The syntax-semantic interface: On-line composition of sentence meaning

Liina Pylkkänen
Department of Linguistics & Department of Psychology
New York University

Brian McElree
Department of Psychology
New York University

Natural language often challenges us to derive salient differences in meanings from superficially similar expressions. For example, consider the resultative, depictive, and small clause constructions in (1) - (3).

(1) The stylist combed the hair straight. (Resultative)

(2) The stylist combed the hair wet. (Depictive)

(3) The stylist considered the hair straight/wet. (Small Clause)

Unlike prenominal adjectives, which mostly serve as noun modifiers in English (c.f., *The stylist combed/considered the straight/wet hair*), postnominal adjectives can relate to the structure within which they are embedded in a variety of ways. In (1), the resultative construction, the adjective *straight* relates causally to the action described by the verb (*viz*., the hair became straight as a result of the stylist combing the hair). The superficially similar depictive in (2), on the other hand, involves no causal relation but rather asserts that the direct object was in the state described by the postnominal adjective during the event described by the verb. Finally, the small clause construction in (3) likewise involves no causal relation, but unlike the depictive, it does not assert that the hair was straight or wet during the act of considering, or ever as a matter of fact.

How are we able to compute these types of distinct interpretations? Part of the answer is that our language comprehension system includes compositional operations that are able to assemble meaning from our knowledge of words and how our grammar codes the types of relations implicit in (1) - (3). However, although psycholinguists have made remarkable progress over the past few decades in understanding the mechanisms involved in identifying the meaning of words and those involved with incrementally assigning grammatical structure, very little attention has been devoted to semantic composition.

Consequently, we understand comparatively less about this component of the language processing system. This is somewhat surprising since compositional processes sit at the interface between lexical and syntactic processing on the one hand and higher-order discourse processing on the other. A complete theory of comprehension requires an understanding of how comprehenders rapidly assemble an interpretation as lexical and syntactic constraints become available incrementally in a left to right analysis of the input.

In this chapter, we draw from current perspectives in semantics to outline a taxonomy of compositional operations that can serve as a framework for psycholinguistic and neurolinguistic investigations of semantic interpretation. We highlight research that finds natural expression within this framework, and, perhaps as importantly, given the paucity of relevant psycholinguistic research on composition, we point to areas that need further investigation.

Interpreting an expression typically requires integrating it into an evolving discourse model. Frequently, this requires resolving ambiguities at conceptually distinct levels, fixing reference, and drawing inferences to align local and global aspects of the discourse. To fully accomplish this task, it is uncontroversial that, in addition to lexical and syntactic constraints, comprehenders must draw upon pragmatic knowledge. The importance of high-level constraints has been illustrated by findings that comprehenders sometimes adopt a pragmatically plausible interpretation even if it is incongruent with lexical and syntactic constraints. For example, Wason and Reich (1979) and Natsopoulos (1985) documented "verbal illusions" in which comprehenders interpret sentences such as *No head injury is too trivial to be ignored* in a manner paraphrased by "No matter how trivial it might appear, a head injury should be treated." This interpretation cannot be derived from any veridical

application of local compositional operations (for an overview, see Sanford & Sturt, 2002). Similarly, Ferreira and colleagues (Ferreira, 2003; Ferreira, Bailey, & Ferraro, 2002) have shown that comprehenders often fail to accurately interpret surprisingly simple and common sentences, apparently opting for shallow forms of processing that are "good enough" for some purposes.

For some researchers, these findings challenge the existence of a compositional mechanism. For example, Ferreira et al. (2002, p. 11) state "[the] assumption of compositionality seems eminently plausible, but results in the literature on the psychology of language call it into question." Shallow and incorrect interpretations indicate that language stimuli, like stimuli in other domains, can be processed to different depths depending on task demands and subjective criteria, but they provide virtually no grounds on which to motivate extreme approaches that would eschew altogether the notion of a compositional mechanism. For, without such a mechanism, it is difficult to imagine how comprehenders would be able to understand novel utterances or exploit the lexical properties and syntactic constraints necessary to derive interpretations in cases such as (1) - (3).

Others have used a failure to find unique brain (ERP and fMRI) responses to violations of semantic and pragmatic knowledge to mount an attack on a distinct level of semantic integration. Hagoort, Hald, Bastiaansen, and Petersson (2004) contrasted sentences that were either false for Dutch speakers given their culture-specific knowledge, e.g., *The Dutch trains are white...,* or false based on more general knowledge, e.g., *The Dutch trains are sour....* They suggested that the latter was anomalous for "semantic internal reasons" because "the core meaning of sour is related to taste and food…[and] semantic features related to taste and food do not apply to trains" (p. 438). ERP responses to the two

anomalies patterned differently in the spectral domain, but both anomalies elicited similar N400 effects. This led Hagoort et al. to argue that there is no principled distinction between semantic and world-knowledge integration.

However, it is entirely unclear whether compositional processes would draw a distinction between general and culture-specific knowledge by including, for example, feature checks on properties such as "taste" and "food." Indeed, to do so would preclude constructing various "figurative" interpretations that routinely require comprehenders to compose interpretations with just these types of feature mismatches (cf., *Some lawyers are sharks*, Glucksberg, 2001). Even if compositional processes respect such a distinction, comprehenders might have used pragmatic knowledge in this task to attempt to construct a plausible construal of *The Dutch trains are sour*, which would engender the same N400 component observed with *The Dutch trains are white*. Hence, we see no reason to draw strong conclusions concerning the nature of semantic processing from this type of null result.

We believe that a complete understanding of language comprehension requires a detailed understanding of the set of composition operations that comprehenders have at their disposal, and how these operations interface with the rest of the language processing system. This will require psycholinguistic and neurolinguistic research to isolate the different operations involved in semantic composition, to identify which are costly for the language processor to compute, and why some might be more costly than others. However, a necessary first step is to clearly identify candidate operations.

1. COMPOSITIONALITY

The principle of compositionality, usually attributed to Frege (1892), can be stated as follows:

> The meaning of an expression is a function of the meanings of its parts and the way they are syntactically combined.

Given the systematicity and productivity of language, some version of compositionality is assumed by all linguistic theories of semantic interpretation. However, a major point of controversy is how strictly compositionality applies. In some theories, there is a complete homomorphism between syntax and semantics. Under this strong interpretation of compositionality (henceforth *strong compositionality*), the meanings of sentences are *fully* determined by the meanings of their constituents and by the syntactic way the constituents are combined (e.g., Fodor and Lepore, 2002; Montague, 1970). Alternative theories allow for semantic rules that do not correspond to any syntactic process (henceforth *weak compositionality*). In these theories each syntactic step still corresponds to a semantic step, but there is an additional inventory of purely semantic rules that may serve to change the meaning of a constituent to "fit" that of another (e.g., Barker, 2002; Hendriks, 1988; Jacobson, 1999; Partee & Rooth, 1983).

Whether natural language obeys strong or weak compositionality is an unresolved question about the syntax-semantics mapping. In what follows we first sketch three basic rules of compositional interpretation that are taken to be basic, strongly compositional, and generally uncontroversial (Heim & Kratzer, 1998; Jackendoff, 2002). These rules specify the basic means of composing meanings from the products of lexical and syntactic analysis, and we believe they represent the types of operations that should be incorporated into

psycholinguistic models of interpretation, either explicitly or functionally. We then review well-known phenomena that challenge strong compositionality and outline the main approaches to these phenomena in linguistic theory. Throughout, we discuss extant psycholinguistic results that bear on these issues, situating them in the outlined framework.

2. BASIC RULES OF COMPOSITIONAL INTERPRETATION

2.1. SEMANTIC TYPES AND FUNCTIONAL APPLICATION

The task of a theory of semantic interpretation is to characterize how elements in a syntactic string semantically relate to one another. Clearly, this depends on how we conceptualize the meanings of the elementary building blocks, the *terminal nodes* of a syntactic tree. Informally, an unchallenged view of lexical meaning is that the meanings of words have "holes" in them, which need to be filled by other words. For example, *destroy* has no valid interpretation unless it occurs in the context of a noun phrase that can be interpreted as the entity undergoing destruction, c.f. (4). Since *destroy* is not complete without a destroyee, the meaning of *destroy* can be taken to contain a variable standing for the destroyed entity.

(4) The boy destroyed the sand castle.

(5) *The boy destroyed.

Whether the meanings of transitive verbs also contain variables for the subject argument is controversial, but for ease of exposition, let us assume that they do. Hence, the meaning of *destroy* would be a formula such as *y destroy x*. Since *destroy* cannot by itself be used to assert that someone destroyed something, a standard way to characterize its meanings is as a rule or *function*. In other words, *destroy* is a function that requires two individuals as its input and produces as an output a description of a destruction of something by someone.

This is formally expressed in the lambda calculus in (6). The input, or the *arguments* of the function, is prefixed with lambdas, and the value, or the output of the function, follows the lambda terms.

(6) **destroy:** $\lambda x.\lambda y.$ destroy(y, x)

The function in (6) specifies that *destroy* requires two individuals as its arguments in order to yield a sentence that can be evaluated as either true or false. Given this, its *semantic type* is said to be a function from individuals to truth-values. Individuals and truth-values are taken to be basic, or so-called "saturated" types; they take no arguments, and therefore they can only function as arguments, not as predicates. The type of individuals is labeled 'e' and the type of truth-values 't'. All other semantic types can be derived from these basic types, according to the rule given in (8) (Montague, 1974).

(7) Basic types: *e* (individuals) and *t* (truth-values)

(8) If σ and τ are semantic types, then $<\sigma,\tau>$ is a semantic type.

This system of semantic typing assumes that the only *ontological* categories that natural language cares about are individuals and truth-values. It assumes that all meanings can be expressed in terms of these two categories. For example, the meaning of *destroy* would be described in terms of the derived type <e<e, t>>.

Most semanticists would agree that more ontological categories are required to explain semantic phenomena. In this system, for example, all sentences have denotations that we should be able to evaluate as either true or false in the actual world. However, a hallmark of natural language is its ability to describe situations other than the actually occurring ones, e.g., *I believe it is raining*. The verb "believe" creates an "opaque" or "intensional" context, and this requires enriching the basic ontology with a notion of

possible worlds. Although intensionality is a core topic in natural language semantics, we will not discuss it here, as it has not figured prominently in psycholinguistics (but see Clark & Haviland, 1974).

More relevant to extant psycholinguistic work on compositionality is the enrichment of the basic ontology by events (Davidson, 1967; Kratzer 1996, in press; Parsons, 1985, 1990; Rothstein, 1998; Tenny & Pustejovsky, 2000). This work captures the intuition that verbs essentially describe properties of events. Formally, this can be realized by adding an event variable to the verb's argument structure. In (9), the revised entry for *destroy*, 'e' is a variable ranging over events. The two individuals that are realized in syntax as subject and object of the verb are participants in the event e, and they relate to the event via the thematic relations 'agent' and 'theme'. In this analysis, the semantic type of *destroy* would be <e<e<s, t>>>, where *s* is the semantic type for events.

(9) **destroy:** λx.λy.λe. destroy(e) & agent(e, y) & theme(e, x)

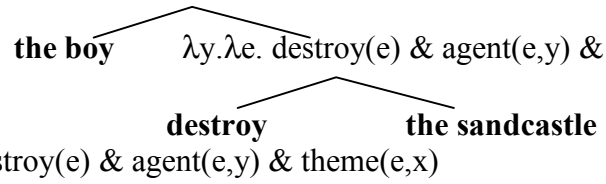
The event argument has become an indispensable tool in investigations of verbal semantics and argument structure. Therefore, in the discussions to follow, we will assume that the basic ontology of natural language includes events. In this framework, the types of some common expressions would then be as in (10):

(10)

Semantic type	Expression
e	names (*John*), definite noun phrases (*the cat*)
<e,t>	common nouns (*cat*), adjectives (*tall*)
<e<s,t>>	intransitive verbs (*run*)
<e<e,<s,t>>>	transitive verbs (*kick*, *destroy*)

Equipped with this system, we can return to our example sentence *the boy destroyed the sandcastle*. In order to interpret this sentence, we first apply the function in (9) to the two noun phrases *the boy* and *the sandcastle*.[1] This saturates the two individual arguments of *destroy*.[2]

(11)



The tree in (11) illustrates the basic compositional rule of *Functional Application*, so called because it refers to the application of a function to its arguments. Every time we apply a function to its syntactic sister, the denotation of the sister replaces the variable in the function, and we erase the corresponding lambda term to indicate saturation of the argument. The formal definition of Functional Application is given in (12).

(12) *Functional Application* (FA):

> If α is a branching node, $\{\beta, \gamma\}$ is the set of α's daughters, and $[[\beta]]$ is a *function* whose domain[3] contains $[[\gamma]]$, then $[[\alpha]] = [[\beta]] ([[\gamma]])$.
>
> (Heim & Kratzer, 1998)

In sum, when the relationship between two syntactic elements is one of *selection*, such that one element denotes a function that selects another element as an argument, the mode of semantic composition is functional application. A large body of psycholinguistic

---

[1] We will ignore the interpretation of tense, aspect and various other higher inflectional heads.

