Priming Unaccusative Middles: Flexible Argument Structures and the Online Processing of Middle Constructions.

Juliana Meyohas

A Thesis

In the Department

of

Psychology

Presented in Partial Fulfillment of the Requirements

for the degree of Master of Arts (Psychology) at

Concordia University Montréal, Québec, Canada

September 2012

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: Juliana Meyohas Moreira Maia

Entitled: Priming Unaccusative Middles: Flexible Argument Structures and the Online Processing of Middle Constructions

and submitted in partial fulfillment of the requirements for the degree of

Master of Arts (Psychology)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Chair: Wayne Brake, Ph. D.

External to Program: Anna Maria Di Sciullo, Ph. D

Examiner: Norman Segalowitz, Ph. D

Examiner: Aaron Johnson, Ph. D.

Thesis Supervisor: Roberto G. de Almeida, Ph. D.

Approved by:

Graduate Program Director: Wayne Brake, Ph. D.

Dean of Faculty: Brian Lewis, Ph. D.

Date: September 21, 2012

ABSTRACT

Priming Unaccusative Middles: Flexible Argument Structures and the Online Processing of Middle Constructions.

Juliana Meyohas, M.A.

Middle constructions, such as *This book sells well*, are formed with a transitive verb, but with an object figuring in subject position (*This book*). It has been proposed that these constructions either project a transitive structure (e.g., Bowers, 2002) or an unaccusative one (e.g., Rapoport, 1999). Recent studies (Di Sciullo, de Almeida, Manouilidou, & Dwivedi, 2007; Maia, Oliveira, & dos Santos, submitted) seem to suggest that middle constructions are more complex to process than their transitive counterparts. We contrasted middles constructions against syntactically simplex transitive constructions, and syntactically complex transitive and unaccusative constructions, namely passives and predicatives. We kept the verbs constant across conditions, and controlled for subject animacy. Results from Experiment 1 (self-paced reading) and Experiment 3 (eye-tracking) indicated a facilitation on the processing of middles when compared against their transitive counterparts. Results from Experiment 2 (an off-line rating task), indicated greater complexity associated with the comprehension of middles when compared against transitives, but not when compared against their passive and predicative counterparts. We suggest that, once materials were composed of two clauses in a contrastive coordinated relation (e.g., *That stucco is rough, but this wall paints smoothly*), information pertaining to the unaccusative predicative clause in the first conjunct (*That stucco is rough*) influenced the processing of the target sentences in the second conjunct. Our findings are consistent with the proposal that middle constructions

project an unaccusative structure and that argument structure information is available during online processing.

Acknowledgements

I thank my supervisor Dr. Roberto G. de Almeida for his support and insight during the conduction of this project.

I thank my loved ones, Regina, Sassi, and Chomsky—the feline—for their unconditional support.

I would like to thank my dear peers Clarisse Longo dos Santos, Christophe Steele, Carlos Roncero, Levi Riven, John Brand, Nabil Khoury, and Julianne Pohlman for all the academic and non-academic exchanges. I also thank our undergraduate volunteers, especially those more directly associated with this project, namely Anna Koziol, Emmanuela Tedone, and Marinee Kavoukjiam. A very special thanks goes to Cynthia Jabre for her commitment and assistance with data collection and data coding.

I would like to thank Dr. Anna Maria Di Sciullo for the valuable theoretical debate.

I would like to thank Dr. Aaron Johnson for the joy in sharing his knowledge of eye-tracking technology.

I would like to thank Dr. Norman Segalowitz for his kindness and academic guidance.

And I could not forget to thank Shirley Black for handling the biggest file in her drawer.

This research has been funded by the Conselho Nacional de Desenvolvimento Científico e Tecnológico - "National Counsel of Technological and Scientific Development"(CNPq), the Social Sciences and Humanities Research Council of Canada, and the Concordia School of Graduate Studies.

Table of Contents

List of Tables

List of Figures

Priming Unaccusative Middles: Flexible Argument Structures and the Online Processing of Middle Constructions

In the present study, we investigated the psychological reality of middle constructions, such as *This wall paints smoothly*. These seemingly intransitive constructions are in fact formed with verbs that usually appear in transitive constructions (e.g. *paint* in *John painted this wall*). In middles, it is the affected entity or object of the verb (e.g., *this wall*) what appears in subject position. This phenomenon is of particular interest to psycholinguistics because the online language comprehension system (in particular, the syntactic *parser*) is presented with a challenge: Is the constituent appearing in subject position the actual subject or is it the object of the sentence? Middle constructions should force a syntactic reanalysis and thus should be more complex to parse than more canonical transitive sentences.

Middle constructions can be defined (i) as having an entailed but not expressed agent, (ii) as often presenting a verbal adverb, e.g., *smoothly*, and (iii) as not allowing for specific time reference, e.g. * *This wall paints frequently* (Levin, 1993; but also see: Ackema & Schoorlemmer, 1994, 2003; Lekakou, 2005; Marelj, 2004; Steinbach, 2002; Wenzhong, 2005). In the current research, we investigated whether these constructions are more complex structures to parse than their transitive counterparts. This is done also as a way to shed light on the linguistic debate as to whether the verb in middle constructions projects to syntax a transitive (e.g., Boweres, 2002), or an intransitive structure (e.g., Rapoport, 1999). We contrasted the acceptability, and the online processing patterns of middles constructions against those of more canonical transitive constructions, such as *This worker paints smoothly*, and *This brush paints smoothly*

which include an understood object (an *implicit argument*; what is painted, in the examples).

Before we present further details, it is important to lay out some basic linguistic concepts regarding argument structures, verb types, and syntactic operations, which are required to the understanding of the topics discussed in the present research.

