Coercion without Lexical Decomposition: Type-Shifting Effects Revisited

ROBERTO G. DE ALMEIDA
Concordia University

VEENA D. DWIVEDI
Brock University

Remotely plausible theories are better than no theories at all.
Jerry Fodor (1975)

A key issue for both psycholinguistic experimentation and theoretical linguistics is the nature of the information used in interpreting indeterminate constructions. For example, one can ask how a sentence such as (1a) comes to have a meaning that parallels that of (1b). Does this reflect lexical-semantic operations (as often claimed in the literature) or does it arise via pragmatic-inferential processes? In this article, we argue for the latter view.

(1) a. Mary began the book.
    b. Mary began reading the book.

Sentences such as (1a) are indeterminate in that one cannot determine the state-of-affairs that makes the sentence true (Gillon 2004). “Indeterminate” and “underspecified” are sometimes used interchangeably to refer to a state of uncertainty regarding the truth-value or the semantic representation of the sentence (see Gillon 2004). We reserve “underspecified” for the more general phenomenon of semantic uncertainty. Thus, (1a) is indeterminate but not necessarily underspecified: if the context provides information about a particular event—that is, an event that specifies what Mary began doing with the book—then semantic uncertainty is resolved.

We thank Brendan Gillon, Anna Maria Di Sciullo, and two anonymous reviewers for comments. We are indebted to Lisa Travis for extensive discussions on the syntactic account of type-shifting predicates presented in section 3. Since the time that this paper was originally submitted in February 2005, both authors have learned of related studies and theoretical accounts, including Pylkkänen and McElree (2006), which did not justify altering substantially the nature of our arguments or our view on type-shifting operations and compositionality. Research for this paper was supported by a grant from the Social Sciences and Humanities Research Council of Canada.
An indeterminate sentence such as Mary began the book (1a) is not equivalent in meaning to the counterpart sentence Mary began reading the book (1b), nor is the interpretation in (1b) the only one licensed by (1a). A key question is how the meaning of indeterminate sentences is represented—via Fregean compositionality (Fodor and Lepore 2002; Partee 1984) or via operations that go beyond compositionality, such as “cocomposition” (Pustejovsky 1995) or “enriched composition” (Jackendoff 1997). Recently, there has been a surge of psycholinguistic work on the nature of the interpretation of the so-called type-shifting constructions (McElree et al. 2001; Traxler et al. 2002; de Almeida 2004; Pickering et al. 2005), in large part motivated by the lexical-semantic work of Pustejovsky (1995) as well as by the earlier work in the Montague semantics framework (Partee and Rooth 1983; Partee 1987).

The article is organized as follows. We review theoretical accounts of type-shifting (section 1) and discuss psycholinguistic experiments designed to address the nature of type-shifting operations (section 2). We discuss two levels of representation and processing involved in these operations: we consider what may happen at a linguistic level of analysis and what may happen at higher levels, perhaps conceptual and pragmatic. We make two claims: (i) indeterminacy is generated by linguistic structure, which in turn reflects the nature of the predicates involved in type-shifting (section 3); and (ii) the basis for the interpretation of indeterminate linguistic structures is inferential or pragmatic in nature (section 4). In particular, pace Pustejovsky (1995) and Jackendoff (1997), we argue for the atomic representation of lexical concepts.

1. Type-Shifting and Compositionality

Type-shifting effects may be a consequence of computations effected at three distinct levels of representations. First, type-shifting may reflect syntactic computations, where empty elements in the argument structure of verbs such as begin or enjoy may lead to type-shifting effects. Second, type-shifting may reflect computations at the syntax-semantic interface; these are true cases of type-shifting but with no computational cost. Third, type-shifting may be a consequence of computations at the conceptual-pragmatic level, for example due to indeterminacy in token expressions. These three alternatives may overlap, as when gaps in verbal argument structure trigger pragmatic inferences. Before exploring these three distinct types of type-shifting operations in detail, we briefly survey different approaches to type-shifting.

1.1. Nominals and type-shifting

Let us assume that nominals (hereafter NPs) can belong to (at least) three different semantic types: referential \(<e>\) (e.g., Mary); predicative \(<e,t>\) (e.g., a fool); or quantificational \(<<e,t>,t>\) (e.g., every woman) (Partee and Rooth 1983; Partee...
Different type-shifting operations yield the appropriate NP type in different contexts. These semantic types can apply to the same NP token in different contexts. For example, the proper name Mary can refer to the individual entity (type $<e>$) or to the set of properties that are true of Mary (type $<<e,t>,t>$). Whereas the context in (2a) requires type $<e>$ for Mary, the context in (2b) requires the quantificational type for Mary and every man in the room. In (2b) the proper name Mary is conjoined with a quantificational NP (every man in the room): if constituent conjunction requires an identity of types, then, in this context, Mary must be a quantificational NP.

1. Mary sat down.

2. Mary and every man in the room sat down.

Partee (1987) proposes that an expression such as Mary is assigned a different semantic type via the semantic operation of type-lifting, a subtype of type-shifting. In particular, an expression of referential type $<e>$ can be lifted to a more complex type such as $<<e,t>,t>$. Of interest to the present discussion is the fact that Partee’s analysis foreshadows some of the findings of our research:

Even the most general type-shifting principles, such as the ‘lifting’ operation that maps $j$ (type $e$) onto $\lambda P [P(j)]$ (type $<<e,t>,t>$), need not be universal, but I would expect such a principle to be universally available at ‘low cost’ or ‘no cost’ for any language that has NPs of type $<<e,t>,t>$ at all. (Partee 1987:120)

Partee is primarily concerned with the universality of type-shifting principles, but we also understand her to be suggesting that type-shifting a nominal from a referential to a quantificational type should yield no computational cost. This is consistent with Partee’s view that types such as $<<e,t>,t>$ and $<e,t>$ are acquired by any basic type $e$. If so, then the meaning of a nominal — and the determination of its proper semantic type — is a function of its sense and sentence context. If computational cost involves mental/computational resources, then we expect that it will correlate with behavioural measures such as extra processing time. Moreover, if type-lifting is a low-cost (or no-cost) operation, then it should not impose any demands on parsing — that is, linguistic — resources.

