Complement Coercion

Distinguishing between type-shifting and pragmatic inferencing

Argyro Katsika[1, 2], David Braze[2], Ashwini Deo[1], and Maria Mercedes Piñango[1]
[1]Yale University / [2]Haskins Laboratories

Although Complement Coercion has been systematically associated with computational cost, there remains a serious confound in the experimental evidence built up in previous studies. The confound arises from the fact that lexico-semantic differences within the set of verbs assumed to involve coercion have not been taken into consideration. From among the set of verbs that have been reported to exhibit complement coercion effects we identified two clear semantic classes — aspectual verbs and psychological verbs. We hypothesize that the semantic difference between the two should result in differing processing profiles. Aspectual predicates (*begin*) trigger coercion and processing cost while psychological predicates (*enjoy*) do not. Evidence from an eye-tracking experiment supports our hypothesis. Coercion costs are restricted to aspectual predicates while no such effects are found with psychological predicates. These findings have implications for how these two kinds of predicates might be lexically encoded as well as for whether the observed interpolation of eventive meaning can be attributed to type-shifting (e.g., McElree, Traxler, Pickering, Seely, & Jackendoff, 2001) or to pragmatic-inferential processes (e.g., De Almeida, 2004).

Keywords: complement coercion, semantic composition, syntax–semantics interface, type-shifting, pragmatic inferences

The meaning arising from a complex expression is sometimes not fully derivable simply from the meaning of the parts that it is composed of. One way of dealing with this apparent violation of the compositionality hypothesis has been by introducing non-syntactic meaning shifting mechanisms that allow syntactic components with incompatible meanings to compose with each other. Within the linguistics literature, *coercion* is the cover term used for such resolution mechanisms

The Mental Lexicon 7:1 (2012), **58–76**. DOI 10.1075/ml.7.1.03kat
ISSN 1871–1340 / E-ISSN 1871–1375

(e.g., Jackendoff, 1997; Klein & Sag, 1985; Partee, 1987; Pustejovsky, 1995). The label has been applied to mass-count mismatches (e.g. Michaelis, 2005; Pustejovsky, 1995; Wiese & Maling, 2005), aktionsart mismatches (aspectual coercion) (e.g., De Swart, 1998; Michaelis, 2003, 2004; Piñango, 2003; Piñango, Winnick, Ullah, & Zurif, 2006; Piñango & Zurif, 2001; Piñango, Zurif, & Jackendoff, 1999; Todorova, Straub, Badecker, & Frank, 2000) and selectional mismatches (complement coercion) (e.g., Jackendoff, 1997; McElree, Pylkkänen, Pickering, & Traxler, 2006; McElree et al., 2001; Pickering, McElree, & Traxler, 2005; Pustejovsky, 1991, 1995; Traxler, McElree, Williams, & Pickering, 2005; Traxler, Pickering, & McElree, 2002). As a phenomenon involving semantic mismatches and their extra-syntactic resolution, coercion provides a window into the connection between meaning structure and morphosyntactic structure. Therein lies the significance of this constellation of phenomena.

Consider the contrast in (1-a) and (1-b). In (1-a), the aspectual verb *begin* combines with a VP that denotes the event of reading a book and the resulting sentence refers to a time at which the girl began the event of reading the book. In (1-b), *begin* combines with an NP that denotes a particular book, but the resulting sentence refers to a time at which the girl began some event involving the book.

(1) a. The girl began reading the book.
b. The girl began the book.

The observation is that the explicit morphosyntax of (1-b), [NP[V[NP]]], does not contain expressions that make reference to any event whose beginning the sentence refers to. Yet, the only coherent interpretation demands that there be such an event. Further, to the extent that *begin* is an aspectual verb that modifies temporal reference, it can be posited that the verb *begin* is restricted to arguments of a temporal or eventive nature. *The book*, which denotes an event participant, does not directly provide an argument that satisfies this lexical restriction. That (1-b) receives a coherent interpretation with an eventive component despite this mismatch indicates that a temporal/eventive argument is supplied to *begin* at some point during the interpretation of the sentence.

The experimental literature on Complement Coercion assumes that this phenomenon is exhibited by several more verbs where the presence of an NP-complement requires the interpolation of an event involving the NP-denotation. Besides *begin*, verbs taken to participate in this phenomenon include *finish*, *enjoy*, *start*, *complete*, *endure*, *master*, *attempt*, *expect*, *try*, *resist*, and *savor* (De Almeida, 2004; Frisson & McElree, 2008; McElree et al., 2001; Traxler et al., 2005; Traxler et al., 2002). Processing correlates of Complement Coercion have been detected through a variety of real-time comprehension measures such as self-paced reading and eye-tracking (Lapata, Keller, & Scheepers, 2003; McElree et al., 2006; McElree et

al., 2001; Pickering et al., 2005; Traxler et al., 2005; Traxler et al., 2002), MEG patterns (Pylkkänen, Martin, McElree, & Smart, 2009) and speed-accuracy trade-off (SAT) (McElree et al., 2006). This body of evidence has systematically shown that it is harder for readers to process expressions involving Complement Coercion than expressions that do not.