Verbs, Arguments, and Syntax

Among linguists, the verb is believed to be the key constituent of any given sentence. It is the verb that determines the required grammatical constituents of a sentence. This information, possibly encoded with the verb, is what is known as argument structure (for an overview, see Haegeman, 1994). It is also believed that verb meaning specifies the very semantic nature of argument structure (Jackendoff, 1989; Levin, 1993). The argument structure of a verb would, in summary, contain information regarding the number, the position, the type, and the semantic properties of the arguments required by that verb. For instance, the argument structure of the transitive verb *paint* by hypothesis contains two arguments: a Noun Phrase (NP) with an agent, e.g., *John*, as the external argument, and a NP with an affected entity, e.g. *the wall*, as the internal argument. From a linguistics perspective, argument structure is first generated in the syntactic derivation of a sentence (Chomsky, 1981). The external argument may occupy the sentential subject position and the internal argument may occupy the sentential object position. In the syntax, some operations (such as move-α, Chomsky, 1981), involve argument movement. These operations can produce a surface sentence structure which displays a different linear order than that initially projected to syntax. We provide an example of syntactic derivation in the following paragraphs, where we present the derivation of constructions

containing certain types of intransitive verbs. Such linguistically based derivational analyses can provide useful support for supposing that the distinctions being made among different types of constructions might manifest themselves in a meaningful way in psychological tests of language processing.

Under the government and binding theory (Chomsky, 1981), it is the verb that assigns a semantic role, known as a *theta-role* (θ-role)—such as that of *agent* (the "doer" of the action), *instrument* (the "tool" used to perform the action), or *theme* (the affected entity)—to each of its arguments (see also Dowty, 1991, and Gruber, 1965). In the syntax, besides θ-role, each argument is to receive *case* according to the position it takes in the syntactic structure. The verb assigns *nominative case* to the *specifier* (Spec) position, position in which the external argument is generated, and assigns *accusative case* to the complement position, position in which the internal argument is generated. So, a sentence would have the basic configuration as:

[Spec [$_{\text{Verbal Phrase (VP)}}$ [V Complement]]].

Ultimately, for an argument to be realized, it must have *case* and *θ-role*.

There are two basic types of verbs: transitives and intransitives. Transitive verbs require both an internal and an external argument. In the transitive construction *John painted the wall*, the NP subject *John* is the verb's external argument, generated in Spec position. *John* receives both θ-role of *agent* and *nominative* case from the verb. The NP object *paper* is the verb's internal argument, generated in complement position. *Paper* receives the θ-role of *theme* and *accusative* case.

Conversely, intransitive verbs require only one argument, which occupies the subject position in the surface structure of the sentence. In the intransitive constructions

John smiled, and *John fell*, the NP *John* is the only argument required by the intransitive verbs *smile* and *fall*. However, there are two distinct types of intransitive verbs: unergatives, such as *smile*, and ergatives, such as *fall*. Unergatives, also known as real intransitives (Kayser & Roeper, 1984), require only an *agent* external argument. In the sentence *John smiled*, the NP *John* is the agent preforming the action of smiling. In this example, the unergative external—and only—argument *John* is generated in Spec position, and receives θ-role of *agent* and *nominative* case. In contrast, ergative verbs require only a *theme* internal argument. In the sentence *John fell*, the NP *John* is the entity affected by the action of falling. In this example, the ergative internal—and only—argument *John* is generated in complement position, where it receives θ-role of *theme*. But, because ergatives are believed to fail to assign *accusative* case, their internal argument is raised in syntax to Spec position in order to get the *nominative* case (Burzio, 1986; Perlmutter, 1978). The surface structure of constructions containing ergative verbs, then, displays a *theme* NP in subject position, even though such argument was generated in complement position. We explain the derivation of ergatives in greater details, in the following section.

Returning to the topic of the mapping between argument structure information and syntactic structures, an important issue is that of θ-role hierarchy. Many linguists (e.g., Baker,1987; Grimshaw,1990; and Jackendoff, 1990) have proposed that θ-roles obey a hierarchy with the most prominent roles being assigned consecutively to the different arguments specified by a verb. For instance, the role of *agent* should take precedence in occupying the prominent subject position of a structure (e.g. *John cut this bread easily*). In the θ-role hierarchy, the role of *agent* is followed by that of *instrument*

(e.g. *The knife cut the bread easily*), which is then followed by that of *theme* (e.g., *The bread cuts easily*) (Baker, 1987). Psychologically speaking, if it is true that the θ-role hierarchy has psychological reality, then such a hierarchy can be interpreted in terms of canonicity (Manouilidou, de Almeida, Schwartz, & Nair, 2009). This way, the most canonical, and preferred, construction would display an *agent* as the subject, and a *theme* as the object. An example of a canonical and preferred construction can be given by the transitive construction *John transported the cargo*, in which *John* is the *agent* subject and *the cargo* is the *theme* object. A less canonical construction would display an *instrument* in subject position. This is the case of the transitive construction *The knife cut the paper*, in which *The knife* is the *instrument* subject. And the least canonical construction would display a *theme* in subject position. This is the case of the passive construction *The paper was cut into pieces,* in which *paper* is the *theme* appearing in subject position. And this is also the case of middle and ergative constructions, such as *This paper reads easily* and *The vase broke yesterday,* in which the respective subjects, *This paper* and *The vase,* are *themes.*

The Middle Problem

Middle constructions have been traditionally described in normative grammars of different languages in terms of grammatical voice. The grammatical voices in which a transitive verb can appear are the active voice, the passive voice, and the middle voice. The active voice, e.g., *John read the paper easily*, displays an agent, e.g. *John*, in subject position, and an affected entity, e.g. *this paper*, in object position. The passive voice, e.g., *This paper was read easily by John*, displays the affected entity, e.g., *This paper*, in subject position and allow for an optional agent as an adjunct in the *by* phrase.

Constructions such as *This paper reads easily* fall under the scope of the middle voice because they only display the affected entity, e.g. *This paper*, in the subject position, but prevents the realization of an agent—half way through the derivation from the active to the passive voice—thus, the denomination of middle voice.

In fact, middle constructions, such as in (1a) can be understood as an alternation from canonically transitive verbs (Levin, 1993). That is, verbs, which usually appear in transitive constructions of the type NP-Verb (V)-NP, as shown by the transitive active (henceforth referred to as transitive) sentence in (1b) and by the transitive passive (henceforth referred to as passive) sentence in (1c), can also appear in a seemingly intransitive construction of the type NP-V-(XP[1]), as it is the case of the middle construction shown in (1a).

(1) a. This wall paints easily.

b. John painted this wall.

c. This wall was painted by John.