[2] The event argument remains unsaturated, and so at some point it must be existentially closed either by another predicate or by a default operation. However, these details are not important for our discussion.

[3] The "domain" of the function refers to the set of all possible input values to the function.

research, briefly reviewed below, has investigated the role of argument structures in parsing and interpretation. Within the framework proposed here, this research can be viewed as investigating the real-time deployment of functional application. Next, we turn to situations where two elements combine in the absence of a selectional relation.

2.2. PREDICATE MODIFICATION

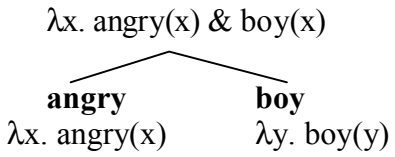
How do we interpret modifiers, such as the adjective *angry* in (13), which presumably does not stand in a predicate-argument relation with either *boy* or *the*?

(13) The angry boy destroyed the sandcastle.

For (13) to be true, the individual who destroyed the sandcastle must have both the properties of being a boy and the properties of being angry. Thus, a rule of composition that formed the intersection of boyness and angriness when given these two properties as its input would account for (13). Predicate modification, stated in (14), achieves exactly this, and the tree diagram in (15) illustrates the application of this rule to (13).

(14) *Predicate Modification* (PM):
If α is a branching node, $\{\beta, \gamma\}$ is the set of α's daughters, and $[[\beta]]$ and $[[\gamma]]$ are both of type <e, t>, then $[[\alpha]] = \lambda x.\ [[\beta]](x) \ \&\ [[\gamma]](x)$.
(slightly modified from Heim & Kratzer, 1998)

(15)

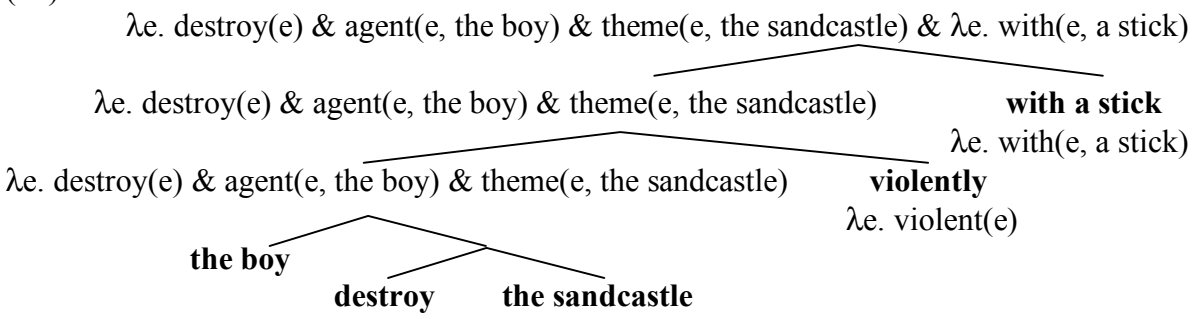


In an event semantics framework, a rule resembling predicate modification can capture intersective modification in the verbal domain as well. For example, all entailments

in (16) can be accounted for by treating VPs and VP-adverbs as the same type of predicates of events, as shown by the step-by-step derivation in (17).

(16) The boy destroyed the sandcastle violently with a stick.
→ The boy destroyed the sandcastle.
→ The boy destroyed the sandcastle violently.
→ The boy destroyed the sandcastle with a stick.

(17)



Alongside functional application, predicate modification is considered a basic composition rule that maintains full transparency of meaning. Of course, not all modification is intersective. For example, an *alleged murderer* is not necessarily a murderer. Thus, predicate modification is not an across-the-board interpretive rule for adjuncts. Space does not allow us to fully discuss nonintersective modification (but see Kamp & Partee, 1995; Frazier, 1999; Murphy, 2002).

2.3. PREDICATE ABSTRACTION

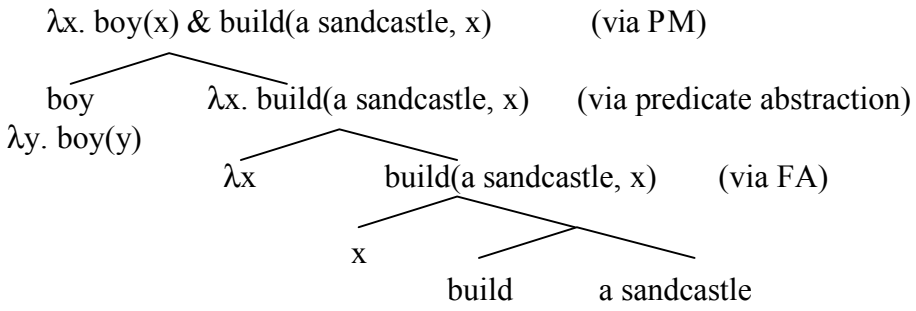
Relative clauses have the surface syntax of sentences, but semantically they function as modifiers. For (18) to be true, the individual liked by Mary needs to have both properties of being a boy and of having built a sandcastle.

(18) Mary likes the boy who built a sandcastle.

We could derive this meaning if *who built a sandcastle* was converted into a predicate of type <e, t> and then combined with *boy* via predicate modification. For this,

most theories include a rule called *Predicate (lambda) abstraction*, which converts a formula into a predicate by binding a free variable within the formula with a lambda-operator. In our example, the relative pronoun *who* is thought to originate in the subject position of the verb, leaving behind a trace, represented as a variable. In its moved position, the relative pronoun is represented and interpreted as a lambda abstractor (for simplicity, we ignore event variables here).

(19)



Predicate abstraction is said to be involved in many types of constructions (for a review, see Partee, TerMeulen & Wall, 1990). Recently it has also been proposed to be *the* variable binding mechanism in movement (Heim & Kratzer, 1998). Strictly speaking, predicate abstraction is not a compositional rule, but rather an operation of merging a covert variable binding operator into the syntactic tree.[4]

2.4. REAL-TIME PROCESSING OF STRONGLY COMPOSITIONAL EXPRESSIONS

The three operations outlined above are generally viewed as strongly compositional. Although they do not exhaust the inventory of composition rules, they provide a formal account of many of the primary ways that meanings are composed. As such, they provide a

[4] In this sense, predicate abstraction is more similar to, say, existential closure, except that the lambda operator changes the type of its complement.

detailed set of hypotheses for the investigation of the actual processing operations in semantic interpretation.

2.4.1. *Differences among basic operations?*

If we take the view of word meaning described above quite literally, and regard lexical items as functions, it would seem intuitive that functional application should be perhaps the most basic composition operation that comprehenders have at their disposal. This operation enables comprehenders to exploit lexical properties to compose basic predicate-argument relations. There now exists a large body of psycholinguistic evidence indicating that lexical information is used to rapidly converge on an interpretation. Carlson and Tanenhaus (1988) provided initial demonstrations that the thematic grid of a verb is used to assign thematic roles to arguments immediately after processing the verb and to resolve different types of ambiguities. Over the past decade, researchers have amassed considerable evidence supporting this claim (e.g., Boland, Tanenhaus, & Garnsey, 1990; Boland, Tanenhaus, Garnsey, & Carlson, 1995; Carlson & Tanenhaus, 1988; Trueswell, Tanenhaus, & Kello, 1993).

Neither predicate modification nor predicate abstraction establish basic predicate-argument relations. If functional application is indeed the more basic operation, we might expect that it would be preferred to other operations. Current evidence is suggestive, although less than definitive.

Several studies have found that argument phrases are easier to process than adjunct phrases in otherwise neutral contexts (e.g., Boland, 2005; Boland & Boehm-Jernigan, 1998; Clifton, Speer, & Abney, 1991; Kennison, 1999, 2002; Liversedge, Pickering, Clayes, &

Branigan, 2003; Liversedge, Pickering, Branigan, & Van Gompel, 1998; Schütze & Gibson, 1999; Speer & Clifton, 1998). For example, Clifton et al. (1991) found that arguments (e.g., *The saleswoman tried to interest the man in a wallet…* or *The man expressed his interest in a wallet...*) were read faster than adjuncts that modified either a preceding verb (e.g., *The man expressed his interest in a hurry…*) or a preceding noun (e.g., *The saleswoman tried to interest the man in his fifties…*). Schütze and Gibson (1999) reported that participants read prepositional phrases (PP) in verb-NP-PP sequences faster when they were arguments of a NP (e.g., *The company lawyers considered employee demands for a raise…*) as opposed to adjuncts of the verb (e.g., *The company lawyers considered employee demands for a month…*). Finally, Kennison (1999) reported that participants read agentive by-phrases in complex event nominals *faster* when they were arguments (e.g*., The frequent collection of butterflies by the kids…*) rather than adjuncts (e.g., *The numerous collections of butterflies by the kids…*).

To the degree that adjuncts involve predicate modification, the privileged processing status for arguments suggests a preference for functional application. A notable complication is that most published studies have investigated sentences with temporary structural ambiguities. For example, in both the Clifton et al. (1991) and the Schütze and Gibson (1999) studies, the PPs were ambiguous between an argument and adjunct analysis until the noun was encountered. Hence, the slower processing times for adjuncts may only reflect parsing preferences.

It is difficult to cleanly disentangle syntactic and semantic effects in strongly compositional constructions, because of the tight linkage between syntactic and semantic composition. However, Schütze and Gibson (1999) argue that simple structural parsing

strategies (e.g., *Minimal Attachment* or *Late Closure*, Frazier and Rayner, 1982) cannot explain their findings. They suggest that a better account is one that posits "an argument preference strategy" (e.g., Abney 1989). This strategy essentially grants priority to functional application in proposing that comprehenders strive to saturate the arguments of a function rather than pursuing analyses that involve operations such as predicate modification. Other findings also partially implicate interpretative factors in the argument advantage. Speer and Clifton (1998) found that the plausibility of a constituent as an argument or as an adjunct dramatically modulated overall reading times. Plausibility did not fully eliminate the argument advantage, but it did interact with argumenthood such that the argument advantage was less with high plausibility. This partly implicates interpretative factors, as plausibility reflects the semantic fit of an item as an argument or as a modifier.

An alternative means of investigating whether functional application has a privileged status is to determine whether it is computationally less costly than other operations. Indeed, functional application might be preferred because it is less costly. Unfortunately, we do not currently have good direct measures of the relative cost of different basic compositional operations. Studies demonstrating that arguments are easier to process than adjuncts in structurally ambiguous contexts do not necessarily bear on the issue, since the processing cost might reflect an initial bias for arguments. Hence, a challenge for future research is to construct and examine relatively unambiguous structures that vary the mode of composition holding other factors constant (e.g., frequency of occurrence).

2.4.2. *Intrinsic processing issues*

In addition to investigating potential differences between basic composition operations, research is needed to systematically examine how different factors affect each operation. The expectation is that such research will provide the foundation for developing explicit processing models, as analogous research has done in the investigation of other components of language comprehension (e.g., word recognition). Below, we note a few salient issues.

The intrinsic complexity of the compositional operation may be an important determinant in performance. Concerning functional application, for example, two recent papers have argued that the complexity of a verb's event structure affects processing speed. McKoon and Macfarland (2000, 2002) reported longer recognition times for verbs with externally caused events (e.g., *break*) than for verbs with internally caused event (e.g., *grow*), which under their analysis (following Levin & Rappaport Hovav, 1995) have a simpler noncausative event structure. Gennari and Poeppel (2003) report that lexical decision and self-paced reading times were longer for eventive verbs (e.g., *destroy*, see (9)), which describe causally structured events, than for stative verbs (e.g., *love*), which involve a simpler event structure.

Natural language can contain ambiguities at any number of conceptually distinct levels of representation, and the interpretive level is no exception. A major issue for any processing model is how comprehenders resolve ambiguities in the application of an operation. For example, there can be uncertainties about whether an adverb relates to the event, as in the resultative construction in (1), or to one of individuals in the event, as in object depictive construction in (2) (for preliminary findings, see Frazier & Clifton, 1996;

Pylkkänen & McElree, 2004). Likewise, in noun modification, there are often uncertainties about whether the modification is intersective or nonintersective (e.g., *a beautiful dancer*).

Several factors are likely to conspire to promote one analysis over another. Some analyses might be preferred because they are inherently more amenable to incremental analysis. For example, in noun modification, there are reasons to suppose that intersective interpretations may be less costly and more compatible with rapid incremental interpretation than nonintersective interpretations, as, in principle, the interpretation of an intersective adjective need not be relativized to the head noun (Frazier, 1999). Currently, however, there is no clear evidence for a general preference for intersective modification (Frazier, 1999).