1. In the Partee/Montague notation, $e$ designates terms (entities), while $t$ designates formulas (sentences with a truth value).

2. Other instances of low-cost type-shifting include mass/count shifts (i), metonymy (ii), and genitives (iii). See Gillon (1992) and Partee and Borschev (1999).

(i) I want three beers (beer $<\text{mass}> \rightarrow \text{beer }<\text{count}>$)

(ii) Many people protested against Vietnam (Vietnam $<\text{entity}> \rightarrow \text{Vietnam }<\text{event}>$)

(iii) Mary’s movie
    (her favourite/the one she acted in/the one she often rents, etc.)
Furthermore, Partee’s type-shifting operation maps syntactic structure onto semantic/conceptual structure by establishing a correspondence between univocal syntactic categories and different semantic types. If semantic interpretation is fed by the syntactic component (Chierchia and McConnell-Ginet 1990), then this kind of type-shifting occurs at the level of syntax or at the syntax-semantics interface.

The type-shifting operations that by hypothesis apply to cases such as (2b) do not resort to semantic composition beyond Fregean compositionality. Rather, type-shifting is a mapping function that establishes the mode of interpretation for a particular nominal given its position within syntactic structure: for example, by virtue of type-lifting, the two NPs in (2b) are accorded the same status in LF (logical form) and beyond. We take LF to be a level of linguistic description represented in a disambiguated language, which allows for the denotation of a sentence; as such, it is not a full specification of intended meaning. Moreover, type-shifting operations do not rely on internal analyses of lexical items in order to derive the proper semantic type for interpretation. Semantic interpretation applies to atomic lexical-semantic representations, and relies on sentence context (to a first approximation, LF) to yield the proper semantic type.

We now turn to another kind of type-shifting operation, one that takes into account computations within the domain of verbal argument structure.

1.2. Type-shifting within VPs

Some type-shifting operations appear to not be computationally low-cost. One such case is the type-shifting operation that occurs within VPs — and more specifically the shifting of semantic types for nominals — as a function of the interaction between the demands of verbal argument structure and the kinds of complements that verbs take in particular sentences. For example, in (3) it is not clear what type of interpretation should be assigned to the VPs begin the book and enjoy the book.

(3) a. Mary began the book.
   b. John enjoyed the book.

Here and throughout, we do not distinguish between sentences (Mary began the book) and their meanings/senses. Mary began the book means, roughly, that MARY BEGAN THE BOOK, which is constituted by the meanings/denotations of the lexical items (Mary, begin, book) and their function in expressing the proposition that Mary began the book.

If the main verbs begin and enjoy in (3) simply require an event verb within the VP, then the interpretations in (4) and (5) should be equally felicitous, yet they are not. Specifically, one more easily understands Mary began the book (3a) as describing a situation that corresponds to Mary began reading the book (4a), than as Mary began eating the book (5a). Similarly, one more easily understands John enjoyed the book (3b) as describing a situation that corresponds to John enjoyed reading the book (4b), than as John enjoyed eating the book (5b).
(4)  a. Mary began reading the book.
    b. John enjoyed reading the book.

(5)  a. Mary began eating the book.
    b. John enjoyed eating the book.

1.3. **Is enriched composition necessary?**

Pustejovsky (1995) proposes that the interpretation of sentences such as (3)—with the selection of an appropriate event such as (4)—must take into account the analysis of the NP complement *the book*. This is based on the idea that nominals are represented as *qualia* structures, where information about the noun referent’s “purpose and function” (its “telic role”), the factors “involved in its origin” (its “agentive role”), are part of the lexical entry of the nominal—that is, are represented in the lexicon. In Pustejovsky’s theory, nominals belong to a family of semantic types including entity (*the book*) and event (*the fight*). In (3), the lack of an event complement for the verbs *begin* and *enjoy* leads to an operation of coercion that type-shifts *book* from entity into event. This type-shifting (type-coercion) takes into account the lexical representation for *book*, which has among its telic roles the reading event that interpolates the semantic structure of the sentence. Consequently, *Mary began the book* (3a) is interpreted as *Mary began reading the book* (4a).

Pustejovsky’s coercion applies to the complement NP so that its compositional structure is enriched by telic role information. Crucially, interpolation is not an operation that fills in a gap in the argument structure of a verb. Pustejovsky (1995) rejects the notion that verbs such as *begin* and *enjoy* have multiple argument structures or even an argument structure that allows for an optional argument. He treats such verbs as polysemous, with different argument structures referring to different senses. This approach is common to his treatment of lexical representation. For Pustejovsky, cases such as *Mary broke the window* and *Mary passed through the window* reflect two senses of *window*: they are instances of logical ambiguity, where one meaning refers to the physical object denotation and the other refers to the aperture denotation. One could then argue that *Mary hit the window, Mary targeted the window, Mary admired the window* would also be treated as different “denotations” of *window*: the “obstacle” versus “target” versus “object of art” denotation. It is unclear to us why these so-called denotations are not treated as characteristics or “features” or part of the content of a univocal *window*. What is clear is that *window* seems to have a constant (referential) meaning across all these contexts.

For Jackendoff (1977), the interpretation of *Mary began the book* (3a) is a function that assigns the activity interpretation to the complement NP by virtue of its missing event argument, as in (6).

(6) Interpret NP as $[\text{Activity} F(\text{NP})]$
The operation in (6) allows the activity/event requirement of a verb such as *begin* to be interpreted as “an unspecified activity involving NP” (Jackendoff 1997:61).

At a second stage of interpretation, the unspecified activity needs to be determined. Although Jackendoff does not endorse Pustejovsky’s *qualia* theory, he concedes that enriched composition is involved in the interpretation of sentences such as *Mary began the book* (3a), such that the nominal complement determines the reading of such sentences.

We agree with Jackendoff that a function may apply to the interpretation of NP to license its activity meaning in the absence of an explicit activity in the context of the sentence. And we also agree that the interpretation operation in (6) “leaves open what the activity in question is” (Jackendoff 1997:61). But while Jackendoff determines the nature of the activity from the content of the nominal, we propose that the unspecified activity is determined only by the wider context.