The fact that the external argument of these canonically transitive verbs is not realized raises questions as to whether (i) the verb in middle constructions project a two-place structure, similarly to transitive constructions, as it is the case for the passive construction in (1b) (Alexiadou & Doron, 2012; Bowers, 2002), or whether (ii) the middle verb selects only an internal argument, as it is proposed to be case of ergative verbs (Fagan, 1988; Di Sciullo, 2005). If middle verbs project to syntax a one-place

[1] The XP stands for any non-obligatory phrasal constituent. In the case of middles, the XP can represent AdvP (Adverbial Phrase), referring to the adverbial adjunct, e.g., *easily* in (1c).

structure, this scenario would preclude an alternation, or a shift from the canonical transitive representation of the verbs undergoing middle construction.

From a psycholinguistic perspective a challenge to the language comprehension mechanism is how to structurally analyze and understand a construction that displays the object in subject position while blocking the realization of the agent. If middle constructions do indeed constitute a type of verb alternation, then linguistic information pertaining to the alternating verb has to somehow be made available during online processing. Thus, the investigation of middle constructions becomes of relevance in furthering our knowledge of how linguistic principles are represented and how they operate in real time.

In the linguistic literature, middles have been proposed to map onto syntax a transitive structure (e.g., Bowers, 2002). But middles have also been proposed to map onto syntax a reduced argument structure, in which only the complement is required (Ackema & Schoorlemmer, 1995, 2003).Thus far, the incipient psycholinguistics literature on the topic seems to present evidence that middles are more difficult to process than their transitive counterpart (Di Sciullo, de Almeida, Manouilidou, & Dwivedi, 2007; Maia, Oliveira, & dos Santos, submitted). However, the impasse concerning whether (i) middles constructions project to syntax a two-place transitive structure, as do transitive constructions, or whether (ii) they project a one-place intransitive argument structure, as do some intransitives, still persists.

In the present thesis, we present three experiments designed to probe for possible confounds influencing the results from previous studies, and to obtain more detailed information concerning the psychological reality of middles constructions. In Experiment

1, by contrasting the online processing patterns of transitive and middle constructions, we attempt to partially replicate the findings from Di Sciullo et al.'s (2007) self-paced reading moving-window task. Their results showed that at the post-verbal position middle constructions exhibit greater processing cost than their transitive counterparts. Such results, however, are liable to be confounded by the high rejection rates attributed to middle sentences in the two acceptability tasks of that study. In an attempt to control for such possible confounds, we added to our materials a supportive context for both transitive and middle constructions. If experimental conditions are equally well accepted, then greater reading times for middles could be interpreted as evidence that middle constructions are more complex to process than transitive constructions are.

In Experiment 2, we tap into the nature of the information offered by off-line judgment tasks. Employing an off-line questionnaire (Gruber & Gibson, 2004), we assess the extent to which an off-line measure can be used to scrutinize different aspects influencing sentence processing—such as semantics, syntactic complexity, and plausibility. We investigate transitive and middle counterparts, as well as other related complex structures, such as passives and adjectival predicative constructions. The latter constructions are argued to be derived syntactically via argument movement, and thus exhibit greater syntactic complexity than transitive constructions (Alexiadou, & Doron, 2012; Bowers, 2002).

In Experiment 3, we employ an eye-tracking technique, which offers a more comprehensive array of data as captured not only by measures of initial parse (e.g., first pass time) but also by revision measures (e.g., re-reading duration). Because participants are allowed to return to previously read regions of the sentences, we can assess which

regions of the sentences trigger revision more frequently. This allows us to pinpoint the onset of structural revision. We can also assess which regions require revision the most. This might reveal which structural nodes require revision. In Experiment 1, mirroring Di Sciullo et al.'s (2007) materials, the critical post-verbal region occupies the end-of-sentence position. In Experiment 3, materials display a prepositional phrase (PP) at the end of the sentences, thus eliminating a wrap-up effect as a possible confound affecting the measures collected for the critical post verbal region.

In the following section, we will first review the core linguistic proposals regarding middle formation, as the linguistic debate offers the theoretical background motivating our investigation. We will then proceed to review the evidence in the psycholinguistics literature concerned with the processing of middle constructions. All three experiments and final discussions are presented subsequently.

Linguistic Analyses of Middles

Within the linguistic literature we find support for both syntactic and pre-syntactic formation of middle constructions. Bowers (2002), Hoekestra and Roberts (1993), Keyser and Roeper (1984), and Stroik (1992, 1995, 1999, 2006), among others, offer varied accounts as to how a two-place transitive verb would be derived syntactically into a middle construction. That is to say, they propose that the external argument of the transitive verb is mapped onto syntax, but due to syntactic operations, it fails from being realized. In contrast, Ackema and Schoorlemmer (1994, 1995), Fagan (1988, 1992), Levin (1993), and Zribi-Hertz (1993), among others, propose middles to derive pre-syntactically, that is, middle constructions would reach syntax with only the complement. In fact, up to date there is still little consensus in the linguistic literature regarding how

middle constructions are ultimately formed, be it in terms of verb conceptual representation, or be it in the terms of syntactic and pre-syntactic derivations. More critically, on what structures ought to constitute a middle construction. (for a review, see Alexiadou & Doron, 2012; Ackema & Schoorlemmer, 2003; Lekakou, 2005; Marelj, 2004; Steinbach, 2002; Wenzhong, 2005).

Properties of middles. Levin (1993) defines middle constructions as (i) having an understood, yet unexpressed agent, (ii) often including an adverbial modal element, and (iii) lacking specific time reference. Levin (1993) proposes that alternations in a verb canonical transitivity, diathesis alternations, are associated with an alteration in the meaning the verb. Under this lexicalist approach, verbs that display diathesis alternations are in fact polyssemic verbs. The verb clean, for instance, in a transitive construction has such a meaning that it would require two arguments to be grammatical (x cleans y). The other meaning of the verb clean—such as the one appearing in middle constructions—requires only one argument to be grammatical (y cleans). Thus, all possible realizations of clean are to be listed in the lexicon. According to this approach, verbs undergoing middle construction are those transitive verbs whose meaning is associated with notions of motion, contact, and change of state, such as the verb paint. Note, however, that although some verbs such as read, or even sell, form the perfectly grammatical middle constructions shown in (2), these verbs do not conform with Levin's classification of verbs undergoing middle formation, since that they do not express notions of motion, contact, or change of state.

(2) a. This paper reads easily.

b. This book sells well.

In addition, some verbs such as *break,* whose meaning is also associated with notions of motion, contact, and of change of state, fall under the causative-ergative alternation, appearing in a transitive construction, as shown in (3a), and in an intransitive ergative construction, as shown in (3b).