Lexical and pragmatic properties are clearly an important determinant. For example, in depictive structures, the semantic and pragmatic properties associated with an adjective can induce an interpretation where the adjective is taken to modify the subject (e.g., *The man ate the meat angry*) or, conversely, the object (e.g., *The man ate the meat raw*). In noun phrase modification, the salience of the modified dimension appears to be important. Murphy (1990) found that nouns modified by adjectives were understood faster than nouns modified by other nouns or by adjectives derived from nominals (e.g., *corporate check*). He argues that these differences reflect the fact that adjectives (e.g., *blue sky*) often denote the dimension they modify, whereas one noun can often modify another on several dimensions (see Murphy, 2002, Ch. 12 for an excellent overview).

Given the rather ubiquitous effects of frequency on other components of the language processing system, it would be surprising if compositional operations were not sensitive to differences in frequency at various grains of analysis. For example, Pylkkänen and McElree (2004) found that resultatives were read faster and more accurately than both

subject depictives (e.g., *The teenage boy painted the wall turquoise…* versus *The teenage boy painted the wall bored…*) and object depictives (e.g., *The artist knocked the frame crooked…* versus *The artist returned the frame crooked…*), even though the adjectives were matched on relevant properties such as frequency, length, and similarity to the modified noun. These differences could reflect architectural differences, but a corpus analysis of 1,462 V-NP-AP sequences in the Penn Treebank corpus (Marcus, Santorini, & Marcinkiewicz, 1993) revealed that resultatives are nearly an order of magnitude more frequent than depictives (370 versus 39 instances). Thus, it is quite likely that the frequency of an operation facilitated performance. Additionally, it is reasonable to expect that finer grain co-occurrence frequencies will also affect performance. For example, in a corpus analysis, Spivey-Knowlton and Sedivy (1995) found that *with*-PPs are more likely to modify a verb than a direct object with action verbs (e.g., *smashed*), but that the opposite pattern held with psychological state and perception verbs (e.g., *hoped* and *heard*). Self-paced reading times reflected these differences: *with*-PPs were read faster when the semantic properties of the PP were consistent with the more frequent modifiee.

Finally, discourse constraints are likely to induce different compositional operations. In a visual world paradigm, in which eye movements were monitored while comprehenders interact with visual displays, Sedivy, Tanenhaus, Chambers, and Carlson (1999) reported striking evidence that comprehenders rapidly interpret scalar adjectives (e.g., *the tall glass*) contrastively when the referential domain supported this interpretation (see Phillips, 2001 for discussion of the implications of these findings for theories of incremental compositional interpretation). Other findings also indicate that basic interpretive operations interact with general contextual and discourse factors. For example, discourse information can eliminate

the processing disadvantage for adjuncts. Liversedge et al. (1998) found that, in isolation, locative by-phrases following passive verb forms (e.g., *The shrubs were planted by the greenhouse*) were more difficult to process than agentive by-phrases (e.g., *The shrubs were planted by the apprentice*), replicating the basic argument advantage. However, the adjunct disadvantage disappeared when context set up an expectation for a location (e.g., *The man was wondering where to plant the shrubs*). Collectively, these results point to the need to systematically investigate how different types of constraints can modulate processing costs and the likelihood of pursuing various compositional operations.

3. CHALLENGES FOR COMPOSITIONALITY

Although the interpretive rules sketched in Section (2) have significant empirical coverage, many semantic phenomena seem to call for more powerful interpretive mechanisms. We provide an overview of constructions that challenge the simple view described so far. First, we discuss cases of *type-mismatch*, where elements can combine even though they do not constitute an appropriate structural environment for either functional application or predicate modification. Second, we review cases where a part of the meaning of a sentence is not overtly expressed by any of its syntactic constituents, and the unexpressed meaning cannot be generated by any basic compositional operation.

3.1. TYPE MISMATCH

3.1.1. *Quantifier interpretation*

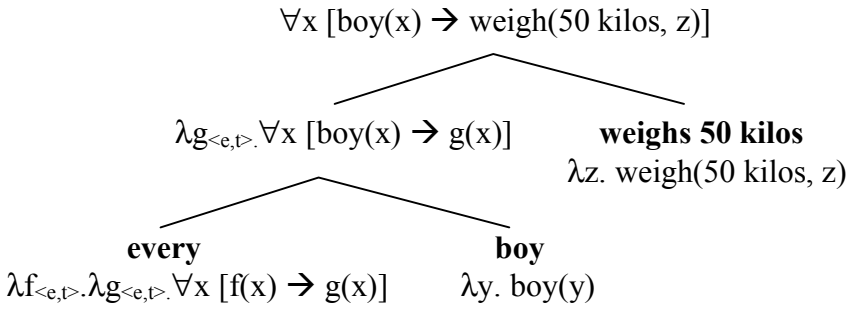
The system described so far cannot account straightforwardly for even the simplest cases of quantification. We have described a view where verbs are functions that can take individuals, such as *the boy*, as their input. Since quantified noun phrases such as *every boy*

combine with verbs without problem (*every boy ran*), the approach suggests that quantifier phrases denote individuals. However, such an approach gets the meaning of quantifier phrases wrong in many ways. Consider an example from Sauerland and Stechov (2001): if a quantified phrase such as *every boy* denoted a group consisting of every boy, then it should be possible to predicate properties such as *weighs 50 kilos* of that group. However, (20) clearly does not mean that the group of boys collectively weighs 50 kilos, rather that every (contextually relevant) individual who is a boy weighs 50 kilos.

(20) Every boy weighs 50 kilos.

To derive this interpretation, the notion that *weighs* takes *every boy* as it argument is relinquished. Instead, *every* is be treated as a function that takes two predicates of individuals as its arguments: The first one (e.g., *boy*) restricts the domain of quantification, and the second one attributes a property (e.g., the property of weighing 50 kilos) to those individuals satisfying the restrictor predicate. This is shown in the tree diagram in (21).

(21)

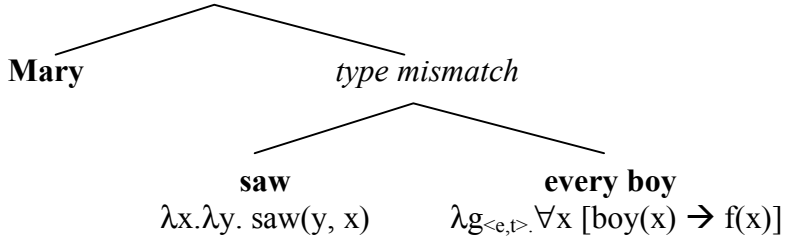


This standard treatment of subject quantifiers essentially dates back to Frege (1879).

By treating quantifier phrases as higher order predicates that take their verbs as their argument, instead of vice versa, we have an account of quantified subjects. However, this account cannot straightforwardly explain how we compose an interpretation of quantified *objects.* (22) illustrates the problem. If the object is quantified, as in (22), then the second

argument of the quantifier needs to be the transitive verb. But the lexical entry for *every* given in (21) requires the second argument of the quantifier to be of type <e,t>, i.e., an intransitive verb. A type mismatch results.

(22)

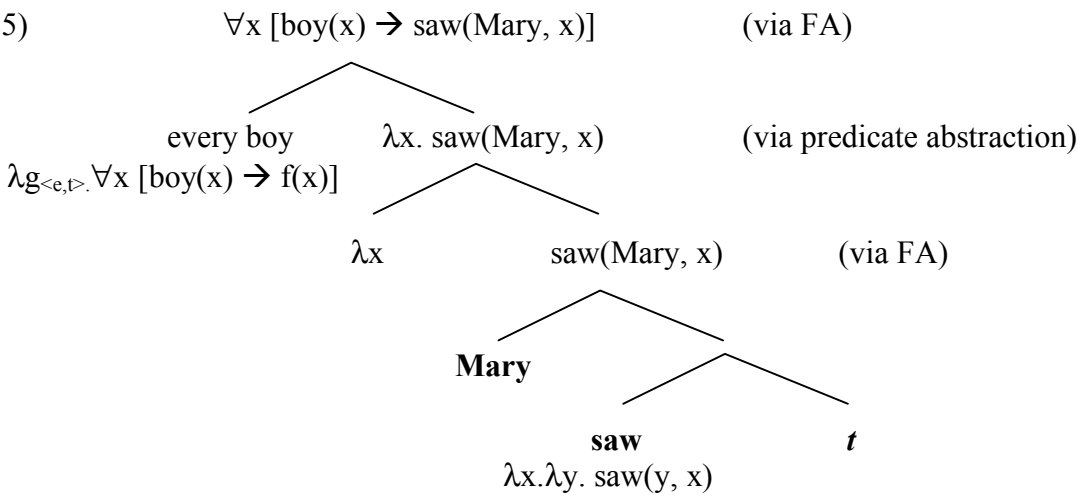


Two main approaches to this type of type mismatch have been proposed. One involves simply shifting the object to a higher type. In this account, there is a freely available type shifting rule that serves to lift the type of every (and other quantifiers) from (23) to (24).

(23) **every$_1$:** $\lambda f_{<e,t>}.\lambda g_{<e,t>}.\forall x\ [f(x) \rightarrow g(x)]$ (basic type)
(24) **every$_2$:** $\lambda f_{<e,t>}.\lambda g_{<e,et>}.\lambda x.\forall y\ [f(y) \rightarrow g(y, x)]$ (lifted type)

The other approach involves syntactic movement, or quantifier raising (QR) (Heim & Kratzer, 1998). The object is raised out of the VP, and in doing so it creates an abstract predicate above the VP. In other words, the movement *creates* a predicate that is of a suitable type for the quantifier.

(25)



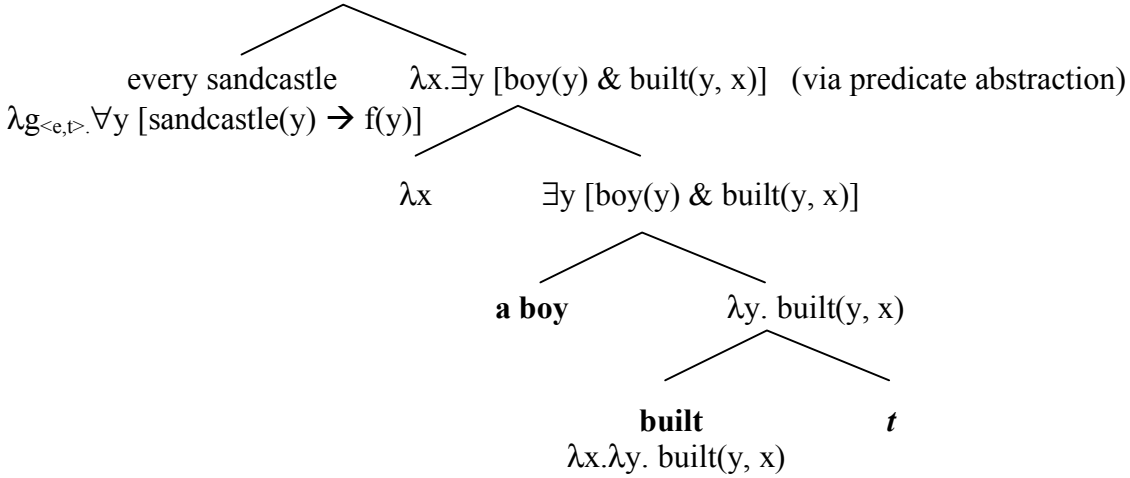
For our purpose, the essential point is that all extant theories involve some additional operations to account for the interpretation of object quantifiers.

Matters are complicated further if two or more quantifiers are present, such as (26), which includes an indefinite noun phrase. Cases such as this are inherently ambiguous, as the quantifiers can take different scope with respect to each other.

(26) A boy built every sandcastle.
- (i) Surface scope: There exists an x such that x is a boy and x built a sandcastle.
- (ii) Inverse scope: For all x, such that x is a sandcastle, there exists an individual y, such that y is a boy and y built x.

How is inverse scope computed? Under the assumption that object quantifiers can raise, inverse scope is easily obtained.

(27) $\forall x$ [boy(x) → $\exists y$ [boy(y) & built(y, x)] (via FA)



Under a QR approach, we must assume that the subject quantifier phrase can also move, in order to account for surface scope. Thus surface scope would be computed by further raising the subject quantifier over the object.

3.1.2. *Real-time processing of quantified expressions*

Only a few studies have investigated the interpretation of quantified expressions. So far, researchers have not addressed the basic question of whether quantifiers in object position are more costly to process than quantifiers in subject position, as most theories would predict. Instead, studies have focused on doubly quantified structures to determine whether one scope assignment is preferred or less costly to compute (e.g., Anderson, 2004; Filik, Paterson, & Liversedge, 2004; Gillen, 1991; Kurtzman & MacDonald, 1993; Tunstall, 1998). The guiding intuition is that the latter should be preferred or consume less processing resources.