It is important to note here that the literature has not demonstrated that the class of Complement Coercion verbs exhibits any shared properties beyond the observation that in each case, the presence of an NP-complement gives rise to the inference of a related event. For instance, a sentence like *The surfer endured the tuxedo* (sample stimulus in McElree et al., 2001) gives rise to an inference that the surfer endured wearing or, perhaps, observing someone's wearing of the tuxedo. While this inference does seem to arise with all the verbs in the coercion verb set, in the absence of more robust syntactic or semantic shared properties, the hypothesis that eventive inferences must be attributed to the *same* mechanism of building meaning (coercion + type-shifting) is too strong. The possibility that these inferences arise due to type-shifting in some cases but are the result of some other mechanism in other cases cannot be eliminated.

The goal of this paper is to better understand the nature of the meaning-interpolating mechanism(s) associated with Complement Coercion. The literature presents two main hypotheses: (a) a semantic repair mechanism of *type-shifting* that coerces entity-denoting complements into event-denoting ones (e.g. Jackendoff, 1997; Pustejovsky, 1991; 1995), and (b) *pragmatic inferencing,* which involves no semantic meaning-shifting operation but only the contextualization of the result of the compositional derivation (e.g., De Almeida, 2004; De Almeida & Dwivedi, 2008).[1] By contextualization, these authors refer to the contextual retrieval of an event associated with the NP-complement rather than the semantic repair of a mismatch caused by the presence of the NP-complement. Crucially, these two possibilities have distinct computational implications: type-shifting is obligatorily triggered by mismatches between the semantic selectional restrictions imposed by a verb and the NP-complement denotation; pragmatic inferencing is not obligatory for semantic composition and occurs as sentence meaning is contextualized. The former automatically adds to cost during the processing of a sentence, while the latter may generate at most only a negligible cost (cf. Frazier & Rayner, 1990; Frisson & Pickering, 1999). Which of the two mechanisms underlies Complement Coercion?

One implicit assumption characterizing previous Complement Coercion studies is that a morphosyntactically unresolvable incongruity either exists across the board for all Complement Coercion predicates (forcing type-shifting) or does not exist at all. A closer look at the predicates employed in these studies (De Almeida, 2004; Frisson & McElree, 2008; McElree et al., 2001; Traxler et al., 2005,

Table 1. Classification of Verbs Used in Previous Complement Coercion Studies

Aspectual	Psychological	Other
begin, finish, start, complete, continue, end	endure, prefer, resist, savor, enjoy, survive	master, attempt, expect, try

2002), however, reveals that these are semantically heterogeneous (see Table 1) — falling into two broad classes — aspectual and psychological. The distinction has primarily to do with the argument structure of the predicates belonging to the two classes.

Aspectual verbs are inherently temporal in that their meaning introduces existential quantification over some event whose (initial, medial, or final) subpart is referred to by the VP containing the verb (ter Meulen, 1990). These verbs then semantically require an event-denoting complement. Their external argument (the subject) is understood to be an agentive participant in the event. The restriction to event-denoting complements observed with these verbs, thus, has a clear basis in their argument structure. The psychological verbs considered to belong to the class of Complement Coercion verbs, on the other hand, are Subject Experiencer verbs (e.g., *enjoy, tolerate*). In contrast to aspectual verbs, their external argument is an experiencer while their object argument (the complement) is entailed to be either a target of emotion or subject matter of emotion (Pesetsky, 1996, pp. 55–69).[2] The target/subject matter of emotion may be realized by different sorts of individuals such as an entity, an event, or a state-of-affairs. That is, the complement of psychological verbs *may* be event-denoting, but this is not necessarily encoded as a semantic selectional restriction in their lexical meaning. Regardless of whether the complement of a psychological verb denotes an entity or an event, as long as it expresses a target or subject matter of emotion, the compositional process involved is the same. So, for instance, with a sentence like *The composer tolerated the music score*, the music score is the target of emotion, the argument towards which the composer's emotion of toleration is directed. The fact that in order to have such an emotional response there needs to be some event in which the composer accesses the music score (by reading, playing, or listening to it) does not mean that the verb *tolerate* obligatorily selects for an event-denoting complement.[3]

To summarize, the aspectual verbs and the psychological verbs within the Complement Coercion verb set differ with respect to their argument structure, with only the former requiring event-denoting complements.

Our claim is that recognizing this distinction in argument structural properties is crucial to a better understanding of Complement Coercion and its processing correlates. Note that the verbs claimed to exhibit Complement Coercion in previous studies are a semantically heterogeneous class, unified only by the observation

that the presence of an entity-denoting complement gives rise to an eventive inference. Within this class, we have identified two clearly distinct semantic subclasses — the aspectual and the psychological verbs — which differ with respect to whether their complement is obligatorily constrained to be event-denoting. If these classes are studied in isolation from each other rather than being conflated, then Complement Coercion effects, characterizable as *type-shifting triggered in the presence of compositional mismatches,* should only be observed with aspectual verbs and not with psychological verbs. Our experiment tests this hypothesis, and in this way it seeks to sharpen our current understanding of coercion, teasing it apart from the linguistically related but computationally distinct (at the processing level) mechanism of pragmatic inferencing.

Methods

Participants

Thirty-six university students, native English speakers with reported normal or corrected-to-normal vision, participated in this study (mean age = 21.15 years, *SD* = 2.27; 17 female). Participants gave informed consent and were paid for their participation. The Yale University Human Investigation Committee approved the protocols reported here.