(3) a. John broke the vase.

b. The vase broke.

Levin (1993) characterizes the ergative alternant as (i) neither requiring an understood agent, (ii) nor requiring a final adverb, but also as (iii) subject to display specific time reference. The latter definitional feature is also proposed by Keyser and Roeper (1984, p. 384), as it can be demonstrated by the comparison between the middles in (4 a, b) and the ergatives in (5 a, b):

(4) a. *Bureaucrats bribe abruptly.

b. *This wall paints abruptly.

(5) a. The vase broke abruptly

b. The door opened abruptly.

However, regardless of these definitional differences, verbs that undergo ergative alternation can also undergo middle alternation. The verb *melt,* for instance, can be realized in a transitive construction, as shown in (6a), in an ergative construction, as shown in (6b), and also in a middle construction, as shown in (6c) (Keyser & Roeper, 1984, p. 381).

(6) a. The sun melted the ice.

b. The ice melted.

c. The ice melted easily.

This is so despite the fact that not all middle verbs can undergo ergative formation (Keyser & Roeper, 1984; Levin, 1993). This would be the case of the verb *paint,* which render grammatical transitive and middle constructions, as shown in (1c, a) respectively, but which would render ungrammatical an ergative construction, as shown by the sentence **The wall painted.*

Notwithstanding these definitional difficulties, attempts to describe middle formation have been developed. As we detail below, middle formation can be devised in terms of a voice derivation. Under this view, middle constructions, like passives, as shown in (1b), are derived syntactically from a transitive representation. Middle formation is also approached under the comparison established between the middle and the intransitive ergative constructions.

Middles, passives, and ergatives.

Passives. Based on the minimalist approach (Chomsky, 1995), Bowers (2002) argues that even though some constructions with transitive verbs may not realize both arguments as a lexicalized NP, as is the case for middles and some passives, these constructions still possess a transitive aspect about them. Such phenomenon can be explained by the fact that transitivity itself is a syntactic category, TrP (transitivity phrase). TrP would only be present in the syntactic representation of transitive verbs (e.g., *kick,* as in the sentence *John kicked the ball*), but not in that of intransitive verbs—that is, ergatives (e.g., *fell* as in the sentence *John fell*), and unergatives (e.g., *cry* as in in the sentence *John cried*). In (7) we show the proposed syntactic projections for transitives and intransitives:

(7) a. transitive:$[_{PrP}$ DP $[_{Pr'}$ Pr $[_{TrP}$ Tr $[_{VP}$ V DP]]]]

b. ergative:[$_{PrP}$ Pr [$_{VP}$ V DP]]

c. unergative:[$_{PrP}$ DP [$_{Pr'}$ Pr [$_{VP}$ V (PP)]]]

The syntactic derivation of middle and passive constructions which only spell out their internal argument in subject position, respectively as in (8a, c), is explained by the presence of a voice morpheme instantiated at Tr. Bower (2002) poses that in languages such as English, this voice morpheme (realized as –EN for passives, and as μ for middles) would disable the θ-features of Tr's probe. Consequently, the NP complement would cyclically move upwards the structure, until finally merging at Spec position. The representation of English middle and passive derivations is shown respectively in (8b) and (8d). Ultimately, these voices would differ in that -EN requires an auxiliary verb, *was*, whereas μ is to be phonetically null.

(8) a. Bureaucrats bribe easily.

b. [$_{TP}$ T [$_{PrP}$ [Pr' [$_{TrP}$ [$_{Tr}$ μ] [$_{VP}$ easily [$_{V'}$ bribe bureaucrats]]]]]]

c. This bureaucrat was bribed.

d. [$_{TP}$ T [$_{PrP}$ [Pr' [Pr be][$_{TrP}$ [$_{Tr}$ –EN] [$_{VP}$ bribe this bureaucrat]]]]]

Under this approach, at the same time that what differentiates middles from passives is the nature of the voice morpheme, what differentiates middles form ergatives is transitivity itself. It would, thus, be Tr's semantic properties which are to account for the transitive feel present in middle constructions, but absent in the intransitive ergative constructions.

Bowers (2002) conceptualize transitivity as "an independent syntactic category in its own right, with interpretive properties that are independent of the presence or absence of an external argument" (2002, p. 216). Moreover, even though we understand that

within this program θ and case are grammatical features, and as such should be present at different levels of processing, it is unclear in Bower's work whether he also assumes derivations outside syntax. Attending to the fact that some ergatives can appear in middle constructions (Keyser & Roeper, 1984, also see [6c]), Bowers (2002) maintains that middles are derived syntactically whereas ergatives are not, and proposes that the surface structure of such constructions is indeed ambiguous, as exemplified in (9a). It would be Tr's semantic properties that would confer the middle interpretation of (9a): it is easy for someone to gallop the horse. The middle version would be syntactically represented as in (9b). The ergative interpretation of the ambiguous (8a), would express the horse's ability of a good gallop, and would have the syntactic representation shown in (9c).

(9) a. The horse gallops well.

b. middle: [$_{PrP}$ Pr [$_{TrP}$ [$_{Tr}$ μ] [$_{VP}$ well [$_{V'}$ gallop the horse]]]].

c. ergative: [$_{PrP}$ Pr [$_{VP}$ well [$_{V'}$ gallop the horse]]]

Alexiadou and Doron (2012) offer a similar approach to voice derivation, and utilize Doron's (2003) notion that voice can be articulated in terms of syntactic functional heads: the voice head π derivates passives, and the voice head μ derivates middles. They propose that because voice heads are elements of syntax —being thus independent of the verbal root— these heads can constraint argument realization in the sense that they may block merge. Differently from Bowers (2002), Alexiadou and Doron (2012) sustain that argument structure is determine by the verb root's semantics, and propose that not only middles, but also ergative (e.g. *The ice melted*) and reflexive (e.g. *John washed himself*) constructions would be derived syntactically the same way. Middles are derived syntactically the same way as ergative, and reflexive constructions are. Note that both

ergative and reflexive constructions display a canonically transitive verb with a theme in subject position[2], and both constructions have been proposed to be unaccusative (respectively: Burzio, 1986; Friedmann, Taranto, Shapiro, & Swinney, 2008; Grimshaw, 1990; and Pesetsky, 1995; but Reinhart & Siloni, 2004, 2005).