Moreover, we take the operation interpreting the VP to be a gap-filling operation on argument structure whereby the specification of the activity is inferentially determined on the basis of discourse information.

### 1.4. Simple composition enforced

Fodor and Lepore (2002) offer a different view of how indeterminate sentences such as *Mary began the book* (3a) and *John enjoyed the book* (3b) are interpreted. Their discussion of expressions such as *x want y* gives us clues about how the interpretation of indeterminate constructions in general might occur. Consider the examples in (7).

\[
\begin{align*}
7(a) & \quad \text{Mary wants a beer.} \\
7(b) & \quad \text{Mary wants to drink a beer.} \\
7(c) & \quad \text{Mary wants to have a beer.}
\end{align*}
\]

In Pustejovsky’s theory, (7a) is interpreted as (7b): the analysis of the nominal complement gives rise to its telic role, which enriches the compositional process with interpolated information. Fodor and Lepore, however, propose that (7a) should be interpreted as (7c). In their account, the interpretation of *x wants y* as *x wants to have y* involves two stages. First, lexical items (*x, want, y*) are assigned their denotations (are interpreted). At this stage, the argument structure of the verb is determined, that is, the arguments *x* and *y* and their syntactic/semantic functions are determined, such that it is *x* that wants *y* and not the other way around. Second, type-shifting applies, such that *x wants y* is interpreted as *x wants to have y*; this is achieved by assigning to *y* (the object of *want*) a function that establishes *y* (NP) as the object that *x* wants, as in (8).

\[
\{x: x \text{ wants to have F(NP)}\}
\]

Fodor and Lepore (2002) acknowledge that, for this solution to hold, nominals must be attributed semantic types: if *x wants y*, *x* is taken to be a “creature” and *y* a “state of affairs”, with such types being “entirely heuristic” (Fodor and
Lepore 2002:113, n. 24). Type-shifting takes into account properties of the verb in selecting the proper type for its arguments; for example, in selecting for its NP complement a state-of-affairs reading rather than a creature or entity reading. In the Fodor and Lepore analysis, the putative complexity of lexical representations is subsumed by two hypotheses:

(i) lexical meaning is roughly denotation (that is, what the item means is determined by its *nomic*, referential properties (Fodor 1990)); and

(ii) a lexical item’s compositional properties are determined by compositional rule, that is, how the item contributes to its place or function in a logical form of which it is a constituent.

We take (i) to be atomic rather than definitional or complex. We understand (ii) to be the argument structure of a lexical category, where arguments are positions licensed by the verb — regardless of whether or not they receive theta-roles (Grimshaw 1990) — and where argument structure specifies (and orders) the categories that form grammatical sentences with the verb.

Conceptually, the distinction between the Pustejovsky (1995) and Jackendoff (1997) proposals on the one hand, and the Fodor and Lapore (2002) proposal on the other hand is that the latter does not resort to supplementary internal analyses of the object NP to interpret the structure that holds between the argument and the verb. In this respect, their proposal is similar to Partee’s in assuming that arguments of a verb are assigned a semantic type that establishes conditions for their interpretation. Accordingly, we regard Fodor and Lepore’s position as retaining the compositionality of Partee’s account of type-shifting: arguments of a verb are assigned particular types depending on the role they play in a given structure.

1.5. Where do telic roles come from?

We briefly address the nature of Pustejovsky’s (1995) telic roles that allegedly contribute content to the resulting interpolated constructions. Given that telic roles are the source of information in enriched composition, it behooves us to discuss their status as regards linguistic and psycholinguistic theorizing about type-shifting and semantic composition.

If any NP referent has a particular function, it is designated by the proper context. Given that objects can appear in \( n \) contexts (with \( n \) being a very large number), no specific role can be encoded in the lexicon. Pustejovsky (1995) proposes that an expression such as *window* has different senses in the different contexts given in (9).

\[
\begin{align*}
(9) & \quad \text{a. Mary passed through the window.} \\
& \quad \text{b. Mary closed the window.}
\end{align*}
\]

As discussed above, all senses of *window* can be treated univocally as contributing the same content to the different expressions — namely, the epistemic (world-knowledge) properties that trigger the concept *window* under different conditions
and across different contexts. If the lexicon does not encode world-knowledge (Jackendoff 1997:61–62; Asher and Pustejovsky 2000), then lexical information is about aspects of word meaning that are “linguistically active” (Grimshaw 2005). Thus, the distinction between melt and freeze is linguistically inactive as the two verbs behave similarly across linguistic contexts. Similarly, the occurrence of window in (9a) versus (9b) is linguistically inactive.

There remains the persistence of telic-role effects. Why can Mary began the book (3a) be interpreted as Mary began reading the book (4a)? It is plausible that if Mary began the book, then what Mary began doing with the book was something like reading it. If the notion of a telic role has any status whatsoever, we conjecture that it is only because of typicality effects. It is well known that lexical-conceptual representations show pervasive prototypicality effects (Murphy 2002): for example, when people think PET they think DOG; when they think BIRD they think more frequently ROBIN and less so PENGUIN. Similarly, when people think BOOK they think READ more often than they think EAT, presumably because books are often read and rarely eaten. Such effects point to lexical/conceptual associations created by frequency of use, or by what we know to be true about the world. We hold that such categorization effects say nothing about how lexical concepts are represented and moreover, that they are opaque to linguistic structure.

2. Psycholinguistic Empirical Support and Lack Thereof

Some psycholinguistic studies are taken as providing support for type-shifting effects of the kind advocated by Pustejovsky (1995) and Jackendoff (1997). We show that these research results are questionable both in terms of the strength of the effects and the nature of the experimental materials employed.