However, because doubly quantified sentences are ambiguous, researchers can only infer what interpretations comprehenders consider while processing the quantifiers. A typical research strategy has been to co-opt an experimental protocol used to investigate syntactic ambiguities. This is illustrated by (28) and (29), taken from a study by Kurtzman and MacDonald (1993).

(28) A kid climbed every tree. The kid was full of energy.

(29) A kid climbed every tree. The kids were full of energy.

The singular continuation in (28), *the kid,* forces a surface scope interpretation of the first sentence, whereas the plural continuation in (29) forces an inverse scope interpretation. Differential difficulty at one of these continuations provides evidence for the relative dominance of one of the scope readings. Conversely, no difficulty at this region suggests one of three options: Either comprehenders are capable of computing both readings in parallel (e.g., Kurtzman & MacDonald, 1993), comprehenders compute an underspecified interpretation compatible with both readings (e.g., Gillen, 1991; see also Sanford & Sturt, 2002), or comprehenders stochastically compute either readings with approximately equal frequency.

Some studies have reported faster reading or judgment times for a singular continuation such as (28) than for a plural continuation such as (29), consistent with a preference for the computationally simpler surface scope reading (Anderson, 2005; Tunstall, 1998). Unfortunately, other results suggest that this difference may not reflect an overarching surface scope preference. Tunstall (1998) found an advantage in response times for grammaticality judgments of singular continuations in sentences analogous to (28), but comparable latencies for singular and plural continuations when the indefinite quantifier (*a*)

followed the universal quantifier (*every*) in surface order. In a self-paced (sentence-by-sentence) reading time study, Gillen (1991) found that, irrespective of surface order, singular continuations were read more quickly than plural continuations. Filik et al. (2004) also found a general advantage for singular continuations in an eye tracking study that factorially varied the type of quantifier, surface order of the quantifier, and ordering of grammatical arguments (direct and indirect object). These results undermine the claim that the advantage for singular continuations in structures like (28) reflect a strong preference for surface scope.

Interestingly, Anderson (2005) conditionalized reading times for doubly quantified sentences on responses to post-sentence questions, which were designed to indicate whether readers interpreted the sentences with surface or inverse scope. Sentences conditionalized on inverse responses were read more slowly than those conditionalized on surface responses. Hence, although there might not be a general preference for surface scope, this finding suggests a cost for computing a scope configuration other than what is represented in surface syntax.

However, Filik et al. (2004) found across the board differences in quantifier orders, with indefinite-universal orders being more costly to process than universal-indefinite orders. This finding is consistent with a prediction of Fodor (1982; see also Ioup, 1975) that comprehenders will initially assume that an indefinite NP refers to one entity. In an indefinite-universal ordering, this initial bias will limit comprehenders' ability to compute an inverse interpretation when the universal quantifier is encountered, as they would need to revise their interpretation of the indefinite to denote multiple entities. Note that no such bias would be found in a sentence with a universal-indefinite ordering. The implication of the Filik et al. (2004) finding is that the observed difficulty of computing an inverse scope

reading in indefinite-universal orders can only be attributed to the cost of computing inverse scope relations *per se* if there are also measurable effects in constructions with universal-indefinite orders. Otherwise, we cannot rule out the possibility that such effects represent reanalysis of initial semantic commitments.

Unfortunately, Anderson (2005) did not examine universal-indefinite orders in her conditional analysis. Nonetheless, she did find that, with highly constrained contexts, sentences with both relatively unambiguous inverse scope readings (e.g., *A different member tested every recipe*) and ambiguous inverse scope readings (e.g., *A member tested every recipe*) were read slower than sentences with both relatively unambiguous surface scope readings (e.g., *Every member tested a different recipe*) and ambiguous surface scope readings (e.g., *Every member tested a recipe*). To the degree that the indefinite NP, *a different member*, would be unlikely to be interpreted as one entity in a context that introduced several members (e.g., *...Members who nominated recipes were required to test the recipes to make sure that the instructions were correct. A different member…*), these results suggest that computing inverse scope is more costly than surface scope.

A host of factors may affect the likelihood that comprehenders commit to an initial interpretation of an NP when the sentence is ambiguous. This makes it rather difficult to construct clean tests of whether inverse scope is costly to compute across different constructions. Anderson's (2005) results with relatively unambiguous structures are suggestive. An alternative approach would be to investigate unambiguous structures with a singly quantified expression in object position. Such cases would provide a simpler test of whether the hypothesized type-shifting operation is taxing to perform.

3.1.3. *Type mismatch in conjunctions*

As so far described, noun phrases can be of two different types: individuals (type *e*) or higher order functions taking a predicate as their argument (type <<et>, t>). The higher order functions, such as quantifiers, take verbs as their arguments whereas the e-type noun phrases serve as the arguments of verbs. Under this hypothesis, one might expect that e-type and higher order noun phrases should not be able to conjoin, as this would create a conflict for the directionality of functional application. But this prediction is clearly not born out: proper names and definite noun phrases can conjoin with quantified noun phrases without problem as in (30) and (31).

(30) The teacher and every student left the classroom.
(31) Sally and some boy skipped class.

How is type-mismatch in conjunctions resolved? Two main approaches have been proposed. One is a so-called "generalizing to the worst case" solution where all noun phrases are treated as higher order predicates, or as so-called *generalized quantifiers* (Montague, 1970). The consequence of this is that noun phrases *always* take verbs as their arguments rather than vice versa. The other approach is to shift the types of e-type noun phrases to a higher type only when necessary (Partee & Rooth, 1983). Thus, the type-lifting rule shown in (32) would serve to shift the e-type argument to a higher type in both (30) and (31).

(32) lift: e → <<e,t>, t>
Sally → $\lambda f_{<e,t>}$. f(Sally)

Partee and Rooth (1983) propose type shifting as a last resort processing strategy to resolve type-mismatches. Type-shifting is a purely semantic operation with no syntactic reflex and therefore sacrifices strong compositionality. Although most semantic theories

incorporate type-shifting rules in some form or other, their existence is controversial (Heim & Kratzer, 1998). The lack of consensus about type-shifting may be at least partly due to the fact that it is difficult to find linguistic evidence that sharply distinguishes between different approaches. Given the controversy, type shifting would be a natural domain for psycholinguistic investigation. However, there have been no experimental studies investigating whether these traditional cases of type-shifting such as (30) and (31) are costly in-real time processing. Instead, an increasing body of experimental work has researched a closely related phenomenon, called "coercion." In coercion, the semantic types of syntactic sisters mismatch, and in addition, some part of the meaning of the sentence is syntactically unexpressed.

3.2. COERCION

Psycholinguistic research on compositionality has mainly investigated the processing of two types of expressions: *complement coercion* and *aspectual coercion*. Both complement coercion and aspectual coercion involve a semantic mismatch and have been argued to challenge strong compositionality. Work on these constructions has focused either on their representation or on their processing with little synthesis. In what follows, we consider both traditional linguistic evidence and psycholinguistic and neurolinguistic data in evaluating to what extent these cases indeed require non-syntactic interpretive mechanisms.

3.2.1. *Complement coercion*

So far we have only considered verbs that take individuals, such as *John* or *the book*, as their arguments. However, many verbs select, not for individuals, but for verb phrases (VPs). Aspectual verbs such as *begin* and *finish* are good examples of VP-selecting verbs. They

denote functions that, when given a VP describing an event as their input (such as *writing a book*), return a predicate that denotes the initial or final part of that event (*began/finished writing a book*). Given these selectional restrictions, strong compositionality would seem to predict that these verbs should be ungrammatical if combined with non-event-denoting objects. But interestingly, these verbs are entirely natural and grammatical with objects that denote individuals rather than events, as shown in (33) and (34).

(33) The author began the book.
(34) The student finished the thesis.

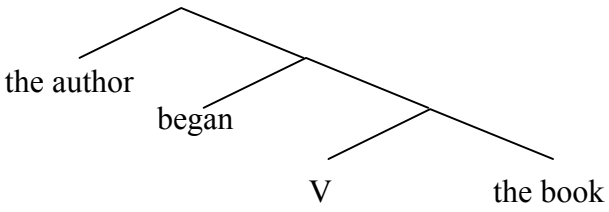
Indeed, corpus analyses demonstrate that expressions such as (33) and (34) are greater than 9 to 1 more frequent than forms that fully specify the event (e.g., *The author began writing a book*) when, as here, the event is commonly associated with the noun (e.g., reading or writing a book). Fully specified events frequently occur with less predictable activities, such as *translate the book* (Briscoe, Copestake, & Boguraev, 1990; Lapata & Lascarides, 2003; Lapata, Keller, & Scheepers, 2003).

Processing of these sentences has been hypothesized to involve a process of *complement coercion*, a non-syntactic process that converts the entity-denoting object into an event description that satisfies the selectional demands of the verb (Jackendoff, 1997; Pustejovsky, 1995). In (33) and (34), the end result of the coercion process would be an interpretation in which, for example, the author began writing the book and the student finished writing the thesis. Complement coercion has been taken to be an obvious disconfirmation of strong compositionality, but possible syntactic mechanisms for repairing the semantic mismatch have not been seriously considered. In what follows we sketch a syntactic solution to coercion and evaluate its feasibility.

3.2.2. *Coercion as VP-insertion*

A possible solution would be to assume that a syntactically silent VP has been inserted into the structure, as shown in (35), and that this circumvents the need for an extra-syntactic interpretive mechanism.[5]

(35)



3.2.2.1.*Evidence in favor of VP-insertion*

If coercion is based on VP-insertion, it should be blocked exactly in those environments where overt VP complements are ungrammatical. Many event-selecting verbs participate in the causative-unaccusative alternation, as shown in (36) and (37). (37) shows that an event-denoting nominal object can raise to subject position if an agent does not already fill that position. However, this is possible only when the direct object is a noun-phrase, not when it is a VP, as illustrated in the gerundive and infinitival VP complements in (39) and (41).[6]

---

[5] Much of natural language has been hypothesized to involve syntactic heads that do not have an overt pronunciation and thus there is nothing exotic about (35). For example, although English present tense is morphologically unmarked in all persons except the third singular, no one would claim that the syntactic tree of a sentence such as *I love Lucy* does not involve a tense node.

[6] The ungrammaticality of (39) and (41) is not due to a general constraint against gerundive and infinitival subjects (c.f. *Reading this book will be enjoyable* or *To read this book would be a mistake*), nor to a general prohibition against event-denoting subjects with unaccusatives (c.f. *The reading of the verdict began*).

Event-denoting NP-complement:

(36) We began the war. (causative)
(37) The war began. (unaccusative)

Event-denoting VP-complement (gerund):

(38) We began reading the book.
(39) *Reading the book began.

Event-denoting VP-complement (infinitive):

(40) We began to read the book.
(41) *To read the book began.

If coercion is purely semantic, there should be no particular reason for it to be blocked with unaccusatives. However, the data patterns clearly favor a VP-hypothesis. While (42) can be used to assert that Sarah began a book-reading event, (43) cannot be used to assert that a reading of a book began. Similarly, (44) can be used to convey that Bill continued writing a thesis, but (45) cannot be used to express that a thesis-writing event continued.[7]

(42) Sarah began the book.
(43) *The book began.
(44) Bill continued the thesis.
(45) *The thesis continued.

These data speak against an *a priori* exclusion of a syntactic mechanism for coercion. A further distributional test can be devised by embedding event-selecting verbs inside adjectives. In (46), the event-selecting verb *finish* is combined with an entity denoting

[7] (43) and (45) are only grammatical under the irrelevant reading where *the book* and *the thesis* are interpreted as texts: *the book began with a prologue*.

argument. Here, composition is mediated by the adjectival suffix *able* (and a semantically vacuous copula). The natural interpretation of this sentence is that some activity involving the book is finishable. Thus it appears that, in contrast to unaccusatives, coercion is possible in the subject position of deverbal adjectives (for supporting processing evidence, see Section 3.2.3.1).

(46) This book is finishable.
It is possible to finish [Ving] this book.

Now, if coercion occurs in the subject position of (46), and if coercion is VP-insertion, overt VP-arguments should also be possible in the subject position of able-adjectives. In other words, sentences such as those (47) and (48) should be well formed.

(47) Reading this book is finishable.
(48) Climbing this wall may be survivable.

Most native speakers seem to judge these sentences as marginal. Importantly, however, they appear better than the unaccusative examples in (43) and (45). If sentences such as (47) and (48) are indeed acceptable, there would seem to be a correlation between the distribution of coercion and the distribution of possible VP-arguments: When VP arguments are ungrammatical, coercion is blocked, and when VP arguments are possible, so is coercion. The hypothesis that coercion is VP-insertion derives this pattern for free.