Note that some linguists, such as Correa (2007), Kemmer (1993), Reinhart and Reuland (1993), and Steinbach (2002), do assume that generic constructions containing a referring clitic, such as *Paul shaves (himself)*, in English, *Paulo se barbeia* (literally, *Paulo self shave*), in Brazilian Portuguese, and the German equivalent *Herr Rossi rasiert sich* are middle constructions. Even though we acknowledge the plurality of definitions of what would constitute a middle construction (cf. Marelj, 2004), our focus is on structures of the type NP-V-ADV, such as *This book sells well*.

According to Alexiadou and Doron (2012), the different possible cross-linguistic manifestations of the middle voice (Kemmer 1993; Keyser & Roeper, 1984; Reinhart & Reuland, 1993; Steinbach, 2002) would be explained by the interaction between the voice head μ and a given verb's semantic content. It follows that the absence of an external argument or its manifestation as referring clitic (e.g., *itself*) (Rizzi, 1986) would depend on the verb root semantic requirements.

Ergatives. Concerning the formation of ergative constructions, the unaccusative hypothesis (Burzio, 1986; Perlmutter, 1978) posits that the ergatives' only argument is born in complement position. That NP complement would then raise in the syntax to

[2] Maia et al. (submitted), for instance, argues that reflexive constructions such as (i), are identical in meaning to ergatives in Xavante, language in which the reflexive morpheme *tsi-* is infixed to the verb:

(i) "ridawa ma tsi–tsitowa
Porta 3a/pass refl.–abrir
"A porta se–abriu" (The door self–opened)

Spec position because ergatives fail to assign accusative case. Hence the label unaccusatives. Under the government and binding framework (Chomsky, 1981), the unaccusative hypothesis conjugates two principles in order to explain the syntactic derivation of ergatives: (i) Burzio's generalization: "All and only the verbs that can assign a θ-role to the subject can assign accusative Case to an object. [subject = external subject (agent)]" (Burzio 1986:178); (ii) the filter case rule (Chomsky, 1981), which imposes that no argument is to be left without case. Because ergative verbs can't assign θ-role of agent—they only select a theme born in complement position—, these verbs are then blocked from assigning accusative to its internal argument, according to (i). In order to satisfy (ii), the internal argument has to move to subject position to get case: nominative case, which is offered at specifier position.

The unaccusative hypothesis was developed based on the assumption of a close relationship between θ-role and case assignment. Even though it remains influential (Friedmann et al., 2008), Burzio's generalization is currently being questioned and novel theoretical approaches regarding how arguments get case and θ-role are being proposed (see Bowers, 2002, Laka, 2000, and Woolford, 2003). Also, see Chierchia (2004), and Embick (2004), and Alboiu, Barrie and Frigeni (2004) for a debate on unaccusativity.

Contrasting middle and ergative constructions, Rapoport (1999) proposes that both middles and ergatives are unaccusative verbs, which select only a complement. Attending to middles' agentive character—which would pose the distinction between middles and ergatives— Rapoport proposes that such characteristic of middles does not derive from the existence of an external argument. What makes middle constructions implying a "protoagent" is the entailment derived from the instrument or manner (IM)

component associated with the meaning of the action denoted by the verb.[3] If the action denoted by the verb involves an instrument, it consequently also involves the notion of an agent, which manipulates such instrument. This way, ergatives would be grammatical with simple change of state verbs, as exemplified by the verb *brake* in (10a), and middles would be grammatical with change of state verbs with the IM component, as exemplified in (10b) by the verb *cut*, whose action entails the need of an instrument, such as a knife.

(10) a. The glass broke.

b. *The bread cut.

A piece of evidence for this argument comes from the observation that the addition of the expression *all by itself* renders the middle in (11a) ungrammatical, whereas the ergative in (11b) remains grammatical. The expression *all by itself* corefers to theme internal argument, *the bread* in (11a), and *the glass* in (11b). Therefore it would be disallowed in the middle construction because the IM component of middle verbs would require it to be an agent performing the entailed action.

(11) a. *This kind of bread cuts easily all by itself.

b. This glass breaks easily all by itself.

However, Rapoport's proposal seems to fail to accommodate for middle verbs such as *read* and *sell*, which, at least prima facie, do not entail a change of state.

It is worth noting that middle sentences such as (10b) and **The bread cuts* could be considered grammatical. This could entail (i) that middles and ergatives share the same underlying configuration, i.e. argument structure, and (ii) that the differences

[3] "The agentivity aspect of some middles, then, is due to the presence of the I/M component in the verb base. Thus, it is a characteristic of particular verbs and not a property of the middle construction itself …." Rapoport, 1999, p.151).

between these two constructions lie outside the realm of grammar, and are derived from the lexical root itself. As Di Sciullo (2005) posits, a shift in the verb argument structure *may* entail changes in the verb's content. We present this argument in more details in the following section. In fact, proponents of a pre-syntactic derivation of middles suggest that the putative requirement for a modifier derives from reasons detached from grammaticality conditions. Ackema and Schoorlemmer (1994), Fagan (1988, 1992), and Zribi-Hertz (1993), for instance, agree that the adverb, an adjunct in nature, serves to pragmatically anchor middle sentences which could be odd, some in a greater degree than others, when taken in isolation in a utterance, e.g., *This dress buttons*. (Fagan, 1988), *This bureaucrat bribes.* (Ackema & Schoorlemmer, 1994). Given the appropriate or supportive context, the verb's reduced argument structure should suffice. Concerning Rapoport's proposal, if the root's I/M content is not altered in the middle alternant, then, in the absence of contextual support, adverbs may provide pragmatic anchoring to such middle constructions.

Hale and Keyser (2002) also propose that middle constructions are represented in the syntax with an unaccusative structure of the type [$_{\text{head}}$ Head [$_{\text{Comp}}$ [Head Comp]]], which configures a lexical head taking a complement, but no specifier. Following theory-internal mechanisms, accusative case binding would be cancelled in such configuration, and, consequently, the complement NP would raise to Spec position. According to this approach, both middle and ergative verbs would display a one-place unaccusative argument structure, thus extending the principled elements of the unaccusative hypothesis to middle formation. Such appreciation of middles is indeed coherent with Rapoport's proposal that middles are unaccusatives.