Apparatus

Binocular eye-movements were recorded at a sample rate of 250hz with an Eyelink II eye-tracker. Sentences were presented in a monospace font on a monitor positioned nominally 68 centimeters from subjects' eyes. Font size was such that each character subtended about 17 minutes of visual arc.

Materials

The experimental stimuli consisted of 36 triads built out of 36 two-sentence vignettes. Each vignette was composed of a context sentence and a target sentence. Each triad contained three vignettes corresponding to the three conditions tested: 1) the aspectual verb condition (requiring event-denoting complements), 2) the psychological verb condition (compatible with event- and entity-denoting complements), and 3) a control condition containing an entity-selecting verb (see Table 2). Within each triad, the vignettes were identical except for the verb in the target sentence. This generated a total of 108 vignettes. The stimuli were

Table 2. Example Stimulus Item

***Context** Sentence*	
The new interns, Alexandra and John, loved to read novels.	
Target Sentences:	
Alexandra was completing a sci-fi book when the secretary announced the meeting.	(aspectual)
Alexandra was enjoying a sci-fi book when the secretary announced the meeting.	(psychological)
Alexandra was shelving a sci-fi book when the secretary announced the meeting.	(control)

composed by native speakers of English and checked for acceptability through norming questionnaires administered to Yale undergraduate students (n = 56). Other design features of the stimuli were as follows: (a) the NP-complement denoted a physical entity with informational character (e.g., *book*, *menu*), wherein the physical facet was predominantly targeted in the control sentences and the information facet is targeted in the aspectual and psychological sentences;[4] (b) at least two words preceded the critical verb; and (c) at least five words followed the NP-complement head.

Most vignettes contained a different control verb in order to maximize variability across the test stimuli (5 verbs were used twice and 1 verb was used three times). The majority of the control verbs (31/36 instances) selected for the physical, rather than the informational aspect of the complement's meaning (e.g., *trash the newsletter* rather than *read the newsletter*). This was done for two reasons. First, the physical facet of NPs like *book*, invokes reference to their concreteness and tangibility, thus making them better members of the "entity" set. Second, the predictability of the exact complement to a verb that selects for a physical NP-complement,

Table 3. Verbs Used in each Condition (Times Used). Most of the Aspectual and Psychological Predicates in this Table Are Selected from the Pool of Predicates Used in Previous Studies (see Table 1).Other Verbs (e.g., Stomach and Face) Were Introduced After Early Norming Questionnaires Showed that some Previously Used Verbs Received Consistently Low Ratings from Native Speakers (e.g., endure).

Aspectual:	begin (8), start (7), finish (7), continue (7), complete (7)
Psychological:	enjoy (7), tolerate (6), resist (6), prefer (6), favor (5), stomach (4), face (2)[5]
Entity-selecting:	trash (3), submit (2), write (2), misplace (2), shred (2), shelve (2), sell, find, peruse, send out, send, subscribe to, auction, drop, inspect, open, destroy, buy, unpack, unearth, contribute, rent, fax, work on, photocopy, access, purchase, deliver, conduct[6]

e.g., *trash*, is less as compared to the predictability of the complement to verbs like *read* that select for an NP-complement of the informational type. This is because there is a limited number of the sorts of objects that can be "read" in contrast to the sorts of objects that can be "trashed". Following previous studies, psychological and aspectual verbs were re-used several times across the 36 vignettes. Table 3 lists the verbs used for each condition. The two main criteria for inclusion were: (1) that the presence of an NP-complement gave rise to an eventive interpretation and (2) that the verb combined grammatically with a VP-complement of the form *V-ing NP* (e.g., *begin/tolerate [reading Anna Karenina]*$_{VP}$).

Scripts

All three conditions were counterbalanced across 12 scripts or presentation lists using a blocked design to ensure even distribution of the conditions. Only one member of each triad occurred in a presentation list. Thus, each list contained 12 items from each experimental condition. Participants were randomly assigned to the lists.

Materials for this experiment were interspersed with 41 filler trials (from an unrelated study). There were additionally 4 practice items at the beginning of each session, resulting in a total of 81 trials. 35 of the 81 trials, including 12 of the critical trials, were followed by a yes/no comprehension question. These were evenly distributed across sentence conditions (aspectual, psychological and control). For example, the comprehension question following the vignette in Table 2 was *Was John a new intern?* (expected response: *yes*). Due to a programming error, 0–2 of the 36 critical trials (which included 12 each of control, psychological, and aspectual trials) were omitted from each subject's session. This eliminated 0 trials in 8 cases, 1 trial in 24 cases, and 2 trials in 4 cases. The omitted trials were evenly distributed across conditions. Every subject saw at least 11 trials for each condition.

Procedure

Participants were seated in front of the PC-monitor in a height-adjustable chair. They were briefed on the procedure, and instructed to read each vignette for comprehension with normal speed, as some would be followed by comprehension questions. The eye-tracker was calibrated using a series of nine fixed targets across the display.