3.2.2.2.*Evidence against VP-insertion*

Although the VP-hypothesis fares well in accounting for some distributional properties of coercion, it runs into problems with modification. If the structure of coerced expressions is as in (35), there should be two attachment sites for VP-modifiers: the higher VP and the lower VP. The examples in (49) and (50) illustrate the availability of the lower VP for VP-

modification in those cases where the VP is overt. The lower VP is modified by the manner adverb *slowly* in (49), and by an instrumental phrase in (50).

Overt VP: Adverbial can modify the lower VP

(49) We finished eating the meal slowly.
True if a slow meal comes to an end quickly.

(50) I started cutting a loaf of bread with a knife.
True if an event of cutting the loaf of bread with a knife was initiated.

If the structure in (35) is right for coercion, these VP-modifiers should be able to take lower scope in coerced expressions as well. However, this seems impossible, as shown in (51) and (52).

Coercion: Adverbial cannot modify the "lower VP"

(51) We finished the meal slowly.
False if a slow meal comes to an end quickly.

(52) #I started a loaf of bread with a knife.
False if an event of cutting the loaf of bread with a knife was initiated.
The knife must be an instrument of the initiation.

The VP-hypothesis also makes wrong predictions about passivization. With overt VP complements, raising of the object of the lower VP to the subject position of the matrix clause results in deviance, as shown by the (a) examples below. If the coerced versions of these sentences are syntactically identical to the overt versions, they should also be ill formed in the passive. But, contra this prediction, the coerced versions are entirely natural.

(53) a. ?The book was begun to be written by the author.
b. The book was begun by the author.

(54) a. ?The album was completed to be recorded by the artist.
b. The album was completed by the artist.

(55) a. ?The problem was attempted to be solved by the minister.

b. The problem was attempted by the minister.

In sum, if coerced expressions involve a silent VP, that VP for some reason cannot be adverbially modified and it does not constitute an intervener for A-movement in passivization. In other words, with regard to these phenomena, the VP would have to behave as if it was not there. Thus, despite the positive results regarding the distribution of coercion, the VP hypothesis is problematic in other domains. Consequently, a type-shifting analysis of coercion may be necessary, even if at the cost of sacrificing strong compositionality. Collectively the evidence suggests complement coercion constitutes a strong candidate for a purely semantic interpretive process.[8]

If coercion requires an extra-syntactic process that is not part of the default repertoire of interpretive rules, coerced sentences should be more costly to process than non-coerced controls. We next turn to processing measures of whether coerced sentences involve online construction of an event predicate from an entity-denoting noun phrase. Further, if the resolution of the type-mismatch in coercion involves *composition* of an event structure rather than just retrieval of a suitable activity, it should be possible to obtain a processing delay that is not simply due to activity retrieval.

3.2.3. *Real-time processing of complement coercion*

3.2.3.1. *Basic findings*

Reading time investigations of complement coercion have repeatedly found that they are indeed more difficult to process than various types of control expressions, consistent with

[8] A rigorous formal examination of the representation of coercion is beyond the scope of this chapter; for proposals, see Egg (2003), Jackendoff (1997) and Pustejovsky (1995).

the idea that they require costly operations to repair the type-mismatch between an event-selecting verb and an entity-denoting object. (McElree et al., 2001; McElree, Frisson, & Pickering, in press; Pickering, McElree, & Traxler, 2005; Traxler, Pickering, & McElree, 2002; Traxler, McElree, Williams, & Pickering, 2005). Initial demonstrations compared sentences such as (56) and (57).

(56) The carpenter began the table during the morning break. (coerced)
(57) The carpenter built the table during the morning break. (control)

Coerced expressions such as *began the table* were compared to control expressions, which used the verb that readers most often ascribed to the eventive interpretation of the coerced expression (determined by completion norms, see McElree et al. 2001). Reading time measures indicated that processing (56) was more costly to process than (57), even though the sentences were rated as equally plausible: In self-paced reading, participants took longer reading *table* and *during* in (56) than in (57) (McElree et al, 2001); in eye tracking, reliable differences first emerged at the NP (e.g., *the table*; Pickering et al, 2005) or on the two words following the NP (e.g., *during the*; Traxler et al., 2002).

These studies provided the initial evidence that expressions requiring complement coercion are indeed taxing to process. However, conclusions based on comparisons between two different of types constructions can be confounded by uncontrolled factors. In what follows, we outline findings indicating that the slower reading of coerced expressions is *not* due to uncontrolled semantic properties, differences in cloze probabilities, or differences in co-occurrence frequencies, and that the effect clearly reflects the slower computation of a coerced interpretation.

Contrasts such as (56) and (57) are not fully synonymous. Minimally, they differ in aspectual properties. However, Pickering et al. (2005) found the same coercion effect when (56) was compared to (58), which overtly expresses the interpretation that comprehenders report giving to the coerced form.

(58) The carpenter began building the table during the morning break.

Eye tracking measures showed that control expressions such as (57) and (58) did not differ at *the table* or beyond, but readers spent longer processing *the table* in (56) than both (57) and (58).

Another concern is that eventive verbs such as *begin* could be generally more semantically complex than control verbs such as *build* (c.f. Gennari & Poeppel, 2003), and hence they might be more taxing to process with any type of complement. A related concern is that cloze probabilities are likely to be higher for an expression such as *built the table* than for *began the table*. Indeed, in the Traxler et al. (2002) materials, the average cloze values were .03 for coerced expressions such as (56) and ranged from .14 to .19 for controls such as (57). These values are all quite low, well below the range that typically affects eye-tracking measures (Rayner & Well, 1996), but they do trend in the direction of the processing cost.

Traxler et al. (2002) addressed both concerns by using both self-paced and eye-tracking measures to examine the processing of eventive verbs (e.g., *started*) paired with entity-denoting NP complements (e.g., *the puzzle*) and event-denoting NP complements (e.g., *the fight*). *The boy started the puzzle* was more difficult to process than *The boy started the fight*, despite the fact that cloze values were nearly identical (.03 and .02 for event and entity nouns). Moreover, entity-denoting NPs and event-denoting NPs did not differ as the complements of verbs such as *saw* (e.g., *The boy saw the puzzle/fight*), which can

semantically combine with either argument type. The form of this interaction indicates that the observed processing cost is localized to the combination of an eventive verb and entity-denoting complement, and this strongly suggests the effect is linked to the kind of compositional operations comprehenders use to interpret the VP.

Cloze measures may not reflect subtler differences in the frequency with which various constituents co-occur in the language user's experience. Corpus analysis is not useful for assessing the frequencies of particular verb-complement pairings, as the data for many pairings of constituents will be sparse. One solution is to use similarity measures derived from Latent Semantic Analysis (LSA) (Landauer, Foltz, & Laham, 1998), which gives an approximation to co-occurrence—the cosine between pairs of text—reflecting the degree to which the constituents appear in similar contexts even if they never appeared together. Using the Traxler et al. materials, we estimated the co-occurrence of event- and entity-denoting NPs (e.g., *fight* versus *puzzle*) following the subject and either an eventive verb (e.g., *The boy started…*) or a verb that selects either an event or entity complement (e.g., *The boy saw…*). From a 300-factor LSA database reflecting usage up to 1st year college, the mean cosine for the entity-denoting NPs was estimated at 0.285 (sd=.145) following eventive verbs and 0.280 (sd=.167) following neutral verbs. Overall, event-denoting NPs had *lower* cosine values than entity-denoting NPs, 0.178 (sd=.131) following eventive verbs and 0.169 (sd=.161) following neutral verbs. This analysis suggests that the observed differences cannot be attributed to co-occurrence patterns, as the results are exactly opposite this type of account: The condition that showed the greatest cost—the one which required complement coercion (e.g., *The boy started the puzzle*)—had the highest cosine value (.285).

Coercion effects also appear to generalize to other types of NPs and to languages other than English. Traxler et al. (2002) found a coercion cost for complements with both indefinite and definite NPs (*a book* and *the book*). This suggests that the effect cannot be attributed to the particular types of pragmatic accommodations that might be required by a definite NP. McElree, Frisson, and Pickering (in press) found reliable coercion effects with proper nouns, …*began Dickens* versus ...*met/read Dickens*. Eye movement measures during reading indicated that metonymies (…*read Dickens*) were not more costly to interpret than conventional expressions (…*met Dickens*), but expressions that required coercing *Dickens* into an event were more taxing to interpret than both. Scheepers, Mohr, Keller, and Lapata (2004) replicated the contrast between *began/read the book* in German, with materials tightly matched on overall plausibility ratings and the predictability of the object noun.

Finally, a recent speed-accuracy tradeoff (SAT) study indicates that reading time measures reflect differences in the time needed for comprehenders to build an event interpretation of the complement (McElree, Pylkkänen, Pickering, & Traxler, in press). A notable short-coming of reading time measures is that longer reading times might indicate that readers take longer to interpret one expression than another *or* that readers are simply less likely to accurately process all the information necessary for an interpretation (McElree, 1993; McElree & Nordlie, 1999). The SAT procedure provides a means of directly measuring processing time in the presence of concomitant differences in accuracy (Bornkessel, McElree, Schlesewsky, & Friederici, 2004; McElree, 1993; McElree & Nordlie, 1999).

McElree et al. (in press) used the SAT procedure to investigate the processing of materials such as (56) and (57) above, as well as constructions such as (59) and (60).

(59) The climber proved the ice survivable. (coerced)
(60) The climber proved the fall survivable. (control)

These additional constructions were formed around verbs that select for small clause complements, consisting of an NP subject and an adjectival predicate. Crucially, the adjectives (e.g., *survivable*) were morphologically derived from eventive verbs. As such, they semantically required eventive subject NPs, just as eventive verbs require eventive complements (see Section 3.2.2.1). Composition should be simpler when a subject NP denotes an event (e.g., *the fall*) than when it does not (e.g., *the ice*), as the adjective in the latter case should trigger coercion of the NP into an eventive interpretation (e.g., "climbing of the ice").

In a trial of the SAT procedure, a sentence was presented phrase-by-phrase, and the participants were required to decide whether a critical constituent represented a sensible continuation of the sentence. Participants were trained to respond to a response signal—a tone—presented at several times following the onset of the critical expression (*the table* in examples such as (56) and (57) above, or *survivable* in examples such as (59) and (60)). The response signal was varied across a range of times (0-4900 ms) to fully measure how the different interpretations unfolded over time.

Coerced expressions yielded lower overall levels of performance than minimally contrasting controls, whether coercion was triggered by an event-selecting verb (V-NP) or adjective (NP-AP). This suggests that comprehenders were less likely to compute a sensible eventive interpretation of the subject or object NP when coercion was required. More

importantly, however, measures of how the interpretations unfolded over time demonstrated that comprehenders computed plausible interpretations of the coerced expressions more slowly than control expressions (123 ms slower in complement coercion and 158 ms slower in NP-AP constructions). These measures provide direct evidence that comprehenders required more time to construct a sensible interpretation of the coerced complement.

3.2.3.2.*Locus of the effect*

Why do comprehenders require more time to interpret coerced complements? One explanation of these differences is that they simply reflect disruptions arising from the detection of a semantic-type mismatch. There are at least two reasons to question this account. First, it leaves unanswered the question of how comprehenders actually succeed in recovering a suitable interpretation of these expressions. Second, although disruption in processing does occur when an interpretation does not make sense (e.g., Clifton, 1993; Pickering & Traxler, 1998; Traxler & Pickering, 1996), this type of explanation is not completely consistent with the delayed and sustained effects found in the reading time studies. One might have expected a mismatch effect to be evident when readers first fixate the complement, but such effects have not been found in any of the reading time studies.

Data from a recent magnetoencephalographic (MEG) study of complement coercion provides a more direct means of examining this hypothesis (Pylkkänen, Llinás, & McElree, 2004; Pylkkänen, Llinás, & McElree, submitted). Crucially, complement coercion does not modulate the same brain activity found in clear cases of mismatching semantic relations between a verb and its complement. Relative to control expressions such as *The author wrote the book*, anomalous expressions such as *The author amused the book* increased the

activity in a left temporal source at 300-400ms (M350)—the MEG analogue of an N400 event-related potential (ERP) component. However, *The author began the book* generates the same activity levels in this source as the control expression, *The author wrote the book*. Complement coercion modulated a ventromedial prefrontal source in a later 350-500ms time-window, generating more activity in this source than either the anomalous or simpler control sentences. These findings suggest that the increased activity associated with complement coercion must involve more than simple detection of mismatching semantic properties.

