative and unaccusative conditions. The account proceeds with the strong assumption that all analyses are maintained as long as they are congruent with the incoming sentence, and the long-held view that sentence processing is primarily an ambiguity resolution task. Our account readily accommodates not only the Stevenson and Merlo (1997) data, but more generally the Pritchett (1992) results of a processing advantage for obligatorily transitive verbs over optionally transitive verbs in the reduced relative construction. The obligatorily transitive verbs are never congruent with an intransitive main clause analysis, whereas optionally transitive verbs are congruent with both the intransitive main clause analysis and the transitive, passivized, reduced relative analysis, until the disambiguating token is met.

References

BNC. 2007. Distributed by Oxford University Computing Serv. Version 3.

Chi, Zhiyi. 1999. Statistical properties of probabilistic context-free grammars. *Computational Linguistics* 25:131–160.

Folli, Raffaella, and Heidi Harley. 2006. On the licensing of causatives of directed motion: Waltzing Matlida all over. *Studia Linguistica* 60:121–155.

Grenander, Ulf. 1967. *Syntax-controlled Probabilities.*. Division of Applied Mathematics, Brown University.

Guillaumin, Matthieu. 2004. Conversions between mildly context sensitive grammars.

Hale, John T. 2003. Grammar, Uncertainty and Sentence Processing. Doctoral dissertation, The Johns Hopkins University.

Halle, Morris, and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger.*, ed. K. Hale and S. J. Keyser, 111–176. MIT Press.

Hoekstra, Teun. 1988. Small clause results. *Lingua* 74:101–139.

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

Levin, Beth, and Malka Rappaport-Hovav. 1995. *Unaccusativity: at the Syntax- Lexical Semantics Interface.*. MIT Press.

MacDonald, Maryellen C. 1994. Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes* 157–201.

Merlo, Paola, and Eva Ferrer. 2006. The notion of argument in PP attachment. *Computational Linguistics* 32.

Miller, George, and Noam Chomsky. 1963. Finitary models of language users. In *Handbook of mathematical psychology*, ed. R. D. Luce, R. R. Bush, and E. Galanter, volume 2, 419–491. Wiley.

Pritchett, Bradley L. 1992. *Grammatical Competence and Parsing Performance*. University of Chicago Press.

Shannon, Claude E. 1948. A mathematical theory of communication. *Bell Sys. Tech. Journal* 27:379–423,623–656.

Stabler, Edward P. 1997. Derivational minimalism. In *Logical Aspects of Computational Linguistics.*, 68–95. Springer.

Stevenson, Suzanne, and Paola Merlo. 1997. Lexical structure and parsing complexity. *Language and Cognitive Processes* 12.

Zubizarreta, Maria L., and Eunjeong Oh. 2007. *On the Syntactic Composition of Manner and Motion*, volume 48 of *Linguistic Inquiries Monographs*. MIT Press.

Department of Linguistics
203 Morrill Hall
Cornell University
Ithaca, NY 14853-4701
kwg33@cornell.edu