Trials consisted of several events. First, a fixation target appeared at the screen position to be occupied by the sentence-initial letter. Participants fixed their gaze on this target while pressing a button to initiate the trial, at which point the context phrase appeared, all on a single line. Participants read silently, and clicked a

response key when finished. The context sentence disappeared and the sentence initial fixation target re-appeared. Participants re-fixed their gaze and pressed a button to bring up the critical sentence. They read this sentence and clicked a button when finished. This was followed by either a yes/no comprehension question or a prompt to "Press YES to continue". They responded by clicking YES/NO buttons on a response box. Feedback was provided. Eye-movements were recorded over both the context and critical sentences. Viewing was binocular, but for technical reasons (e.g., difficulty with calibration due to use of eye-glasses) eye-movements were recorded from one eye for 6 participants. Each session lasted about 50 minutes.

Results

Prior to analysis, we removed data from critical sentences with viewing times longer than 10 seconds or containing more than three blinks (1.9% of trials removed). For the purpose of aggregating reading times and classifying leftward eye-movements we divided target sentences into five regions (see Table 4). For each target sentence, corresponding regions contain the same number of words across verb conditions. Region 1 includes words from the beginning of the sentence up to but not including the main verb. Region 2 begins with the main verb and includes all words up to the head of the NP-complement. Region 3 includes the head of the NP-complement and the following word. Region 4 contains the two words that follow Region 3, unless none of them is a content word, in which case Region 4 extends so as to include at least one content word. Region 5 contains the remainder of the sentence. This segmentation allowed us to isolate the head noun of the NP-complement, while maintaining the size of the regions in a scale that would permit calculation of local reading times. The head of the NP is expected to be the trigger of coercion. Experimental findings by Traxler and colleagues (2002) support this expectation. In that study, sentences like *The boy started the puzzle after school today*, which involved entity-denoting nouns in the object position, were found to be more difficult to process than sentences with event-denoting objects, like *The boy started the fight after school today*. Notice that the sentences under comparison are almost identical, differing only in the head noun of the NP-complement.

Table 4. Example of Sentence Regions

1	2	3	4	5
Alexandra was	completing a sci-fi	book when	the secretary	announced the meeting.

Data were analyzed in terms of regional first-pass reading time, incidence of first-pass regressive eye-movements, regression-path time and second-pass reading time. *First-pass reading time* is the summed fixation durations within a region, beginning with the first fixation inside the region and ending with, but not including, the first subsequent fixation outside the region. A region is counted as having a *first-pass regression* if its final first-pass fixation ends in a backward glance to an earlier part of the sentence. Thus, for each region, incidence of first-pass regression is a dichotomous score (0/1). An initial backward saccade is often followed by additional regressive movements resulting in a chain of backward eye-movements, or a regression path (Konieczny, Hemforth, Scheepers, & Strube, 1997). *Regression-path time* includes all fixations from the first fixation in a region until the reader fixates to the right of the region. First-pass reading time, first-pass regression frequency and regression-path time are all contingent upon there being a first-pass fixation of at least 50 ms within the region (Carpenter & Just, 1983). Regions not meeting this criterion are excluded from the analyses. *Second-pass reading time* is the sum of all fixation durations in a region that are *not* first-pass fixations, including time spent in the region after exiting to the left or to the right.

Regional data were analyzed using mixed effects models incorporating a fixed effect of verb type (3 levels) and random intercepts for subjects and items (Baayen, Davidson, & Bates, 2008). Analyses were carried out in the R statistical environment (R Development Core Team, 2009), using the LME4 and LanguageR packages. Our main regions of interest were the verb-containing region 2, and region 3. Analysis of region 4 data was included because some eye-tracking studies of Complement Coercion have found "spillover" effects (e.g., Traxler et al., 2002). Region 5 was analyzed, because some types of semantic processing have been shown to modulate end-of-sentence "wrap-up" effects (Braze, Shankweiler, Ni, & Palumbo, 2002). Table 5 provides regional summaries of eye-movement measures by verb condition.

For each response measure for each region, the main effect of verb type on the response measure was assessed by contrasting a base model with intercept only for fixed effect against a model adding the verb-type factor, both of which included random intercepts for subjects and items. In the case of continuous response measures (first-pass time, regression path time and second-pass time), *p*-values for pair-wise comparisons were estimated using Markov chain Monte Carlo (MCMC) estimation (Baayen et al., 2008). In the case of the first-pass regressions, a similar approach was used, but incorporating a logit link function to accommodate the binomial nature of the response. As MCMC estimation routines are not available for logistic models in the LanguageR and LME4 libraries, we evaluated pair-wise comparisons using Wald's Z-statistic. All significant contrasts are reported.

Table 5. Mean (SD) and Regional Reading Time by Verb Type. Cells in Bold and Superscripts Note Significant Contrasts. A, P, and C Stand for Aspectual, Psychological and Control Verbs Respectively.