Psycholinguistics is full of paradigm disputes — à la Kuhn (1962) — ranging from the debate on the modularity of lexical versus sentence processing to the nature of lexical-semantic versus conceptual representation. Often, debates turn into statistical or methodological qui-pro-quo when the real crux of the matter is the theoretical choices researchers make when they construct their objects of inquiry. We evaluate a recent dispute on the nature of type-shifting (coercion) effects, taking into account a series of experiments aimed at investigating where these effects — if they exist at all — come from. The dispute is (or should be) mainly about three main issues:

(i) The type of theoretical position one takes on the nature of lexical-conceptual representation and compositionality;

(ii) The nature of verbs that engender type-shifting effects; and

(iii) How solid the experimental evidence is.
2.1. Overview of key studies

In a quasi-replication self-paced reading study with sentences such as *Mary began the book* in isolation, semantic coercion effects obtained by McElree et al. (2001) were not obtained by de Almeida (2004). When the same sentences were presented in context — roughly, paragraphs setting the stage of the activity performed by the agent NP of the main sentence — type-shifting effects could not be distinguished from those obtained with low-probability verbs. Pickering et al. (2005) raise the possibility that the failure of de Almeida (2004) to replicate McElree et al.’s (2001) previous finding could be a type-II error: failure to obtain an effect when it is real. They also obtained results compatible with one of de Almeida’s experiments and reported effects that, as we argue below, are moot.

2.2. The alleged coercion effects

We review some of the claimed effects taken to support Pustejovsky’s (1995) view of coercion and reassess the nature of the verbs used in the psycholinguistic studies. We focus on what we take to be the key issue: the nature of the coercion process and how lexical concepts are mentally represented. We argue that even if coercion effects are obtained (*contra* de Almeida’s study, and *pro* McElree et al.’s results), the crux of the matter lies in the explanation provided for the effect. We further argue that the evidence favouring a lexical-semantic decompositional theory is unconvincing.

The original psycholinguistic study on type-shifting was conducted by McElree et al. (2001). Motivated by Pustejovsky’s (1995) approach, their goal was to show that coercing a nominal such as *the memo* when it was preceded by a verb such as *begin*, as in (10a), was associated with greater processing complexity.

\[(10)\]

\[\begin{align*}
\text{a. Coerced/type-shifted:} \\
& \text{The secretary began the memo before the annual sales conference.} \\
\text{b. Preferred:} \\
& \text{The secretary typed the memo before the annual sales conference.} \\
\text{c. Non-preferred:} \\
& \text{The secretary read the memo before the annual sales conference.}
\end{align*}\]

McElree et al. argue that (10a) should result in an increased processing load compared to (10b) (the preferred construction, according to their normative study) and (10c) (the non-preferred construction) (McElree et al. 2001; de Almeida 2004). They reason is that if type-shifting (as defined by Pustejovsky) coerces the NP into the appropriate semantic type, then longer reading times at the nominal or post-nominal positions are expected; we refer to this as a type-shifting effect. In fact, they obtain such an effect, and also find that, in several positions after the verb, (10a) takes longer to process than (10b) or (10c). According to McElree et al., the main reason for this effect is that *memo* type-shifts from an entity to an...
event, and coercion (leading to enriched composition) takes processing resources (thus, longer time).

These findings appear to be consistent with the notion that type-shifting constructions are more complex. In the contrast between (10a) and (10c), it seems that type-shifting constructions exhibit a complexity effect that is above and beyond what one might expect from a non-preferred verb. This is because there is a processing load associated with the non-preferred verb, even though it is congruent from a structural and semantic viewpoint. The cost associated with this verb is that its direct object is not expected, relative to the preferred verb. This surprise effect resolves itself immediately. However, the McElree et al. data show that the type-shifting verb construction exhibits a complexity effect that goes beyond the direct object position. This is a long-lasting effect, suggesting that the longer reading times are not simply due to the presence of an inappropriate word for the sentence.

Traxler et al. (2002) use the same materials as McElree et al. (2001), but employ an eye-tracking technique. They report five measures of eye movement data: first-pass time, first-pass regressions, second-pass time, total time, and regression-path time. Reading times for three regions of the sentences are monitored: the main verb region (began/typed/read); the target region, which includes the determiner and noun (the memo); and the post-target region (the first two words following the direct object, e.g., before the). This is a sensitive technique, yielding 15 different measures of eye-movement behaviour while subjects read sentences online. Of all the analyses, few yield statistically significant effects (assuming the canonical $\alpha < 0.05$ in both subjects and items analyses), with only trends in other analyses taken to support their view. More importantly, none of the first-pass time analyses show significant type-shifting effects.\(^4\)

---

\(^3\)The reading-time measures employed in these studies are, as defined by Traxler et al. (2002:535):

First-pass time is the sum of all the fixation times beginning with the reader’s first fixation in a region until the reader’s gaze leaves the region (. . . ) A first-pass regression occurs when the reader’s gaze crosses the left edge of the scoring region following a first-pass fixation. Second-pass time includes all of the time spent in a region following the first-pass fixations, including time spent in the region after exiting to the left and time spent in the region after exiting to the right. Total time is the sum of all of the fixations within a region. Regression-path time for a region includes all fixation times from the first fixation in a region until the reader fixates to the right of the region.

\(^4\)We refer here to Experiment 1 in the Traxler et al. (2002) study, which is relevant to the present discussion. Their results showed that second-pass and total time reading times associated with the verb region were significantly longer in (11a) compared to (11b) and (11c). At the target region (the memo), where effects were found by McElree et al. (but not replicated by de Almeida 2004; see below), only first-pass regressions were significant, in the same direction as those effects obtained in the verb region, but only in pair-wise
The interpolation effect associated with type-shifting verbs does not show up immediately in the target region and in first-pass reading times. The only significant effect occurs for second-pass reading times, suggesting that the alleged coercion effects, if they exist, are relatively late-occurring, and not necessarily produced by online lexical generation of interpolated structure, as claimed. What is important is that, despite the authors’ interpretation, these eye movement data do not strongly support the findings of McElree et al. (2001).

It therefore comes as no surprise that de Almeida (2004) reports that, when no context is provided before the sentence, there is no significant reading time difference between type-shifting compared to preferred and non-preferred constructions in the first of two self-paced reading experiments. This is important because de Almeida (2004) uses similar materials to those of McElree et al. (2001), modifying them for contextual congruity, as in (11). Crucially, verb triads remain the same.

(11)  
a. **Coerced/type-shifted:**
   The secretary began the memo long before it was due.

b. **Preferred:**
   The secretary typed the memo long before it was due.

c. **Non-preferred:**
   The secretary read the memo long before it was due.