As it is characteristic in linguistics, these approaches to middle formation are not only diverse, but they are also not primarily concerned as to the how their proposals are to be implemented and operationalized in real time during language use. In relation to sentence comprehension, it becomes particularly relevant to understand, for instance, how syntax is informed about the necessity to build a TrP node for the incoming linguistic stream (Bowers, 2002). Nevertheless, some of these proposals indeed find empirical evidence. This is the case, for instance, of the unaccusative hypothesis, which was investigated by Friedmann et al. (2008).

Under the principles and parameters framework (Chomsky, 1981), moved arguments are believed to leave a *trace* (*t*) in the position where they were generated. The investigation of the psychological reality of *traces* has often relied on the cross-modal lexical priming technique (Swinney, Onifer, Prathe, & Hirshkowitz, 1979). The technique relies on the assumption that traces can be reactivated, and that, consequently lexical judgments tasks—word or non-word— at the trace position would be facilitated once the visually presented target word is semantically related to the moved NP in the aurally presented sentences. Friedmann et al. (2008) employed this paradigm to investigate the psychological reality of the unaccusative hypothesis. They contrasted sentences containing (non-)alternating ergatives (e.g., *The tailor disappeared*, and *The table dried*) and unergatives (e.g., *John died*) verbs. Even though ergatives and unergatives have the same surface structure (NP-V), only unergatives have their argument already generated in Spec position, and are thus not derived via syntactic movement. Friedmann and colleagues employed sentences such as those in (12).

(12) (a) Non-alternating unaccusative: The **tailor** (1) from East Orange, New Jersey,

mysteriously **disappeared** (2) when it was (3) time to adjust the tuxedos and dresses for the participants in the wedding party.

(b) Alternating unaccusative: The **table** (1) in the basement of the old house finally **dried** (2) after the leaking (3) window was sealed a month ago.

(c) Unergative: The **surgeon** (1) with a brown felt fedora hat and matching coat eagerly **smiled** (2) when the beautiful (3) actress walked down the corridor to exam room three.

Friedmann et al. (2008) collected data from three probes which were either semantically related or not semantically related to the NP in subject position: probe (1) was displayed immediately after the NP subject, probe (2) was displayed immediately after the verb, and probe (3) was displayed at 750 ms. after probe (2). Because priming effects are better perceived after decay of the target, probes (1) and (2) served as baseline to test priming, and probe (3) was the critical one to uncover priming effects associated with NP trace reactivation. Results showed no reactivation at probe (3) for the unergative condition, thus supporting the idea that the argument of unergatives is not base generated. Concerning the non-alternating ergatives, results showed a priming effect at probe (3), which is consistent with the reactivation of the *trace* left by the moved internal argument, thus corroborating the predictions generated by the unaccusative hypothesis.

Middles and flexible argument structures.

The observation that alternations in the expression of a verb argument structure can be accompanied by changes in meaning (Levin, 1993; Levin & Rappaport-Hovav, 2005) receives a different interpretation under antisymmetry, or asymmetry, theories (Kayne, 1994; Di Sciullo, 2005). Under this perspective, linear relations have direct

implications to structure hierarchy, whereby different linear order corresponds to different hierarchical relations amongst the elements of a given phrase. Hence, the putable changes in verb meaning associated with verb alternations would be understood as a derivative of the changes in relations between the verb and its arguments, and would be reflected on and by the changes in linear order phrases. Because asymmetrical relations are to be present in different domains, such as syntax and morphology (Di Sciullo, 2005), the arising conclusion is that asymmetries are in the grammar, that is, they "are core relations of the language faculty." (Di Sciullo, 2005, p.2).

Hale and Keyser (2002), for instance, propose argument structures to be the syntactic representation of lexical items, and that, as elements of the grammar, argument structures should be independent of specific lexical categories, such as N (noun), V, and P (preposition). Under their view, even though argument structure types are to be universal, some canonicity is bound to take effect in specific systems, that is: specific types of argument structures tend to occur recursively with specific lexical items in specific languages. In English, for instance, the atomic (d)-type [Head] tends to appear with N; and the monadic (a)-type [$_{\text{head}}$ Head [$_{\text{Comp}}$ [Head Comp]]], which configures a lexical head taking a complement, but no specifier, tends to appear with V (Hale & Keyser, 2002, p. 13).

Di Sciullo (2005) operationalizes a similar approach to that of Hale and Keyser's (2002) treatment for argument structures and proposes the existence of three basic argument-structures under which a morphological head can be realized: the unaccusative [$_{x}$ α [$_{x}$ [x [+A]]]], the unergative [$_{x}$ [+A] [$_{x}$ [x β]]], and the transitive [$_{x}$ [+A] [$_{x}$ [x [+A]]]]. Note that [+A] stands for Argument feature (A-feature), and is the interpretable

argument feature, probe. According to Di Sciullo, since asymmetric relations are part of the grammar, these same relations are to be available for other lexical heads, such as V, and are also to be mirrored in other domains, such as that of syntax. Still assuming a modular language system, the syntax domain, more specifically its computational space—as it is conceptualized by Di Sciullo—, is taken to stand in parallel with other computational spaces, such as those pertaining to the morphological, phonological, and lexical domains. The interface between these computational spaces is proposed to be limited, as it would also be limited their interface with language external systems[4]. Attached to the construct of parallel narrower computational spaces, is the notion of flexible argument-structures, which are, in turn, proposed to be limited in number. Flexible argument-structures are conceptualized as part of the grammar and are, consequently, available to all processing spaces.

It is important to differentiate this notion of argument-structures from lexical-conceptual templates (Jackendoff, 1989, 1990; McKoon & Macfarland, 2002; even Pustejovsky, 1995). As Di Sciullo (2005) notes, also in similarity to Hale and Keyser's (2002) model, thematic relations are not primitives associated with lexical items, they are features attached to and derived from configurations, that is, the argument-structure in which those items appear. (Di Sciullo, 2005, p. 194-5). Hence, θ- and case- assignment relations are dissociated from conceptual representation, as these relations are then proposed to be affixed to a given argument-structure.