	First-pass time		First-pass regression		Regression path time		Second-pass time	
Region 2								
Aspectual	400	(261)	.07	(.26)	444	(287)	274	**(366)**$^{>P,C}$
Psychological	403	(255)	.06	(.25)	439	(277)	187	(291)
Control	422	(288)	.08	(.27)	481	**(334)**$^{>A,P}$	185	(303)
Region 3								
Aspectual	379	(246)	.20	**(.40)**$^{>P,C}$	495	(352)	209	(339)
Psychological	380	(218)	.13	(.34)	460	(295)	170	(293)
Control	390	(244)	.16	(.36)	489	(382)	169	(319)
Region 4								
Aspectual	334	(191)	.16	(.37)	477	(491)	169	(262)
Psychological	348	(228)	.14	(.35)	433	(329)	155	(268)
Control	337	(200)	.14	(.35)	430	(445)	158	(279)
Region 5								
Aspectual	688	(500)	.62	**(.48)**$^{>C}$	1410	(1059)	198	(419)
Psychological	698	(529)	.57	(.50)	1254	(884)	163	(346)
Control	695	(506)	.53	(.50)	1247	(882)	181	(422)

Region 2

The main effect of verb type on second-pass time was assessed by contrasting a base model with intercept only for fixed effect against a model adding the verb-type factor, both of which included random intercepts for subjects and items. The result indicates a robust main effect of verb type [$\chi^2(2) = 25.37$; $p < .001$]. *P*-values for pair-wise comparisons were estimated using Markov chain Monte Carlo estimation. Comparisons showed that readers spent more time re-reading the verb in the aspectual condition than the control ($p < .001$) or psychological condition ($p < .001$). Second-pass reading times for the psychological and control verb conditions did not differ reliably. Results for Region 2 suggest that while readers initially had no more difficulty in processing aspectual verbs than they had in processing psychological or control verbs, the aspectual verb condition became more difficult later in the sentence, causing more refixations on the verb, as reflected in second-pass reading times.

A main effect of verb type on regression-path reading time was detected [$\chi^2(2) = 8.87$; $p = .012$]. Pair-wise comparisons showed that regression-path reading times were significantly longer in the control verb condition than both aspectual ($p = .022$) and psychological ($p = .009$) verb conditions. Regression-path reading times for the psychological and aspectual verb conditions did not differ ($p = .669$). We believe that the longer regression-path reading time for the control condition in region 2 emerges from the choice of control verbs, which were deliberately selected so as to be less predictable in the given context.

Region 3

Effects of first-pass regressions were analyzed using a mixed effect model for binomial response (logistic regression), as described above. The result indicated a significant main effect of verb type on first-pass regressions [$\chi^2(2) = 8.58$; $p = .014$]. Pair-wise comparisons revealed that complements to aspectual verbs evoke more regressive eye-movements than complements to psychological [coefficient = -.56; Wald's Z = −2.83; $p = .005$] and control verbs [coefficient = .37; Wald's Z = 1.92; $p = .054$]. The control verb condition did not differ from the psychological verb condition [coefficient = -.20; Wald's Z = -.97; $p = .329$].

The distinction between aspectual verbs on the one hand, and psychological and control verbs on the other, was also observed in second-pass reading times [$\chi^2(2) = 5.21$; $p = .073$]. Pair-wise comparisons indicated that second-pass reading times were significantly longer in the aspectual than the control verb condition ($p = .040$) and marginally longer in the aspectual than the psychological verb condition ($p = .062$). Second-pass reading times for the psychological and control verb conditions did not differ ($p = .822$).

Region 4

No difference was observed in any measure for this region.

Region 5

Here we found strong trends toward significant main effects of verb condition on first-pass regression rate [$\chi^2(2) = 7.36$; $p = .025$]. Based on the pair-wise comparisons, the aspectual condition induced more regressions than the control condition from the sentence-final region [coefficient = .418; Wald's Z = 2.74; $p = .006$]. The psychological verb condition showed an intermediate regression rate, not reliably different from either the aspectual [coefficient = -.246; Wald's Z = −1.57; $p = .116$] or the control [coefficient = .168; Wald's Z = −1.11; $p = .265$] condition. A main

effect of verb condition on regression-path time was also found [$\chi^2(2) = 10.361$; $p = .006$]. Regression-path time was greater in the aspectual than the psychological [$p = .015$] and control verb conditions [$p = .004$]. The psychological and control conditions did not differ from each other [$p = .829$].

In order to better facilitate the interpretation of these results, we conducted an offline judgment test to assess the general plausibility of the experimental stimuli, as well as to determine the relative plausibility of the aspectual and the psychological verb stimuli.

Offline Plausibility Judgment Test

Ten randomized lists were constructed containing the 108 coercion vignettes together with an additional 33 "nonsense" vignettes, which contained a nonsense target sentence. The latter were minimally different from the original test sentences in form. For example, the nonsense counterpart of the sentences in Table 2 was the following: *Alexandra was believing a sci-fi book when the secretary announced the meeting*. Twenty-two native speakers of American English from the same population as the eye-tracking participants rated the 141 vignettes in terms of plausibility. Their task was the following:

> *"We want to know whether the target sentence is one that you might expect a native speaker of English to speak or write, given the relevant context. Please, rate the plausibility of the sentences on a scale of 1–5 (1 = Implausible, 5 = Completely plausible)."*

The ratings were submitted to a mixed effects model analysis with sentence type (4 levels) as fixed effect and including random intercepts for subjects and items (Baayen et al., 2008). Results indicated a main effect of sentence type [$\chi^2(2) = 1549.7$; $p < .001$]. Aspectual, psychological, and control sentences were rated as significantly more plausible than nonsense sentences ($p < .001$), thus confirming our pre-test plausibility evaluations. In particular, the sentences in all three experimental conditions were rated well above the midscale point of plausibility (range 3.73 to 4.09: Aspectual: mean = 3.73, $SD = 1.33$, Psychological: mean = 3.98, $SD = 1.21$ and Control: mean = 4.09, $SD = 1.17$). In contrast, the nonsense sentences received a mean rating of 1.67 ($SD = 1.04$).