Another explanation of the processing time difference attributes it to the time needed to retrieve or infer the activity implicit in the event interpretation of the coerced complement. Traxler, McElree, Williams, and Pickering (2005) tested this notion by investigating the effects of context. The idea was to place the required activity in the preceding context under the assumption that this would eliminate the cost if the difficulty involved retrieving or inferring an appropriate activity. For example, *The carpenter began the table* is most often interpreted as "The carpenter began to build the table." Traxler et al (2005) examine whether the coercion cost would hold in contexts such as *The carpenter was building all morning. Before he began the table...*

Importantly, this type of context manipulation did *not* eliminate the coercion cost. This speaks against attributing the cost to the time needed to retrieve the action implicit in the event sense. The findings also speak against attributing the cost to *selecting* an activity from a set of plausible actions. If expressions such as *began the book* are taxing because it is possible to interpret them in several ways [e.g., began reading, writing, translating (etc) the book], a constraining context should have reduced the ambiguity and eliminated the cost

(c.f. Altmann & Steedman, 1988; Binder & Morris, 1995; Hess, Foss, & Carroll, 1995; Pickering & Traxler, 1998).[9]

These results are consistent with the idea that coercion is more taxing because comprehenders must undertake more complex compositional operations to build a representation for the event sense of the complement. In other experiments, Traxler et al found that the cost was eliminated if the context provided the full event sense of the complement (e.g., *The student started/read a book in his dorm room. Before he started the book…* or *The student started/read a book in his dorm room. Before he started it…*). These experiments show that the cost of coercion can be circumvented if a relevant event sense is in the immediate discourse. Traxler et al. suggest that comprehenders were able to circumvent the costly act of building an event sense for the target expression by linking its interpretation to a relatively abstract event representation in the discourse.

In summary, the evidence indicates that these expressions are costly to process because comprehenders need to engage additional, extra-syntactic processes to generate an interpretation of the mismatched constituents. The evidence is most consistent with the idea that the cost arises from the on-line composition of an event structure.

3.2.4. *Aspectual coercion*

Both complement coercion and type-mismatch in quantifier interpretation involve semantic mismatches between a predicate and its argument. The grammaticality of these cases suggests that natural language has some set of mechanisms for fitting an argument to meet the selectional restrictions of its predicate. Does this also occur in modification?

---

[9] Traxler et al. (2005) provide other arguments against both of these alternative explanations.

Let us first revisit cases where predicate modification applies straightforwardly. In (61), the predicates (a-e) *grey*, *furry*, *four-legged* and *creature* are all intersected to yield one complex predicate of type <e,t>. This works because all the predicates are predicates of individuals, so they are functions that require as their input the same type of thing. Similar reasoning applies in the verbal domain. In (62), the VP *Brutus stabbed Caesar* and the modifiers (a-e) *violently*, *in the back* and *with a knife* are all predicates of events (see section 2.2) and can therefore happily intersect. In contrast, mixing predicates that are unambiguously predicates of individuals with predicates that are unambiguously predicates of events results in anomaly, as shown in (63).

(61) Fido is a grey, furry, four-legged creature.
Entailments:

a. Fido is grey.
b. Fido is furry.
c. Fido is grey.
d. Fido is four-legged.
e. Fido is a creature.
Etc…

(62) Brutus stabbed Caesar violently in the back with a knife.
a. Brutus stabbed Caesar violently.
b. Brutus stabbed Caesar in the back.
c. Brutus stabbed Caesar with a knife.
d. Brutus stabbed Caesar violently with a knife.
Etc…

(63) #Fido is four-legged violently.

Now consider (64), which several researchers have argued to involve a type-mismatch (Chierchia, 1984; Jackendoff, 1997; Klein & Sag, 1985; Partee & Rooth, 1983; Pustejovsky, 1991, 1995). Here the punctual verb *jump* is modified by a durative adverb. One might expect this combination to be ungrammatical, but instead, comprehenders appear

to invoke a process or iterative interpretation of the verb *jump*. This reinterpretation is generally called 'aspectual coercion'.

(64) The girl jumped for three hours.

Crucially, however, (64) does not involve a type-mismatch in the framework we have laid out, as both *jump* and *for three hours* are predicates of events. Consequently, they should be able to intersect. The problem though is that such an intersection does not yield an iterative reading but rather the transparent meaning in (65), which asserts that there exists a single jumping-event that lasted for three hours. This obviously clashes with our real-world knowledge about people's jumping abilities.

(65) $\exists e$ [jumping(e) & agent(e, the girl) & for(e, three hours)]

Thus, in the framework we have described, aspectual coercion and complement coercion are crucially different. In complement coercion, transparent semantic composition is impossible. In aspectual coercion, transparent semantic composition is possible but it yields an anomalous output. Since sentences involving aspectual coercion are nevertheless judged felicitous, there must be some process that repairs this anomaly and changes the sort of the event described by the verb from a punctual event to a process. Given this difference between complement and aspectual coercion, comparison of their online processing is clearly of great interest to theories of the syntax-semantics interface. In what follows we first review psycholinguistic results on aspectual coercion and then return to representational issues pertaining to coercion and type-shifting more generally.

3.2.5. *Online processing of aspectual coercion*

To test whether aspectual coercion imposes a significant processing load, Piñango, Zurif, and Jackendoff (1999) examined cross-modal lexical decisions to unrelated words at different points during the reading of sentences such as (66) and (67).

(66) The insect hopped effortlessly until it reached the far end of the garden…
(67) The insect glided effortlessly until it reached the far end of the garden…

At the durative adverb *until*, lexical decision times were longer for punctual verbs such as *hop* than for temporally unbounded verbs such as *glide*. They argued that this would not be the case if verbs such as *hop* were lexically polysemous between point-action and unbounded-activity interpretations. They suggest that the increased load reflects the deployment of on-line operations to generate a non-lexicalized, iterative sense of the verb.

Todorova, Straub, Badecker & Frank (2000) pointed out some limitations of the Piñango et al. study. One concern was that all the sentences in the critical condition had iterative interpretations, whereas sentences in the control condition did not. It is possible that iterative readings *per se* impose a higher processing burden than noniterative readings.

Todorova et al. employed conditions illustrated in (68) – (71).

(68) Even though Howard sent a large check to his daughter for many years, she refused to accept his money.
(69) Even though Howard sent large checks to his daughter for many years, she refused to accept his money
(70) Even though Howard sent a large check to his daughter last year, she refused to accept his money
(71) Even though Howard sent large checks to his daughter last year, she refused to accept his money

Sentence (68) illustrates the condition predicted to engage more complex compositional processes, as it included a punctual verb (*send*) and a singular direct object (*a large check*) followed by a durative adverb (*for many years*). Todorova et al. point out that when a bare

plural direct object (*large checks*) is an argument of the verb, as in (69), it appears to impose an iterative reading on the predicate immediately. Sentences such as (69) served as one contrast to (68). They also employed two other controls, which were analogues of (68) and (69) with adverbials (*last year*) compatible with both iterative and noniterative predicates. Processing was measured with a "stop-making sense" procedure (e.g., Boland, Tanenhaus, Garnsey, & Carlson, 1995; Mauner, Tanenhaus, & Carlson, 1995), in which participants evaluated whether the sentence continued to make sense as they paced themselves through the sentence phrase by phrase.

Results indicated that participants were over twice as likely to reject sentences such as (68) than any of the control sentences (19% versus 7%, 8%, and 9%), and significantly more likely to reject these sentences at the onset of the region containing the durative modifier. Additionally, the associated self-paced times were significantly longer at the durative modifier for (68) as compared to (69), with no comparable differences evident in (70) and (71).

These findings indicate that iterative readings *per se* are not more taxing to process than noniterative readings, which alleviates one of the concerns with the Piñango et al. study. Importantly, like the Piñango et al. study, a processing cost was evident in (68) where the predicate was incompatible with the durative adverb. This pattern, however, is not consistent with the Piñango et al. claim that the source of the difficulty lies with the lexical representation of the verb. Punctual verbs paired with bare plurals (*...sent large checks*) in (69) were not difficult to interpret. This finding indicates that comprehenders can assemble without cost an interpretation compatible with the durative adverb when the verb is paired with an appropriate argument (see 3.3 for further discussion). Todorova et al. suggest

that comprehenders commit to an aspectual interpretation of the predicate only after both the verb and its arguments have been processed. This result suggests that aspect is not a lexical property stored with the verb, but instead is derived compositionally using features from the argument (see also Seegmiller, Townsend, DeCangi and Thomas, 2004).

Do these results implicate a coercion process for aspectual mismatches analogous to what has been found with complement coercion? Conceivably they might if aspectual coercion is accomplished by introducing new semantic structure, such as an iterative operator that is not implicit in the verb's representation. However, as outlined in Section 3.2.4, there are formal reasons to doubt that the same compositional processes underlie the two phenomena. Additionally, there are other salient differences in the on-line experiments. Note that aspectual studies show evidence of a cost *only* when an interpretation of the predicate turns out to be incompatible with later material in the sentence, namely the durative modifier. In this respect, it appears more correct to view these effects as a type of semantic "garden-path." In contrast, the cost in complement coercion appears to reflect the time it takes to establish an initial interpretation of the complement, not whether that interpretation is compatible with some material occurring later in the sentence. Crucially, if aspectual coercion were processed like complement coercion, effects should have been evident in cases like (69), where comprehenders were forced to unify the argument *large checks* with a "mismatching" punctual verb such as *sent*.

There are also other empirical grounds on which to question the commonality of the processes. Crucially, aspectual coercion does *not* engender any detectable differences in standard reading tasks (Pickering, McElree, Chen, Traxler & Frisson, in prep). We conducted two self-paced reading studies and one eye-tracking study with the Piñango et al.

materials and found no measurable differences in self-paced times or any aspect of the eye movements for the conditions represented in (66) and (67). Additionally, we collected self-paced reading times and eye-tracking measures for these sentences when the adverbial phrase (PP) was fronted, as in (72) and (73).

(72) Until it reached the garden, the insect hopped effortlessly…
(73) Until it reached the garden, the insect glided effortlessly…

If Piñango et al's arguments are correct, there should be a coercion cost associated with generating an iterative sense of *hop.* Hence, analogous to cases of complement coercion, there should be evidence of difficulty (longer times or more regressive eye movements) at and after *hopped* as compared to *glided*. We found no indication of any difference.

Finally, we examined the Todorova et al materials, (68) – (71) above, in another eye-tracking study experiment. Again, there were no differences evident in any of the conditions. Importantly, all these experiments included constructions involving complement coercion, as part of other ongoing research. In each case, we observed reliable effects of complement coercion. Hence, there is no concern that the null results for aspectual "coercion" reflect an aberrant participant population or faulty measurement procedures.

3.2.6. *Neurolinguistic studies of coercion*

Another way to investigate whether aspectual and complement coercion engage similar interpretive operations is to study their respective brain bases. Piñango and Zurif (2001) investigated the comprehension of both types of coercion by Wernicke's and Broca's aphasics. In the aspectual coercion experiment, three Wernicke's aphasics and three Broca's aphasics listened to coerced and transparent sentences such as *the horse jumped for an hour yesterday* and *the horse jumped over the fence yesterday*, respectively. Each sentence was

followed by a comprehension question querying whether the horse jumped once or many times. Wernicke's aphasics performed at chance for coerced sentences and at 87% accuracy for transparent sentences. In contrast, Broca's aphasics performed significantly above chance for both conditions and showed no effect of sentence type. Given the small number of subjects, it is important to note that Broca's aphasics' performance was in fact numerically better on the coerced sentences than on the transparent sentences. Thus, it appears that left temporal areas are indeed important for aspectual coercion, in a way that Broca's area is not. Crucially, however, these deficit-lesion data do not discriminate between the hypothesis that aspectual coercion *occurs* in Wernicke's area and the hypothesis that Wernicke's area *provides crucial input* for aspectual coercion.

For complement coercion, two Wernicke's and three Broca's aphasics again listened to coerced (*The boy began reading the book*) and transparent sentences (*The boy began the book*) but now in a picture-matching task. The picture depicted either the correct scenario of a boy reading a book or an incorrect scenario of a boy buying a book. Contra the aspectual coercion experiment, in this study both patient groups performed numerically worse on coerced than transparent sentences although this effect only reached significance for the two Wernicke's aphasics. Thus these data are more suggestive of a main effect of coercion than an interaction. Nevertheless, the authors concluded "that aspectual and complement coercion involve computations requiring the integrity of the left posterior cortical region associated with Wernicke's area, but not the integrity of the left anterior cortical region associated with Broca's area" (idib, p. 307). Although the aspectual coercion data are quite consistent with this hypothesis, the conclusion is premature for complement coercion.

Further, the complement coercion task may be difficult for Wernicke's aphasics for reasons that are independent of coercion. In the author's procedure, the auditory sentence and the two pictures were presented simultaneously, which means that the pictures had potentially been processed by the time the complement of the verb occurred. Consequently, at the coercion site, the listener may have had the conceptual representations of both reading and buying active. Interpretation of *the boy began the book* would then require inhibition of the competing event description. Given this, a patient population such as Wernicke's aphasics, who have general problems with inhibitory processing (Milberg, Blumstein, Katz, & Gershberg, 1995; Wiener, Connor & Obler, 2004), might perform poorly even if their coercion mechanism was intact. Further, interpretation of *the boy began the book* as *the boy began buying the book* is not entirely impossible. Consider, for example, an auction setting where buying is construed as a process. The fact that the "buying" interpretation is not strictly ungrammatical might increase the likelihood of patients with inhibitory problems to choose that scenario.