[4] Modularity of Computational Space Hypothesis: "The computational space includes interactive types of derivations leading to target types of configurations." (Di Sciullo, 1996c, p. 5 [cf: Di Sciullo, 2005, p. 182])

If phrase meaning and phrase order are to be linked, than the most appropriate argument-structure, even if non-optimal, is to be selected for the appropriate context or intention, in comprehension and in production. It is such flexibility in the attributed argument-structure configuration which, in order to accommodate for non-canonical meaning of phrases, would enable an argument shift (A-shift). However, different from Hale and Keyser (2002), Di Sciullo's (2005) proposal seems to entail no particular canonical, or preferred, association between a specific argument-structure and an specific verb. In relation to predicates, it would be A-shift the mechanism allowing and accounting for natural languages variability in verbs' behaviour, as well as for any possible associated changes in the verb's contents (such as aspect, genericity and compositionality), as it would be the case of diathesis alterations, such as the ergative alternation *John sank the boat* /*The boat sunk*. (Di Sciullo, 2005, p.59). Ultimately, under this framework, it seems to be the case that middle alternation would be explained by the appropriation of an unaccusative argument-structure.

Notwithstanding theory internal distinctions between Di Sciullo's (2005), and Hale and Keyser's (2002) approach, what becomes relevant to a psycholinguistic point of view, what ultimately can be experimentally tested (syntactic movement), is the notion that middles would be syntactically derived by having its internal argument raised to the sentential subject position. In fact, Di Sciullo et al. (2007) interpret their results, which suggest that middles are processed differently form passive and from transitive active constructions, as evidence for A-shift. We discuss next the psycholinguistic evidence on the processing of middle constructions.

Comprehension of Middle Constructions

There are only two psycholinguistic studies, that we know of, which investigate the middle construction: Di Sciullo et al. (2007), and Maia et al. (submitted).

Di Sciullo et al. (2007) investigated whether middle constructions are a product of A-shift. They compared English middle sentences against non-movement (Experiment1-2) and movement (Experiment 3) exhibiting structures: respectively transitive active constructions, as shown in (13 a, b), and passive constructions, as shown in (13 d, e). Transitive active (henceforth transitive) and passive sentences were construed under two conditions in order to control for any potential impact of subject NP animacy on the processing of the experimental sentences: the Animate condition, displaying an animate NP in the subject position, as shown in (13 a, d); and the Inanimate condition, displaying an inanimate NP in the subject position, as shown in (13 b, e). Because middle sentences were devised as NP-V-ADV, in order to maintain a surface parallelism across conditions, both transitive and passive sentences had an end-of-sentence adverb, which replaced the object NP of the transitive conditions, and which replaced the passive agentive *by*-phrase. Verbs and adverbs were kept constant across conditions. See (13) below:

(13) a. Transitive Animate: *The clerk sells steadily*.

b. Transitive Inanimate: *The store sells steadily.*

c. Middle: *The book sells steadily*.

d. Passive Animate: *The clerk was fired quickly.*

e. Passive Inanimate: *The store was sold quickly.*

f. Middle Passive: *The book was sold quickly.*

Experiments 1 and 2 consisted of an off-line acceptability judgment task, in which participants had to answer *yes* or *no* as to whether transitive and middle sentences were acceptable in English. In Experiment 1, materials were presented visually. In Experiment 2, materials were presented aurally. Response times revealed participants took longer to judge sentences in the Middle than in the Transitive Animate condition (Experiment 1, p =.08; Experiment 2, p =.01). With both experiments taken together, rejection rates were grater for middles (around 35%) against each transitive condition (below 15 %). Regarding animacy, no difference was found between the transitive conditions. Results were interpreted as possibly indicating grater processing load on the comprehension of middle constructions. No power or effect size measures were reported.

Experiment 3 consisted of a self-paced reading moving-window paradigm, in which the passive conditions, (13 d-f), were added to the previous materials. In their Experiment 3, the maintenance of the same verb and adverbs across all the conditions would have rendered some sentences in the passive conditions semantically anomalous, (e.g. *The clerk was sold quickly*). Thus, some sentences in the passive conditions were corrected for such anomalies, by displaying different verbs and adverbs, as can be seen in the examples provided in (13).

The reported analyzed regions were that of the verb, and that of the adverb. The only significant results were found at the end-of-sentence position, where the adverb in the Middle condition (i) was read faster than in the Transitive Inanimate condition, but, in turn, (ii) it was read slower in the Middle Passive condition. Di Sciullo et al. (2007) interpreted these results (i) as possibly reflecting a difference between A-shift and

argument movement, and (ii) as possibly reflecting the difference between the processing cost derived from A-shit against the cost of processing non-movement sentences.

However, the only statistically significant differences, which were found at the adverb final position, might be unreliable. First, reading times could be encompassing an end-of-sentence wrap-up effect. Second, reading times at that region could also be reflecting the processing costs derived from a structural revision due to the non-fulfilled expectation of encountering a post-verbal NP object in the transitive conditions, or a post-verbal by-phrase in the passive conditions. Third, lexical items were kept constant across conditions. Ultimately, even though we cannot explain the directionally of the results obtained in their Experiment 3, it is possible to speculate over a lack of power affecting the analyses, especially considering the small sample size reported in their preliminary analyses (N= 20).

Furthermore, Di Sciullo et al. (2007) discarded NP animacy as a factor influencing the results. Concerning the off-line measures, in which the differences between middle and transitive animate conditions were found, results could be interpreted as reflecting a canonicity effect associated with the typicality of the NP found in subject position—which is typically: *agent*, *instrument*, and *theme* respectively for transitive animate, transitive inanimate, and the middle conditions (e.g., Baker, 1987). Regarding the differences found in the online measures between the middle and the transitive inanimate conditions, difference which would corroborate the authors' claim, no definite assertions can made due to (i) the problems associated with the adverb final position, and (ii) the fact there are no data for the NP subject position.

Maia et al. (submitted) presented a larger study of diathesis alternations, replicating Di Sciullo et al.'s Experiment 1. They investigated the middle alternation in Brazilian Portuguese (BP), as shown in (14); and the ergative alternation in other two Brazilian indigenous languages, Xavante and Karajá. Xavante is a language which, unlike English, Portuguese, and Karajá, has verb transitivity morphologically marked in the verb. Materials were constructed based on the parameters used in Di Sciullo et al. (2007)'s Experiments 1 and 2.

(14) a. Transitive Animate: Esse homem vende bem. (*This man sells well.*)

b. Transitive Inanimate: Essa loja vende bem. (*This store sells well.*)

c. Middle: Esse livro vende bem. (*This book sells well*.)