Pair-wise comparisons revealed further distinctions: control sentences received a higher mean rating than both aspectual ($p < .001$) and psychological sentences ($p = .065$), and aspectual sentences received a lower mean rating than psychological sentences ($p < .001$).

To summarize, in confirmation of our pre-test evaluations, the experimental stimuli from all three conditions were judged highly plausible by native speakers. At the same time, speakers do make a systematic distinction between the control,

aspectual, and psychological conditions. However, we note that the difference in plausibility cannot explain the effects of the eye-tracking results reported in the previous section. For example, if plausibility were the crucial driver behind those results, we would expect the control sentences, because of their higher plausibility, to have uniformly shorter reading times than the aspectual and psychological sentences. This was not the case. Further, we conducted an analysis of the effects of plausibility and verb type on the total reading time of the test sentences, while controlling for sentence length (measured as number of characters). First, the effect of plausibility on total reading time was assessed by contrasting a base model with intercept for verb type and fixed effect against a model adding the plausibility factor, both of which included random intercepts for subjects and items. Plausibility did not account for the variance observed in the data beyond the factor of verb type [$\chi^2(3) = 2.7032$; $p = .44$]. On the other hand, contrasting a base model for plausibility against a model adding verb type revealed a significant effect of the latter [$\chi^2(4) = 12.105$; $p = .017$]. These analyses indicate that the factor verb type explains unique variance in the total sentence reading time above and beyond the effect of plausibility, whereas the reverse is not true. Thus, these analyses further suggest that plausibility is not behind the eye-movement results reported in the previous section.

In order to further assess whether the lower plausibility of the aspectual condition sentences contributed to the processing difficulty indicated by the eye-tracking results, we performed an additional screening of the experimental stimuli, selecting those triads from the larger set of 36 triads that exhibited minimal difference in ratings across the three conditions. This criterion allowed us to select 28 out of the 36 triads, which we then subjected to statistical analysis. The analysis corroborated that there was no significant difference across conditions [$\chi^2(2) = 3.37$; $p = .185$]. Pair-wise comparisons showed that the three conditions did not differ from each other. Specifically, the control verb condition did not differ either from the aspectual ($p = .103$) or the psychological verb condition ($p = .937$) and the aspectual verb condition did not differ from the psychological verb condition ($p = .121$).

We then performed a mixed model analysis of the eye-tracking data for these 28 triads. We predicted that if the eye-tracking findings were independent of the acceptability patterns, the contrast between the aspectual and control conditions should persist despite the reduced number of stimuli and the consequent decrease in power. This prediction was borne out by the results.

Specifically, a main effect of verb type on second-pass reading times for region 2 was detected [$\chi^2(2) = 6.24$; $p = .044$], with these times being longer in the aspectual than both the psychological ($p = .041$) and control ($p = .021$) conditions. The psychological and control conditions did not differ from each other ($p = .663$).

The analysis showed a marginally significant effect of verb type on regression-path reading time in region 2 [$\chi^2(2) = 5.04$; $p = .080$]. Pair-wise comparisons indicated longer regression-path times in the control than the psychological ($p = .058$) and aspectual ($p = .090$) conditions. The two latter conditions did not differ from each other ($p = .769$).

Verb type was found to have a marginal effect on regression-path times in region 5 [$\chi^2(2) = 4.92$; $p = .085$]. These times were greater in the aspectual than in either the control ($p = .049$) or the psychological condition ($p = .080$). The control and the psychological conditions did not differ ($p = .816$).

To summarize, the eye-tracking results based on the subsetted data parallel findings of the whole dataset, modulo decrease in power. This eliminates the possibility that the difference between the aspectual and control conditions observed in the eye-tracking results was due to the plausibility differences observed in the offline test. These differences, as has been claimed by others before us, can instead be reliably attributed to the computational cost arising from the composition of the (aspectual) verb and an entity denoting complement.

Crucially for the present effort, these results also failed to show any difference between the behavior of psychological verbs and control verbs. This lack of difference would be totally unexpected in view of previous work, since most of the psychological verbs in our study have been claimed to be coercion verbs (See Table 1).[7] The contrast between the patterning of the aspectual and psychological conditions suggests that the composition of psychological verbs with an entity-denoting complement does not trigger computational cost.

Discussion

Our results show that the cost that has been associated with Complement Coercion is manifested only in the aspectual verb condition, that is, for a semantically isolable class of verbs. The psychological verb condition, when tested against the control and aspectual conditions, does not exhibit this cost.[8] This observation of diverging patterns is not directly reconcilable with an approach to Complement Coercion that attributes the cost to either type-shifting (Frisson & McElree, 2008; McElree et al., 2001; Traxler et al., 2005, 2002) or to the presence of extra syntactic structure (De Almeida & Dwivedi, 2008). The presence of cost in a subclass of verbs suggests that pragmatic inferential processes cannot fully explain the Complement Coercion phenomenon. In order to make sense of this pattern, we must tease apart the syntactico-semantic properties of verbs from the inferences they license in sentential context. This separation has already been articulated in the literature. Pickering et al. (2005), attribute the observed cost of coercion to the