Although deficit-lesion methodology can tell us which areas are necessary for a certain process, identifying the brain areas responsible for that process requires methods that allow localization of function in the intact brain. As noted in Section 3.2.3.2, Pylkkänen, Llinás and McElree (2004, submitted) investigated MEG responses to coerced sentences such as *the author began the book* and transparent sentences such as *the author wrote the book.* Contra Piñango and Zurif's hypothesis that complement coercion occurs in Wernicke's area, left temporal sources showed no effect of coercion. Instead, coercion elicited larger amplitudes in a ventromedial prefrontal area. This area is known to receive direct input from left superior temporal cortex (Rolls, 2004). Thus these results would in fact

be entirely consistent with Wernicke's aphasics having trouble with complement coercion. In our study, the area that was found to be sensitive to coercion is clearly not a traditional language area. Thus our results suggest that semantic non-compositionality might in some cases engage processes that are not specifically "linguistic."

3.3. TOWARDS A TAXONOMY OF TYPE-SHIFTING RULES

We have discussed several cases of nontransparent composition. Experimental investigation of these phenomena has only recently begun, and clearly many questions remain. Nonetheless, as we have emphasized, we believe that the study of the online processes of semantic composition should be guided by the rich and detailed hypothesis space provided by theoretical semantics. To that end, we briefly revisit each of the phenomena in Section 3 to evaluate them on representational grounds. To what extent can we predict similar or different processing operations based on what we know and can reason about their representation?

One way to classify type-shifting rules would be along the following two dimensions:

A. Does the shift change the complexity of the type?

B. Does the shift invoke an ontological change?

By 'ontological change' we mean a change in the basic ontological category that the word or phrase denotes or predicates over. For example, complement coercion involves an ontological change from an *individual* <e> to a predicate of *events* <s,t>.

Type-shifting rules such as those in quantifier interpretation are shifts that change the complexity of the type but do not invoke an ontological change. For example, the sole

purpose of Partee and Rooth's (1993) 'lift' rule, repeated in (74), is to change the directionality of functional application: an individual that would be interpreted as an argument of a function by default is converted to a higher-order predicate that can take a verb as its argument. This introduces no new semantic content, hence no deep change in the meaning of the constituent.

(74) Lift (enables an individual to take an intransitive verb as its argument)
Sally → $\lambda f_{<e,t>}$. f(Sally)

e → <<e,t>, t>

Complement coercion invokes an ontological shift *and* a change in type complexity. Therefore, one might expect it to be more costly than pure type-shifting rules such as lift.

(75) Complement coercion
the book → λe. θ(e, the book)

e → <s,t>

How does aspectual coercion fit into this framework? As discussed in section 3.2.4, aspectual coercion does not involve a type-mismatch in the framework we have proposed, as our ontology does not distinguish between events on the basis of their temporal properties. Research on the internal temporal properties of events, called *aktionsart* or *situation aspect*, has a long history, going back to Aristotle. However, the precise way these properties should be incorporated into a type-driven system of interpretation has not been fully worked out.

Our knowledge about the internal temporal structure of events is largely knowledge of the world. Understanding the interaction of linguistic knowledge and world knowledge is no trivial matter, but an analysis of aspectual coercion requires us to make specific assumptions about it. Pustojevsky (1991, 1995), whose research has informed and inspired much of the psycholinguistic work on coercion, does not distinguish between linguistic and

world knowledge but rather builds them both into complex lexical entries. In a series of papers, Dölling (1992, 1993, 1995, 1997, 2003) has developed an alternative approach that treats world knowledge as presuppositions on lexical entries. Dölling introduces the notation in (76) for this purpose. Here, for example, the adjective *widespread* is treated as a function from individuals to truth-values, but the sorts of individuals that constitute a felicitous input for the function are restricted to kinds (e.g., *the tiger*). If the input does not denote a kind, a presupposition failure results: #*John is widespread*.

(76) λx: x is a kind. widespread(x)

In Dölling's theory, these types of lexical entries figure into a two-stage model of interpretation, in which strictly compositional interpretation is followed by the fixing of various contextual parameters on the basis of world knowledge. An integral part of the second stage is the application of a rich set of productive *sort coercions*. For example, the sortal restrictions of the verb in (77) cause the semantic sort of the subject NP to shift from an institution to a person.

(77) The newspaper telephoned.

This type of shifting has generally been called metonymy, and Dölling is by no means the first or only person to have systematically discussed it. Although it is questionable whether the two-stage aspect of Dölling's theory is strictly applicable to real-time processing, what is appealing about Dölling's approach from a representational standpoint is that it incorporates lexical-semantic properties within a standard type-driven model of sentence-level interpretation whose basic operations are fully compositional.

In the context of this framework it is natural to conceptualize aspectual coercion as a sortal shift within the basic category of events. How might this relate to the observed

processing differences? The lack of reading time effects for aspectual coercion comparable to complement coercion suggests that sortal shifts are, unequivocally, simpler or easier for the processing system to deal with than ontological shifts. Indeed, the contrast between aspectual and complement coercion parallels a similar contrast between metonymy and complement coercion, which likewise suggests that other types of sortal shifts are not costly. Metonymic expressions such as …*read Dickens*, where *Dickens* must be interpreted as any of a number of books written by Charles Dickens, are not more taxing to process than literal expressions such as …*met Dickens*, where *Dickens* can be interpreted as the person Charles Dickens (Frisson & Pickering, 1999). This appears to be true of a wide range of shifts, including PLACE-FOR-EVENT (e.g., *protested during Vietnam*) and PLACE-FOR-INSTITUTION (e.g., *talked to the school*) metonyms (Frisson & Pickering, 1999). Crucially, as noted previously, McElree et al. (in press) showed that coercing *Dickens* to mean reading the works of Dickens, e.g., *began Dickens,* is costly relative to both ...*met Dickens* or ...*read Dickens*.

We suggest that the fundamental reason why aspectual coercion and other types of metonymy are not as taxing to process as complement coercion is that sortal shifts do not block initial composition. In contrast, the view developed here is that the ontological mismatch, such as those found in complement coercion, does block basic composition, and hence immediately mandates the on-line deployment of a costly coercion process.

The Todorova et al. (2000) study suggest that comprehenders do not have difficulty interpreting a predicate in either an iterative or process manner; rather, difficulty occurs *only* when the interpretation turns out to be incompatible with a durative modifier downstream. In this respect, the pattern is similar what has been found in investigations of collective and

distributed NP interpretations. Frazier, Pacht, and Rayner (1999) used eye-tracking procedures to investigate collective and distributed NP interpretations in structures such as (78)-(79).

(78) Lyn and Patrick *each* saved 1000 dollars to pay for their honeymoon.
(79) Lyn and Patrick *together* saved 1000 dollars to pay for their honeymoon.
(80) Lyn and Patrick saved 1000 dollars *each* to pay for their honeymoon.
(81) Lyn and Patrick saved 1000 dollars *together* to pay for their honeymoon.

When disambiguating evidence—either *each* or *together*—occurred immediately after the coordinate NP, there was no difference between (78) and (79). This indicates that distributed readings are as easy to compute as collective readings. However, when the disambiguating evidence occurred late, there was clear evidence for a bias for a collective reading of the coordinate NP: (80) was more difficult than (81). Apparently, comprehenders do not commit to a collective reading of the coordinate NP until after the predicate has been processed. In cases such as (80), this creates a semantic garden-path at *each*, which requires reinterpreting the subject and the predicate to derive the distributive reading.

Aspectual coercion may differ from collective/distributed interpretations and from complement coercion in that comprehenders might be able to underspecify aspectual interpretations to a greater degree. Underspecification has been argued to be possible in several domains, including quantifier scope relations (Filik et al. 2004; see Sanford & Sturt, 2002) and the processing of polysemous and metonymic expressions (e.g., Frazier & Rayner, 1990; Frisson & Pickering, 1999, 2001; Pickering & Frisson, 2001). The reading times result on aspectual coercion suggest that it might be possible for comprehenders to generate a conceptually underspecified representation that does not fix all of the contextual parameters of the predicate's interpretation. Comprehenders may be less inclined to do so in

tasks that involve intrusive decisions, such as acceptability judgments (Todorova et al., 2000) or secondary tasks (Piñango et al, 1999). In short, comprehenders may need to be pressured to make commitments to see the effects of sortal incompatibilities.

4. CLOSING COMMENTS

Our goal in this chapter has been to propose a framework for investigations of semantic composition. Just as formal work in syntactic theory has informed theories of parsing, we believe that the rich and substantial theorizing in semantics has the potential to guide and stimulate psycholinguistic and neurolinguistic research on real-time semantic composition.

5. REFERENCES

Abney, S. P. (1989). A computational model of human parsing. *Journal of Psycholinguistic Research*, 18, 129–144.

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

Altmann, G.T.M., & Steedman, M.J. (1988). Interaction with context during human sentence processing. *Cognition*, *30*, 191-238.

Barker, C. (2002). Continuations and the nature of quantification. *Natural Language Semantics*,10, 211-242.

Binder, K.S., & Morris, R.K. (1995). Eye movements and lexical ambiguity resolution: Effects of prior encounter and discourse topic. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 1186-1196.

Boland, J. E. (2005). Visual Arguments. *Cognition*, 95, 237-274.

Boland, J. E. & Boehm-Jernigan, H. (1998). Lexical constraints and prepositional phrase attachment. *Journal of Memory and Language*, 39, 684-719.

Boland, J., Tanenhaus, M., Garnsey, S. (1990). Evidence for the immediate use of verb control information in sentence processing, *Journal of Memory and Language* 29, 413–432.

Boland, J., Tanenhaus, M., Garnsey, S., & Carlson, G. (1995). Verb argument structure in parsing and interpretation: Evidence from Wh-questions. *Journal of Memory and Language*, 34, 774–806.

Bornkessel, I., McElree, B., Schlesewsky, M., & Friederici, A. D. (2004). Multi-dimensional contributions to garden path strength: Dissociating phrase structure from relational structure. *Journal of Memory and Language*, 51, 495-522.

Briscoe, E., Copestake, A., & Boguraev, B. (1990). Enjoy the paper: Lexical semantics via lexicology. *Proceedings of the 13th International Conference on Computational Linguistics* (pp. 42-47), Helsinki, Finland.

Carlson, G. & Tanenhaus, M. (1988). Thematic roles and language comprehension. In W. Wilkins (Ed.), *Thematic relations*. New York, NY: Academic Press.

Chierchia, G. (1984). *Topics in the syntax and semantics of infinitives and gerunds*. Unpublished Ph.D. dissertation, University of Massachusetts at Amherst.

Clark, H., & Haviland, S. (1977). Comprehension and the given-new contrast. In R. Freedle (Ed.), *Discourse production and comprehension.* (pp. 1-40). Hillsdale, NJ: Erlbaum.

Clifton Jr., C. (1993). Thematic roles in sentence parsing. *Canadian Journal of Experimental Psychology*, 47, 222-246.

Clifton, C., Jr., Speer, S., & Abney, S. (1991). Parsing arguments: Phrase structure and argument structure determinants of initial parsing decisions. *Journal of Memory and Language*, 30, 251–271.

Phillips, C. (2001). Mechanisms for rapid use of focus information: Review of Sedivy (1997). *GLOT International*, 5, 25-32.

Davidson, D. (1967). The logical form of action sentences, in N. Reschler (ed.), *The Logic of Decision and Action*, University if Pittsburgh Press, Pittsburgh, pp. 81-95.

Dölling, J. (1992). Polysemy and sort coercion in semantic representations. In P. Bosch and P. Gerstl (Eds.) *Discourse and Lexical Meaning*. Proceedings of a Workshop of the DFG Sonderforschungsbereich 340, Stuttgart, 61-78.

Dölling, J. (1993). Commonsense ontology and semantics of natural language. *Zeitschrift für Sprachtypologie und Universalienforschung,* 46, 133-141.

Dölling, J. (1995). Ontological domains, semantic sorts and systematic ambiguity. *International Journal of Human-Computer Studies* 43, 785-807.

Dölling, J. (1995). Semantic form and abductive fixation of parameters. In R. van der Sandt, R. Blutner and M. Bierwisch (Eds.) *From Underspecification to Interpretation*, 113-138. Heidelberg: IBM Deutschland.

Dölling, J. (2003). Aspectual (re-)interpretation: Structural representation and rrocessing. In H. Härtl & H. Tappe (Eds.) *Mediating between Concepts and Grammar*. Berlin-New York: Mouton - de Gruyter, pp. 303-322.

Egg, M (2003). Beginning novels and finishing hamburgers - remarks on the semantics of to begin. Journal of Semantics 20, 163-191.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, 47, 164-203.