When the same sentences are preceded by a context providing more information about the nature of the event that *began* could refer to (e.g., to work on the memo), as in (12), preferred constructions (11b) are significantly faster than both coerced (11a) and non-preferred (11c) constructions, while these last two pair together.

(12) **Preceding context:**
   The secretary would always be sure to work ahead of schedule. She was asked to work on the memo.

When contextual information specifies the nature of the event to which the sentence refers, de Almeida (2004) argues that type-shifted constructions behave just like non-preferred ones. He considers constructions with verbs such as *begin* as underspecified when they lack an event complement, with this underspecification being resolved pragmatically, assuming sentences with *begin NP* violate Gricean-type maxims when out of context (a suggestion also made by Fodor and Lepore 2002).

In a more recent study, Pickering et al. (2005) contrast the materials used by de Almeida (2004) with new conditions such as those in (13).
(13)  a. Coerced/type-shifted:
    The carpenter began the table during the morning break.

    b. Preferred:
    The carpenter built the table during the morning break.

    c. Non-preferred:
    The carpenter sanded the table during the morning break.

    d. Full-VP preferred:
    The carpenter began building the table during the morning break.

Pickering et al. (2005) find a similar pattern of results as de Almeida (2004): no statistically significant effects in first-pass reading, with type-shifting effects appearing only in regression data. To our surprise, most of the results taken to support a type-shifting effect are presented by Pickering et al. as “strong tendencies”, with many probability values larger than 0.05. Moreover, the fact that regressions are significant does not necessarily indicate coercion effects, as the latter are also consistent with the detection of parsing gaps or anomalies.

In sum, the empirical results favouring type-shifting effects are weak at best. In addition, the main questions are still unresolved, namely:

(i) What process does type-shifting entail?

(ii) What is the nature and source of the information that aids in the interpretation of coerced/type-shifted constructions?

We take up these questions in the next two sections. We argue that under the umbrella of type-shifting — in particular, the type-shifting effects studied in the experiments discussed above — the syntactic processing and semantic interpretation of indeterminate constructions can be analyzed in a number of different ways. We propose that the inherent structural properties of the verbs are themselves potential sources of type-shifting effects. In addition, we suggest that the source of information used in interpreting indeterminate sentences such as $x$ began $y$ and $x$ enjoyed $y$ is pragmatic in nature.

3. A Structural Account of Type-Shifting Constructions

Empirical support for type-shifting effects within VPs is moot. Even if there were effects at the locus where one expects to find processing differences, one must also take into consideration the materials employed in those psycholinguistic studies. We propose an alternative linguistic account of the alleged type-shifting effects: we claim that (14a) is not structurally analogous to (14b). Rather, (14a) contains an extra VP, with an empty verbal head, as in (14c).

---

5Thanks to an anonymous reviewer for challenging us to clarify our claims.
(14)  a. The secretary began the memo.
b. The secretary typed the memo.
c. The secretary began $[\text{VP} \ [\text{V} \ \text{e}] \ [\text{NP} \ \text{the memo}]]$

We claim that the extra VP structure in (14c) accounts for the extra time taken in some experiments to process such sentences. Before arguing for this, we would like to point out that there are two a priori reasons to search for an account different from McElree et al.’s (2001). First, if one adopts Partee’s view of semantic type-shifting operations as low-cost or no-cost, then from a psycholinguistic perspective it is unclear why a type-shifting operation would be costly for the processor in the first place. By claiming that extra structure accounts for extra processing time, we rely on old yet standard assumptions about sentence processing (Frazier and Fodor 1978): structural computations correlate with processing time. Second, the structure in (14c) maps onto its meaning in a transparent and compositional fashion, allowing us to maintain the principle of compositionality. We see this last point as an advantage of our analysis over competing ones.

3.1. Restructuring verbs

An examination of the verbs used in diverse studies, but in particular those of McElree et al. (2001) and Traxler et al. (2002), reveals that only about 14 verbs have been examined. This is because the stimuli used rely on repetition both within and across experiments. For example, in the three experiments reported in Traxler et al. (2002), only 12 verbs are studied, and just over half of these are aspectual verbs ($\text{begin, start, end, finish, endure, complete}$), with the other five belonging to a different class ($\text{try, prefer, resist, enjoy, expect}$). What is special about these predicates? On independent grounds, Wurmbrand (2001, 2004) argues that German has a class of verbs, which includes $\text{begin}$ and $\text{try}$, that allows for lexical restructuring of infinitives. We show that extending Wurmbrand’s restructuring analysis to English allows us to account for the alleged type-shifting effects. Before providing arguments for this, we examine the syntactic nature of restructuring verbs.

---

6In de Almeida’s (2004) Experiment 1, no differences were found between type-shifting ($\text{begin}$), preferred ($\text{type}$), and non-preferred ($\text{read}$) verbs when sentences were presented in isolation. This may be an indication that if structural differences — as we argue herein — do account for differences in processing times, these differences may be slim. In de Almeida’s Experiment 2, when sentences are presented with context, the differences between preferred and type-shifting/non-preferred appear. These differences may be due to the “deeper processing” (Kemtes and Kemper 1999) that is facilitated by extra context.

7See Newmeyer (1969) and Rochette (1999), where it is argued that $\text{begin}$ is an aspectual verb with a distinct syntax; and Freed (1979), who elaborates on the class of aspectual verbs.

8We thank Lisa Travis for extensive discussion of these issues. All errors are ours.
Restructuring involves constructions where two verbal domains act as if they are clause mates (Evers 1975; Aissen and Perlmutter 1976; Rizzi 1978): certain processes, which normally apply within one clause, apply across the two verbal domains. For example, in Italian and Spanish, an infinitival clause may undergo clausal union with the matrix clause, such that the two clauses are no longer independent (Aissen and Perlmutter 1976). Relevant to our concerns is the fact that aspectual verbs are typically restructuring verbs (Wurmbrand 2001). In addition, Wurmbrand identifies a second, less coherent class—subject to speaker and language variation—which includes verbs such as try, manage, recommend, permit. Of interest to us is the fact that the class of predicates that Wurmbrand identifies as restructuring verbs overlaps with the class of predicates used in the experiments discussed above.