"compositional operations needed to generate a representation for the expression" and not to "the processes that are responsible for retrieving or inferring the activity implicit in the event structure." (p. 14). De Almeida & Dwivedi (2008) attribute the cost of coercion to the processing of the extra compositional structure triggered by restructuring verbs and not to the determination of the event that semantically corresponds to the empty verb head present in the syntactic structure. It follows from this that if there is no apparent cost for a subset of the so-called coercion verbs (i.e., psychological verbs), then it must be the case that there is neither type-shifting nor restructuring involved in the composition of these verbs with entity-denoting complements. How then does the eventive reading arise when these psychological verbs combine with entity-denoting complements? We propose that this eventive reading is the effect of an inferential process that invokes an activity conceptually associated with the target of emotion. Such an inferential process is crucially independent of the cost-engendering type-shifting operation or the more complex syntactic structure of restructuring verbs, and from a processing perspective, cost free (e.g., Frazier & Rayner, 1990; Frisson & Pickering, 1999). In the case of aspectual verbs, on the other hand, where the verb *must* combine with an eventive complement, some such extra apparatus is necessary to trigger the inferential process that facilitates retrieval of the specific event involved.

Thus, by isolating the processing behavior of aspectual and psychological verbs, we are able to tease apart two distinct sources of eventive meaning: compositional versus inferential. Aspectual verbs when combined with entity-denoting complements must take recourse to both compositional and inferential processes; only the former process is costly. For psychological verbs, construal of an event in the presence of an entity-denoting complement is exclusively the result of inferential processes. The absence of type-shifting or restructuring operations is reflected in the absence of processing cost.

Author Note

The research reported here was supported by NSF grant BCS-0643266 to Maria Mercedes Piñango and by NIH grant HD40353 to Haskins Laboratories. The order of authors is random and the presented work reflects a collaborative effort with equal participation of the four coauthors. We thank Ray Jackendoff, James Pustejovsky and participants in the Yale LingLunch audience (Fall 2009) for constructive feedback. All errors remain our own.

Notes

1. De Almeida and Dwivedi (2008) propose that Complement Coercion verbs are syntactically distinct from their control counterparts. They are restructuring verbs and the syntactic structure of sentences containing such verbs with an NP-complement (e.g., *begin the book*) contains an extra verb phrase with an empty verbal head. Pragmatic inferential processes contribute to identifying the contextual predicate (e.g., *read, write*) that can fill this empty head.

2. The distinctions we make here rely heavily on Pesetsky's (1996) analysis of psychological verbs. Pesetsky distinguishes between Subject Experiencer verbs like *fear* and *like* and Object Experiencer verbs like *frighten* and *annoy*. The psychological verbs within the complement coercion set belong to the former class. According to Pesetsky, the complement of Subject Experiencer verbs, when it is the target of emotion, constitutes an argument that is evaluated by the Experiencer as part of the emotional episode. As the subject matter of emotion, the emotion is directed towards the complement denotation.

3. A reviewer expresses the concern that even if the target of emotion might be an entity, entities are necessarily experienced via events or states-of-affairs. We agree that this might be the case and in fact, it is this strong association with the experience of entities via participation in events involving them that has led to the conflation of these classes in previous studies. Our current goal is to investigate whether this association between arguments of psychological verbs and the events they invoke has the same source as the association between the arguments of aspectual verbs and the events they invoke. To the extent that our study indicates a difference between the processing profiles of the two classes, it suggests that the eventive inference that arises in both cases might be due to distinct underlying processes.

4. An anonymous reviewer suggests that this asymmetry in the physical versus informational facets of meanings could create a confound if there is a processing difference between the two facets. While this is a distinct possibility, we note that it is intrinsically controlled for in our design via the contrast between control verbs (which target the physical or material facet of the NP-complement meaning) and psychological verbs (which target the informational facet of NP-complement meaning). As we show below, there is no difference between these two conditions in the crucial windows (regions 2 and 3), suggesting that this difference is not playing a crucial role here.

5. According to an anonymous reviewer, it is not clear whether *resist* and *face* have a psychological reading. A closer look at the meanings of these verbs given in the Oxford English Dictionary indicates that that their psychological reading is quite robust and has been stably attested for several centuries. Moreover, the vignettes for these verbs were carefully constructed so as to induce a psychological interpretation.

6. The verbs *write*, *contribute*, *work on*, *conduct* and possibly *subscribe to* target the informational aspect of their complement. The rest of the control verbs target its physical aspect.

7. The exceptions are *tolerate* (six instances) and *face* (two instances) both of which fit the criterion for the selection of Complement Coercion verbs provided in previous literature, that is that the presence of an NP-complement requires the interpolation of an event involving the NP-denotation.

8. Non-aspectual, non-psychological verbs used in other Complement Coercion studies (e.g., *master*, *try*, *attempt*) were purposely excluded in this experiment. Upon initial inspection, they appear to be semantically heterogeneous and we believe that they warrant closer study for any meaningful prediction to be made about their behavior.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390–412.

Braze, D., Shankweiler, D., Ni, W., & Palumbo, L. C. (2002). Readers' eye movements distinguish anomalies of form and content. *Journal of Psycholinguistic Research*, *31*, 25–44.