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

Filik, R., Paterson, K.B., & Liversedge, S.P. (2004). Processing doubly quantified sentences: Evidence from eye movements. *Psychonomic Bulletin & Review*, 11, 953-959.

Fodor, J. D. (1982). The mental representation of quantifiers. In S. Peters & E. Saarinen (Eds.). *Processes, Beliefs, and Questions*, Dordrecht: D. Reidel. 129-164

Fodor, J. A. & Lepore, E. (2002). *The Compositionality Papers*. Oxford: Clarendon Press.

Frazier, L. (1999). *On sentence interpretation*. Dordrecht: Kluwer Academic Publishers.

Frazier, L. & Clifton Jr., C. (1996). *Construal*. Cambridge, MA: MIT Press.

Frazier, L., & Rayner, K. (1990). Taking on semantic commitments: Processing multiple meanings vs. multiple senses. *Journal of Memory and Language*, 29, 181-200.

Frazier, L., Pacht, J. M., & Rayner, K. (1999). Taking on semantic commitments, II: collective versus distributive readings. *Cognition*, 70, 87-104.

Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosopic und Philosophische Kritif,* 100, 25-50; trans by M. Black as 'On sense and reference', in P. Geach, & M. Black (Eds.) *Translations from the Philosophic Writings of Gottlob Frege*, Basil Blackwell, Oxford, 1960, 56-78.

Frisson, S. & Pickering, M.J. (2001). Figurative language processing in the Underspecification Model. *Metaphor and Symbol*, 16, 149-171.

Frisson, S. & Pickering, M. J. (1999). The processing of Metonymy: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 1366-1383.

Gennari, S., & Poeppel, D.,(2003). Processing correlates of lexical semantic complexity. *Cognition*, 89, B27–B41.

Gillen, K. (1991). *The comprehension of doubly quantified sentences*. Unpublished Ph.D. dissertation, University of Durham.

Glucksberg, S. (2001). *Understanding figurative language: From metaphors to idioms.* New York, NY: Oxford University Press.

Hagoort, P., Hald, L., Bastiaansen, M., Petersson, K.M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, 304, 438-41.

Heim, I. & Kratzer, A. (1998). *Semantics in Generative Grammar*. Oxford: Blackwell Publishers.

Hendriks, H. (1988). Type Change in Semantics: the Scope of Quantification and Coordination. In Klein, E. & Van Benthem, J. (Eds.) *Categories, Polymorphism and Unification*, ITLI, Amsterdam, 96–119.

Hess, D.J., Foss, D.J., & Carroll, P. (1995). Effects of global and local context on lexical processing during language comprehension. *Journal of Experimental Psychology: General*, *124*, 62-82.

Jackendoff, R. (2002). *Foundations of language*. New York, NY: Oxford University Press.

Jackendoff, R. (1997). *The architecture of the language faculty*. Cambridge, MA: MIT Press.

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

Jacobson, P. (1999). Towards a Variable-Free Semantics, *Linguistics and Philosophy*, 22, 117-184.

Kamp, H. and Partee, B.H. (1995). Prototype theory and compositionality. *Cognition*, 57, 129-191.

Kennison, S. M. (1999). Processing Agentive "by"-phrases in complex event and nonevent nominals. *Linguistic Inquiry*, 30, 502–508.

Kennison, S. M. (2002). Comprehending noun phrase arguments and adjuncts. *Journal of Psycholinguistic Research*, 31, 65-81.

Klein, E. & Sag, I. (1985): Type driven translation. *Linguistics and Philosophy* 8, 163-201.

Kratzer, A. (1996). Severing the External Argument from its Verb, in J. Rooryck & L. Zaring (eds.), *Phrase Structure and the Lexicon*. Dordrecht (Kluwer Academic Publishers).

Kratzer, A. (in press). *The Event Argument and the Semantics of Voice*. Cambridge, MA: MIT Press.

Kurtzman, H. S., & MacDonald, M. C. (1993). Resolution of quantifier scope ambiguities. *Cognition,* 48, 243-279.

Lapata, M., & Lascarides, A. (2003). A probabilistic account of logical metonymy, *Computational Linguistics, 29*, 263-317.

Lapata, M., Keller, F., & Scheepers, C. (2003). Intra-sentential Context Effects on the Interpretation of Logical Metonymy. *Cognitive Science*, 27, 649-668.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). Introduction to Latent Semantic Analysis. *Discourse Processes*, 25, 259-284.

Lascarides, A., & Copestake, A. (1998). Pragmatics and word meaning. *Journal of Linguistics*, 34, 387-414.

Levin, B. & Rappaport Hovav, M. (1995). *Unaccusativity: At the syntax–lexical semantics interface*, MIT Press, Cambridge, MA.

Liversedge, S.P., Pickering, M.J., Branigan, H.P., & van Gompel, R.P.G. (1998). Processing arguments and adjuncts in isolation and context: The case of by-phrase

ambiguities in passives. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 461-475.

Liversedge, S.P., Pickering, M.J., Clayes, E.L., & Branigan, H.P. (2003). Thematic processing in adjuncts. *Psychonomic Bulletin & Review*, *10,* 667-675.

Marcus, M., Santorini, B., and Marcinkiewicz, M. (1993). Building a Large Annotated Corpus of English: The Penn Treebank. *Computational Linguistics*, 19, 313-330.

McElree, B. (1993). The locus of lexical preference effects in sentence comprehension: A time-course analysis. *Journal of Memory and Language*, 32, 536-571.

McElree, B., Frisson, S., & Pickering, M.J. (submitted). Deferrred interpretations: Why starting Dickens is difficult but reading Dickens isn't.

McElree, B., & Nordlie, J. (1999). Literal and figurative interpretations are computed in equal time. *Psychonomic Bulletin & Review*, 6, 486-494.

McElree, B., Pylkkänen, L., Pickering, M.J., & Traxler, M. (in press). The time course of enriched composition. *Psychonomic Bulletin & Review*.

McElree, B., Traxler, M. J., Pickering, M. J., Seely, R. E., & Jackendoff, R. (2001). Reading time evidence for enriched semantic composition. *Cognition*, 78, B15-B25.

McKoon, G., & Macfarland, T. (2000). Externally and internally caused change of state verbs. *Language*, 76, 833–858.

McKoon, G., & Macfarland, T. (2002). Event templates in the lexical representations of verbs. *Cognitive Psychology*, 47, 1–44.

Milberg, W., Blumstein, S.E., Katz, D., & Gershberg, F. (1995). Semantic facilitation in aphasia: Effects of time and expectancy. *Journal of Cognitive Neuroscience*, 7, 33–50.

Montague, R. (1970). Universal grammar. *Theoria,* 36, 373-398.

Murphy, G. L. (1990). Noun phrase interpretation and conceptual combination. *Journal of Memory and Language*, 29, 259-288.

Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.

Natsopoulos, D. (1985). A verbal illusion in two languages. *Journal of Psycholinguistic Research*, 14, 385–397

Parsons, T. (1985). Underlying Events in the Logical Analysis of English. In E. LePore and B. McLaughlin (Eds.), *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, pp. 235-267. Oxford & New York: Blackwell.

Parsons, T. (1990). *Events in the Semantics of English: A Study of Subatomic Semantics.* MIT Press, Cambridge, MA.Parsons (1990).

Partee, B. H. & Rooth, M. (1983). Generalized conjunction and type ambiguity. In R. Bäuerle, C. Schwarze, & A. von Stechow, A. (Eds.). *Meaning, Use, and Interpretation of Language.* Berlin: Walter de Gruyter. 361-83.

Partee, B. H., ter Meulen, A., & Wall, R. E. (1990). Mathematical Methods in Linguistics. *Studies in Linguistics and Philosophy*, *Vol. 30*. Dordrecht: Kluwer Academic Publishers.

Pickering, M. J. & Frisson, S. (2001). The semantic processing of verbs: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 556-573.

Pickering, M.J., Frisson, S., McElree, B., & Traxler, M., (2004). Eye movements and semantic composition. In M. Carreiras and C. Clifton, Jr. (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERPs, and beyond*. Hove: Psychology Press.

Pickering, M. J., McElree, B., & Traxler, M. (2005). The difficulty of coercion: A response to de Almeida. *Brain & Language*, 93, 1-9.

Pickering, M.J., McElree, B., Chen, L., Traxler, M.J., & Frisson, S. (in prep). Underspecification in coercion.

Pickering, M.J., & Traxler, M.J. (1998). Plausibility and recovery from garden paths: An eye tracking study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 940-961.

Pylkkänen, L., Liinas, R., & McElree, B. (2004). Distinct effects of semantic plausibility and semantic composition in MEG. In E. Halgren, S. Ahlfors, M. Hämäläinen, and D. Cohen (Eds.), *Biomag 2004: Proceedings of the 14th International Conference on Biomagnetism*, Biomag: Boston, USA.

Pylkkänen, L., Liinas, R., & McElree, B. (submitted). Distinct effects of semantic plausibility and semantic composition in MEG.

Pylkkänen, L. & McElree, B. (March, 2004). Processing secondary predicates: Does locality matter? *Annual CUNY Sentence Processing Conference*, College Park, NY.

Piñango, M. M., Zurif, E., & Jackendoff, R. (1999). Real-time processing implications of enriched composition at the syntax-semantics interface. *Journal of Psycholinguistic Research,* 28, 395-414.

Piñango, M. M., & Zurif, E. (2001). Semantic operations in aphasic comprehension: Implications for the cortical organization of language. *Brain and Language*, 79, 297-308.

Pustejovsky, J. (1991). The syntax of event structure. *Cognition*, 41, 47–81.

Pustejovsky, J. (1995). *The Generative Lexicon*. Cambridge, MA: MIT Press.

Rayner, K., & Well, A.D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review*, 3, 504-509.

Rolls, E. T. (2004). The functions of the orbitofrontal cortex. *Brain and Cognition*, 55, 11-29.

Rothstein, S. (1998). *Events and grammar*. Dordrecht: Boston: Kluwer Academic Publishers.

Sauerland, U. and A. von Stechow. (2001) The syntax-semantics interface. In N. Smelser and P. Baltes (Eds.), *International Encyclopedia of the Social & Behavioural Sciences*, pp. 15412–15418. Pergamon, Oxford, Great Britain.

Sanford, A.J., & Sturt, P. (2002). Depth of processing in language comprehension: Not noticing the evidence. *Trends in Cognitive Sciences*, 6, 382-386.

Scheepers, C., Mohr, S., Keller, F., & Lapata, M. (2004). The cost of enriched composition: Eye-movement evidence from German. *17th Annual CUNY Sentence Processing Conference*, College Park, Maryland.

Schütze, C. T., & Gibson, E. (1999). Argumenthood and English prepositional phrase attachment. *Journal of Memory and Language*, 40, 409–431.

Sedivy, J.C, Tanenhaus, M.K., Chambers, C.G., & Carlson, G.N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, 71, 109-147.

Seegmiller, M.S., Townsend, D.J., DeCangi, D. & Thomas K. (2004). Comprehenders Use Event Structure to Develop Discourse-Level Representations. *Annual Architectures and Mechanisms for Language Processing Conference*, Aix-en-Provence, France.

Speer, S. R., & Clifton, C. (1998). Plausibility and argument structure in sentence comprehension. *Memory & Cognition*, 26, 965–978.

Spivey-Knowlton, M. & Sedivy, J. (1995). Resolving attachment ambiguities with multiple constraints. *Cognition*, 55, 227-267.

Tenny, C. and J. Pustejovsky. (2000). *Events as Grammatical Objects*. Stanford: CSLI Publications.

Todorova, M., Straub, K., Badecker, W., & Frank, R. (2000). Aspectual coercion and the online computation of sentential aspect. *Proceedings of the Twenty-second Annual Conference of the Cognitive Science Society*, (pp.3-8).

Traxler, M., McElree, B., Williams, R. S. & Pickering, M.J. (in press). Context effects in coercion: Evidence from eye-movements. *Journal of Memory and Language.*

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

Traxler, M.J., Pickering, M.J., & McElree, B. (2002). Coercion in sentence processing: Evidence from eye-movements and self-paced reading. *Journal of Memory and Language*, 47, 530–547.

Trueswell, J. C., -Tanenhaus, M. K., & -Kello, C. (1993). Verb-specific constraints in sentence processing: separating effects of lexical preference from garden-paths. Journal of Experimental Psychology: Learning, Memory, and Cognition, 19, 528-553.

Tunstall, S. L. (1998). *The interpretation of quantifiers: Semantics and processing.* Unpublished Ph.D. dissertation, University of Massachusetts, Amherst.

Wason, P. and Reich, S.S. (1979). A verbal illusion. *Quarterly Jounral of Experimental Psychology* 31, 591–597.

Wiener, D. A., Connor, L. T. & Obler, L. K. (2004).Inhibition and auditory comprehension in Wernicke's aphasia. *Aphasiology*, 18, 599-609.