Wurmbrand argues that restructuring predicates combine with “syntactically and semantically very ‘small’ predicates” (2001:17). In her analysis, Wurmbrand adopts a mono-clausal structure for restructuring contexts: these restructured infinitives not only lack a tense and complementizer projection, but also lack a structural object case position (for her this is AgrOP/vP). The structure of the restructuring predicate in (15a) is illustrated in (15b).

(15) German:

a. weil Hans den Traktor zu reparieren versuchte
   since John the tractor to repair tried
   ‘since John tried to repair the tractor’ (Wurmbrand 2001:17, ex. 5)

b. 

\[
\begin{array}{c}
\text{TP} \\
\downarrow \\
\text{John} \\
\downarrow \\
\text{vP} \\
\downarrow \\
\text{t}_{\text{SUBJ}} \\
\downarrow \\
\text{v'} \\
\downarrow \\
\text{VP} \\
\downarrow \\
\text{V}^{0} \\
\downarrow \\
\text{OBJ} \\
\end{array}
\]

\text{the tractor to repair}

\[V_{^0}\]

\[\text{tried}\]

\[\text{Wurmbrand 2001:17, ex. 5}\]

---

9 Other restructuring verbs include modal verbs, motion verbs, and lexical causatives.

10 vP (Chomsky 1995) is a functional projection of VP: as well as assigning a theta-role to the external argument/subject, v also assigns structural accusative case to the internal argument/direct object. Conceptually, vP is akin to Larson’s (1988) VP-shell.

The following abbreviations are used:

- ACC: accusative
- OBJ: object
- NOM: nominative
- SUBJ: subject
Extending Wurmbrand’s proposal, we claim that English has a class of restructuring verbs that consist of aspectual verbs plus another class, which we call the try-class. Accordingly, the structure of (16a) is as in (16b).

(16)  a. The secretary began to type the memo.
    b. 
      \[
      \begin{array}{c}
      \text{TP} \\
      \text{the secretary} \\
      T' \\
      T_0 \\
      \text{vP} \\
      t_{\text{SUBJ}} \\
      v' \\
      v_0 \\
      \text{VP} \\
      V_0 \\
      \text{began} \\
      V_0 \\
      \text{VP} \\
      \text{Obj} \\
      \text{to type} \\
      \text{the memo}
      \end{array}
      \]

The restructuring analysis also applies to begin the memo, where the structure of (17a) is as in (17b), with English allowing an empty V head.

(17)  a. The secretary began the memo.
    b. 
      \[
      \begin{array}{c}
      \text{TP} \\
      \text{the secretary} \\
      T' \\
      T_0 \\
      \text{vP} \\
      t_{\text{SUBJ}} \\
      v' \\
      v_0 \\
      \text{VP} \\
      V_0 \\
      \text{began} \\
      V_0 \\
      \text{VP} \\
      e \\
      \text{the memo}
      \end{array}
      \]

We conjecture that (17) is more complex than (18); the latter corresponds to the “preferred” types of verbs used in the experiments described above. The crucial difference between (17) and (18) is that the former has an empty V head, and in the latter the V head is spelled out.
The secretary typed the memo.

In the present analysis, when restructuring verbs from the aspectual class (e.g., begin) or the try-class are employed, the structure in (17) is required. If so, then it is possible that a sentence with an empty V head (as in (17)) may take longer to process than a sentence without it (as in (18)). This is compatible with a compositional view of semantic form, for the unsaturated empty V head necessarily draws on extralinguistic resources for its interpretation.

3.2. Empty V heads

Consider again the following examples:

The secretary began [VP [V₀ to type] [NP the memo]]

The secretary began [VP [V₀ e] [NP the memo]]

In (19a), the memo is an argument of type. In (19b), we claim that the empty V can assign a theta role to the memo. If so, then the need to coerce the NP in order to obtain a type-shifted event reading for the memo is dispensed with (pace Pustejovsky 1995 and Jackendoff 1997), allowing the retention of classic compositionality.

Empty V is not specific to English: it occurs in other Germanic languages, in contexts that arguably involve restructuring, with the lower predicate empty. This is shown in (20) for German with modal verbs, and in (21) for Swedish (Vinka 1999) with predicative verb-particle constructions.

(20) German:

a. Ich kann Deutsch.
   I-NOM can German
   ‘I can (speak) German.’

b. Ich kann Deutsch sprechen.
   I-NOM can German speak
   ‘I can speak German.’
c. Ich muss heute ins Büro.
   I-NOM must today in.the office
   ‘I must (go) to the office today.’

d. Ich muss heute in das Büro gehen.
   I-NOM must today in/to the office go
   ‘I must go to the office today.’

(21) Swedish:
   Kalle hade på TVn.
   Kalle had on TV
   ‘Kalle switched on the TV.’

In support of our view that type-shifting is simply an effect of linguistic structure due to differences in VPs, we present four distributional arguments:

(i) VP modifiers can scope over an empty V within a VP.
(ii) VPs (with overt or empty Vs) raise to subject position of V-able structures.
(iii) VPs (with overt or empty Vs) don’t raise to subject position in unaccusative structures.
(iv) The object of the lower empty V raises in passive structures.

Arguments (iii) and (iv) for the extra VP structure reconsider points raised by Pylkkänen and McElree (2006), who ultimately decide that a type-shifting analysis — contra a VP with a structural gap — is warranted. We come to the opposite conclusion.

3.2.1. VP-modifiers scope over the empty V

If a VP-modifier can scope over the empty V, then that is evidence for its presence. This is confirmed: even when the lower predicate is silent, VP adverbs such as again and instrumental PPs can modify the lower VP (von Stechow 1996). This is illustrated in (22) with the adverbial modifier again.

(22) a. We watched Fargo last week. I’d prefer Gandhi again.

   b. I’d prefer [VP [VP [V\(^0\) e] [NP Gandhi]]] again]

In (22a), the scope of the adverbial modifier again is not on the higher verb prefer but on the lower empty verb, which is anaphoric to watch. That is, the speaker is not talking about what s/he prefers to do again; rather, the repeated event concerns watching the movie Gandhi, as in (22b). Thus, even when the lower verb is not expressed overtly, it is nevertheless available for modification.\(^{12}\)

\(^{12}\)Pylkkänen and McElree (2006), which came to our attention after we prepared the original version of the present article, consider a similar argument and conclude otherwise. If the empty verbal head is made salient with some discourse context (even the previous sentence), the modifier scope discussed in the main text does indeed hold.
With instrumental phrases, a sentence such as (23a) can be true on a reading where at some early point in the eating event the speaker used a fork. That is, a true reading is possible even if at the very moment that the eating event began, a chopstick was not used. This reading indicates that the instrumental can indeed modify the empty predicate, which is anaphoric on eating, as in (23b).