Carpenter, P. A., & Just, M. A. (1983). What your eyes do while your mind is reading. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 275–307). New York, NY: Academic Press.

De Almeida, R. G. (2004). The effect of context on the processing of type-shifting verbs. *Brain and Language, 90*(1), 249–261.

De Almeida, R. G., & Dwivedi, V. D. (2008). Coercion without lexical decomposition: Type-shifting effects revisited. *Canadian Journal of Linguistics, 53*(2/3), 301–326.

De Swart, H. (1998). Aspect shift and coercion. *Natural Language & Linguistic Theory, 16*(2), 347–385.

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

Frisson, S., & McElree, B. (2008). Complement Coercion is not modulated by competition: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory and Cognition, 34*(1), 1–11.

Frisson, S., & Pickering, M. J. (1999). The processing of metonomy: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory and Cognition, 25*(6), 1366–1383.

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

Klein, E., & Sag, I. A. (1985). Type-driven translation. *Linguistics and Philosophy, 8*(2), 163–201.

Konieczny, L., Hemforth, B., Scheepers, C., & Strube, G. (1997). The role of lexical heads in parsing: Evidence from German. *Language and Cognitive Processes, 12*(2), 307–348.

Lapata, M., Keller, F., & Scheepers, C. (2003). Intra-sentential context effects on the interpretation of logical metonymy. *Cognitive Science: A Multidisciplinary Journal, 27*(4), 649–668.

McElree, B., Pylkkänen, L., Pickering, M. J., & Traxler, M. J. (2006). A time course analysis of enriched composition. *Psychonomic Bulletin & Review, 13*(1), 53–59.

McElree, B., Traxler, M. J., Pickering, M. J., Seely, R. E., & Jackendoff, R. (2001). Reading time evidence for enriched composition. *Cognition, 78*(1), 17–25.

Michaelis, L. A. (2003). Headless constructions and coercion by construction. In E. J. Francis & L. A. Michaelis (Eds.), *Mismatch: Form-function incongruity and the architecture of grammar* (pp. 259–310). Stanford, CA: CSLI Publications.

Michaelis, L. A. (2004). Type shifting in construction grammar: An integrated approach to aspectual coercion. *Cognitive Linguistics, 15*(1), 1–68.

Michaelis, L. A. (2005). Entity and event coercion in a symbolic theory of syntax. In J-O. Östman & M. Fried (Eds.), *Construction grammars: Cognitive grounding and cross-language dimensions* (pp. 45–88). Amsterdam: Benjamins.

Partee, B. (1987). Noun phrase interpretation and type-shifting principles. In J. Groenendijk, D. de Jongh, & M. Stokhof (Eds.), *Studies in discourse representation theory and the theory of generalized quantifiers* (Grass Series, Vol. 8, pp. 115–143). Dordrecht: Foris.

Pesetsky, D. (1996). *Zero syntax: Experiencers and cascades*. Cambridge, MA: The MIT Press.

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

Piñango, M. M. (2003). Events: Processing and neurological properties. In H. Härtl & H. Tappe (Eds.), *Mediating between concepts and grammar* (pp. 277–301). Berlin: Mouton de Gruyter.

Piñango, M. M., Winnick, A., Ullah, R., & Zurif E. B. (2006). Time-course of semantic composition: The case of aspectual coercion. *Journal of Psycholinguistic Research*, *35*(3), 233–244.

Piñango, M. M., & Zurif, E. B. (2001). Semantic operations in aphasic comprehension: Implications for the cortical organization of language. *Brain and Language, 79*(2), 297–308.

Piñango, M. M., Zurif, E. B., & Jackendoff, R. (1999). Real-time processing implications of enriched composition at the syntax–semantics interface. *Journal of Psycholinguistic Research, 28*(4), 395–414.

Pustejovsky, J. (1991). The generative lexicon. *Computational linguistics, 17*(4), 409–441.

Pustejovsky, J. (1995). *The generative lexicon*. Cambridge, MA: MIT Press.

Pylkkänen, L., Martin, A. E., McElree, B., & Smart, A. (2009). The Anterior Midline Field: Coercion or decision making? *Brain and Language, 108*(3), 184–190.

R Development Core Team (2009). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

ter Meulen, A. (1990). English aspectual verbs as generalized quantifiers. In A. L. Halpern (Ed.), *The proceedings of the ninth West Coast Conference on Formal Linguistics* (pp. 347–360). Stanford, CA: Center for the Study of Language and Information.

Todorova, M., Straub, K., Badecker, W., & Frank, R. (2000). Aspectual Coercion and the on-line computation of sentential aspect. *In Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society* (pp. 3–8).

Traxler, M. J., McElree, B., Williams, R. S., & Pickering, M. J. (2005). Context effects in coercion: Evidence from eye-movements. *Journal of Memory and Language, 53*(1), 1–25.

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*(4), 530–547.

Wiese, H., & Maling, J. (2005). Beers, kaffi, & schnaps: Different grammatical options for restaurant talk coercions in three Germanic languages. *Journal of Germanic Linguistics, 17*(1), 1–38.

Corresponding author's address

Argyro Katsika
Department of Linguistics
Yale University
370 Temple Street
New Haven
CT 06520–8366
USA
Phone: 1-203-436-2867. Fax: 1-203-432-2487

argyro.katsika@yale.edu