(23)  a. I started eating with chopsticks but finished with a fork.
    b. . . . but finished \[VP [V_0 e] \] with a fork

3.2.2. VPs raise to subject position in V-able structures

Another distributional test involves event-selecting verbs embedded inside adjectives, as in (24a). For example, when finish, an event-selecting verb, contains the adjectival suffix -able and is combined with an NP in subject position, grammaticality ensues. We posit the structure in (24b).

(24) a. This book is finishable.
    b. \[VP [V_0 e] [NP This book] \] is finishable.

Our analysis of (24) predicts that the subject position of V-able adjectives will support the appearance of overt VP-arguments. This prediction is confirmed:

(25) a. \[VP Reading this book] is finishable.
    b. \[VP Climbing this wall] may be survivable.

(Pylkkänen and McElree 2007:558, ex. 47–48)

Thus, from a distributional standpoint, the extra VP hypothesis is supported.

3.2.3. VPs don’t raise to subject position in unaccusative structures

An event-denoting NP complement in (26) can alternate between a causative and an unaccusative structure: if an agent does not already fill the subject position, an event NP may fill that position.

(26) a. We began \[NP the war\].
    b. \[NP The war\] began.

With this in mind, consider (27) and (28), which show that VPs may not raise to the subject position in an unaccusative structure. In (27), raising the event-denoting gerundive VP to subject position is blocked. In (28), raising the event-denoting infinitival VP to subject position is likewise blocked.

(27) a. We began \[VP reading the book\].
    b. *[VP Reading the book] began.

(28) a. We began \[VP to read the book\].
    b. *[VP To read the book] began.
Now consider (29): while (29a) asserts that a book-reading event was begun by Sarah, (29b) does not mean that a reading of a book began.

(29)  
   a. Sarah began the book.  

The impossibility of the book occurring in subject position in (29b) follows from our analysis. Recall that the book is generated as the complement of an empty V, as in (30).

(30)  
   \[ VP \left[ V^0 \, e \right] \left[ NP \, \text{the book} \right] \]

Just as event-denoting full VPs — be they gerundive (27b) or infinitival (28b) — cannot raise to subject position in an unaccusative structure, so too is it impossible for an event-denoting VP with an empty V to raise (29b).

3.2.4. **Object of lower empty V raises in passive structures**

The fourth and final argument in favour of the VP-structure involves passivization. Pylkkänen and McElree (2006) observe that when the VP is overt, raising the object of the lower VP to the subject position is odd:

(31)  
   a. The author began to write the book.  
   b. ?[NP The book] was[VP begun [VP to be written tNP by the author]].

They argue that if the type-shifted sentence were then structurally identical to the overt version, as argued for here, then it should also be deviant.\(^{13}\) However, this is not the case:

(32)  
   a. The author began the book.  
   b. The book was begun by the author.

The contrast between (31) and (32) would seem to be an argument against our restructuring analysis, which crucially invokes the presence of a lower VP structure with an empty V. However, the restructuring analysis not only accommodates this data point, but actually predicts it. Wurmbrand proposes that restructuring involves a form of Exceptional Case Marking such that case assignment to the (embedded) object occurs in the matrix vP. Consequently, Case assignment for (16) is as in (33).\(^{14}\)

---

\(^{13}\)To our ears, the sentence in (31b) is well formed.

\(^{14}\)Assuming that Case checking occurs via movement such that the embedded object raises to Spec-vP.
The embedded object is reliant upon the matrix vP for Case assignment. However, if the matrix verb begin is passivized, then according to Wurmbrand, Spec-vP is unavailable, and only Spec-TP (the subject position) is available as a Case assigner for the embedded object, as in (34).

(34) Passive restructuring contexts:

If begin and try are passivized, then we expect that the direct object can raise to the subject position of these English restructuring verbs. This elegantly accounts for the data in (32), and further supports the extension of Wurmbrand’s (2001) restructuring analysis to so-called type-shifting verbs in English.

We have called into question the theoretical basis of coercion, as put forth by researchers working within the lexical-semantic tradition of semantic type-shifting. We have also shown that a restructuring analysis is empirically motivated. By positing an extra VP in so-called type-shifting constructions, there is no need to forgo Fregean compositionality. In addition, if there is a processing
cost associated with such constructions, this cost follows from standard assumptions regarding sentence processing: increased processing time is a function of structural complexity.

4. SOURCES OF INFORMATION: BEYOND THE LEXICON

One of our main objections to the proposals of Pustejovsky (1995) and Jackendoff (1997) — and to the empirical support offered by McElree et al. 2001, Traxler et al. 2002; and Pickering et al. 2005 — regards the type and source of information used to interpret (35a) as (35b). Is the information lexical or pragmatic? We argue for the latter.

(35)  

a. Mary began the book.

b. Mary began reading the book.

We spell out an account that requires a strict distinction between semantic and pragmatic types of representation such that the latter are responsible for filling in the verbal gap that we postulate remains in the LF of these constructions. Relevant to our present concerns, there are two issues relating to the interface between semantic and pragmatic or semantic and conceptual representations and processes. One issue has to do with the presuppositions (in particular, those triggered by verbs) that enter into the interpretation of indeterminate expressions. The other issue regards the distinction between lexical decomposition on the one hand, and inferences that are triggered by token lexical items and syntactically complex expressions on the other hand.

4.1. Presuppositions, indeterminacy, and saturation

Certain predicates entail a presuppositional state of affairs. For example, the refusal event in (36) presupposes an offering event such that \([\exists x, \exists y, \exists z, x \text{ refused } z \text{ from } y \rightarrow y \text{ offered } z \text{ to } y]\). The presupposition creates a reference frame — in discourse or in the mind of the speaker/hearer — that allows for the interpretation of (36) to refer to antecedent events.

(36) Mary refused (to accept) the gift.

Similarly, aspectual verbs such as finish entail a beginning for the event that finish applies to. The information that is presupposed in the indeterminate construction is not part of the lexical entry but can be triggered (in the form of inferences) by the tokening of the verb concept. Thus, the indeterminacy of (37) can be characterized as presupposing a relation between an activity and the two NP referents. Such indeterminacy is felicitous in contexts where background information about

\(^{15}\text{It is beyond the scope of this article to elaborate upon the details of the distinction between semantic and pragmatic types of representation that enter into the interpretation of underspecified utterances. See Carston (2002) for a thorough treatment.}\)
the relation between NP1 (John) and NP2 (the book) is known. Accordingly, for someone to utter (37), a relation $z$ (the verb-referring event) between $x$ (NP1) and $y$ (NP2) needs to be determined such that $x \text{ enjoyed } z$-ing $y$.

(37) John enjoyed the book.

In our proposal, $z$ is the filler for the silent lower V of restructuring verbs such as begin and enjoy. When the relation $z$ is not overtly specified in the sentence, it may be presupposed in the discourse. When a sentence or context does not specify such a relation between $x$ and $y$ (such that for $x$ to enjoy $y$, $z$ needs to be invoked), it is left to the hearer to assign to $z$ any given event; that is, $z$ is a semantic variable at the interface between linguistic and conceptual/pragmatic representations.

In this way, one can understand that John enjoyed the book without knowing what John did with it (how he $z$-ed it). If context provides a filler for the empty V (or a potential filler), the argument is satisfied (via discourse). If context does not provide a potential filler, indeterminacy remains up to the point where inferences — perhaps inductive and even abductive ones — are triggered. This is compatible with the claims of relevance theory, where the level of semantic representation or LF “contains variables and/or slots which require filling by pragmatic inference” (Carston 2002:185). This pragmatic process of filling variables or slots is called saturation. Notice that these gaps do not always require saturation, as when the V head has an antecedent in the sentence or context.

The possibility of pragmatically driven saturation of a variable has implications for how we are to understand the results reported in psycholinguistic studies. In the task devised by McElree et al. (2001), where subjects read type-shifting sentences in isolation, it is not necessarily the case that the purported delay found after the NP complement position — and which de Almeida (2004) fails to replicate — is due to the retrieval of lexical information that supplies the $z$ event information. One plausible explanation for the effect is the realization on the part of the reader that the construction is indeterminate because a filler for the empty V is not supplied overtly and because the presuppositional context does not supply a potential filler either. Our suggestion is, then, that increased reading times at post-verbal positions obtained in some of the experiments discussed above could be due to the lack of specification of the $z$ relation.

4.2. Denotation and inferentialism

The relation between $x$ and $y$ in the absence of an explicit $z$ event in context can be conceived of as follows: assume that for an indeterminate sentence such as John enjoyed the book (37), the appropriate gap (in syntax or in discourse) triggers a set of inferences which are pragmatically licensed (and unlicensed as well) by the denotation of the indeterminate construction.

Conceptual systems generate information based on, but going beyond, the boundaries of the meaning of token utterances. This generative capacity takes the form of inductive, deductive, and perhaps above all, abductive, inferential
reasoning triggered by the content of linguistic expressions. Inferences that are triggered by certain mental states are only bounded by the “inferential capacities of the organism” (Fodor 1990:209). We believe that much of the confusion in the lexical-semantic literature — and in particular, much of the confusion about what constitutes the information that enters into the mental states of linguistic/perceptual analyses — is due to the conflation between the denotation of an item $x$ and the set of inferences (/beliefs) about what $x$ stands for.

While thoughts are unbounded, content-bearing utterances are bounded. Because utterances are taken to be linguistic expressions of thoughts, the inferences that draw upon the content of sentences are an integral part of the process of understanding. However, it is important to recognize that there is a distinction between attributing the content of these inferences to:

(i) the semantic representation of sentences; and

(ii) processes that are triggered by the semantic representation of sentences (including structural gaps).

Recognizing this distinction is a key to our proposal for the interpretation of underspecified constructions, which adopts the following assumptions:

- Lexical items are atomistic representations: they do not encode information beyond their own denotations (Dertske 1981; Fodor 1990).
- The conceptual content of lexical item $X$ is established by the conditions under which $X$ is triggered: that is, there is a nomic relation between $X$ and its epistemically appropriately triggering events and contexts (including purely mental events).
- The above conditions suffice for the determination of the denotation of $X$.
- The sets of inferences triggered by concept $X$ — perhaps in the form of an inferential domain for $X$ (de Almeida 1999a, 1999b) — play a role in the interpretation of token linguistic utterances.
- These inferences are not necessarily analytic, although some analytic inferences may be part of the domain of $X$ and triggered by their carrier sentences, that is, in the case of presuppositions.
- The triggering factor in the interpretation of type-shifting constructions takes place in syntax and semantic structure (or LF) in the form of structurally unfulfilled positions (gaps).
- Among the sets of inferences that constitute the domain of $X$ (and the domain of expressions of which $X$ is a constituent) are those that conceptually fill in structurally defined gaps.
- Coercion effects are accounted for by classical compositionality; there is no need to postulate enriched forms of composition and lexical-analytic computations.
There are a number of linguistic and psycholinguistic arguments that could potentially account for alleged type-shifting or coercion effects without appeal to lexical-semantic decomposition. We have shown that verbs used to test online processing of type-shifting constructions are structurally different from other verbs with which they are normally compared. These structural differences challenge the type-shifting analyses proposed by Pustejovsky (1995) and Jackendoff (1997), and call into question the strength of the alleged type-shifting effects reported in psycholinguistic experiments which have been offered in support of enriched composition. In our analysis, the basic ingredients for interpreting indeterminate linguistic expressions reduce to:

(i) lexical-conceptual atomism;
(ii) classical compositionality; and
(iii) pragmatic inferences triggered by the semantic representation of sentences with structural gaps.

REFERENCES


Asher, Nicholas, and James Pustejovsky. 2000. The metaphysics of words in context. Ms., University of Texas, Austin.