The comparison between response times in the BP middle and transitive animate conditions was marginally significant (p =.06), thus reflecting a tendency of participants taking longer to reject Middles than to reject the transitive animate sentences. Rejection rates were found significantly higher for the middle condition (around 40%), compared against each transitive condition (below 20%). Also in Xavante, results revealed participants took longer to reject the Ergative condition than they took to reject the transitive animate condition. Results also revealed greater rejection rates for the Ergative condition against each of the transitive conditions. On the other hand, in Karajá, neither response times nor rejection rates were found to differ across conditions. No differences were found between the transitive conditions in neither language.

Maia et al. (submitted), operating unaccusativity under the distributed morphology framework (Marantz, 1997; Harley, 2006), argues for evidence of syntactic movement in the formation of middle and ergative constructions. They explain these

alternations in terms of the syntactic realization of the lexical features: (i) in the transitive structure, the "change of state" verb would project CAUSE in the little v functional head; (ii) in the intransitive structure, the "stative" verb would project BECOME in the little v functional head. Whilst vCAUSE, may select an external argument, as it would be the case for the transitive counterparts, vBECOME does not select an external argument (Hale & Keyser, 1993). As a consequence, the intransitive counterparts would have an unaccusative syntactic derivation in which the internal NP would move to Spec position in order to get case. Whereas morphologically opaque in BP, and in Xavante, the little vCAUSE would be expressed as the morpheme *–i–* , and little vBECOME would be expressed as the morpheme *–a–* in Karajá's transitive and ergative constructions, respectively. Maia et al. (submitted) pose that the facilitatory effect on the comprehension of ergatives in Karajá would be explained by the computation of morphological information, that is, the morphologically transparent transitivity morpheme in Karajá. Refer to Di Sciullo (1993, 1995, 1997,) for a more detailed discussion on the morphology associated with the realization of different verb types, such as those verbs undergoing the causative-ergative alternation.

However it would seem that Maia et al.'s (submitted) account fails to explain the alternation of non-causative verbs, such as *read, sell,* and *trade,* which undergo middle formation. Besides, the restrictions imposed by the nature of the measures collected (i.e. response latencies), do not allow for direct substantiation of syntactic movement. This remark can also be extended to Di Sciullo et al.'s (2007) Experiments 1 and 2. These factors in part weaken the validity of Maia et al's conclusions. What cannot be ignored, however, is the evidence from Karajá suggesting that morphological information

facilitates the acceptability of non-canonical unaccusative constructions. Moreover, we maintain that the significant differences between Transitive Animate and Middle conditions can be entertained as reflecting typicality effects of the NP occupying subject position.

In summary, results from four different languages converge to demonstrate that, without contextual cues, unaccusative constructions, be it middle or ergative, are less acceptable than their transitive counterparts.

Our goal in the present study is to correct for the higher rejection rates associated with the experimental middle constructions, and clear online data from the possible confounds of these previous studies. In Experiment 1, noting that morphological cues are not an option for English, we investigated whether (i) the appropriate sentential context would improve middles acceptability, and whether (ii) findings from Di Sciullo's (2007) Experiment 3 would be replicated once acceptability is improved. In Experiment 2, we implement a comprehensive off-line protocol (Gruber & Gibson, 2004), aimed to specifically probe for derivational complexity effects. In Experiment 3, we utilize a more sensitive technique—eye-tracking— but also improved materials, in order to monitor for revision patterns in search for evidence of structural complexity associated with the processing of middle constructions. In Experiments 1 and 3, we eliminate the passive conditions, and incorporate the analysis of the pre-verbal position as a way to substantiate our claim regarding typicality effects of the noun in subject position.

Experiment 1 – Self-paced Reading

Normative Study

Method.

Participants. Thirty undergraduate and graduate students from the Concordia University community participated in this experiment. There were 15 females and 15 males, with age ranging from 21 to 35 (*M*=26.97; *SD=3.90*). They were all native speakers of Canadian English, except for one native of the US variant. Participants were asked to fill out a questionnaire aimed to establish fluency, and dominance of their English language. Seven of them were English monolinguals, and 23 spoke one or more languages other than English. From the non-monolinguals, 14 out of 23 were early bilinguals, having acquired both their languages before the age of five. The others were late learners, having learned their other language(s) after the age of six. All participants reported have been exposed to English the majority of their lives (*M* = 26.73 years; *SD* = *3.78*).

Materials and procedure. Experimental materials were adapted from those used in Di Sciullo et al. (2007). We used the same 15 verbs to construct 45 sentences with structure NP-V-AdvP, 15 in each of the three conditions: Transitive Animate, as shown in (14a); Transitive Inanimate, as shown in (14b); and Middles, as shown in (14c). An adjectival predicative clause, with structure NP-*be*-AdjP (adjectival phrase), preceded each experimental clause, in an attempt to reduce high rejection rates previously found for middle constructions (Di Sciullo et al., 2007; Maia et al., submitted). First and second conjuncts were posed in a contrastive coordination linked by the conjunction *but*. The resulting structure was: NP1-be-AdjP, *but* NP2-VP-AdvP (adverbial phrase). Constant

across conditions were the verb and their associated adverb. The nouns in NP1 and NP2 were semantically related, but varied according to sentence type: the Transitive Animate condition presented animate *agent* NPs (*cook* and *chef*, in [15a]), the Transitive Inanimate condition presented inanimate *instrument* NPs[5] (*blade* and *knife*, in [15b]), and the Middle condition presented inanimate *theme* NPs (*croissant* and *bread*, in [15c]). The full set of materials is provided in Appendix A.

(15) a. Transitive Animate: *That cook is messy, but this chef cuts neatly.*

b. Transitive Inanimate: *That blade is dull, but this knife cuts neatly*.

c. Middle: *That croissant is flaky, but this bread cuts neatly*.

Experimental sentences were composed of 8 to 10 words (M = 9.09, SD = 0.36). Word length varied from 4 to 10 letters (M = 5.87; SD =1.60) for the verb, from 4 to 9 letters (M = 6.80; SD =1.37) for the adverb, and from 4 to 9 letters (M = 5.58; SD =1.44) for the NP2. NP2 length did not vary across conditions ($F2$ (2, 28) = 1.09, $p = .350$, $\eta_p^2 = .07$, $\varepsilon = .22$, t-tests were all n.s.). As expressed by frequency of occurrence[6], word familiarity displayed $M_{\text{Log10}} = 2.13$ (SD = 0.58) for verbs, $M_{\text{Log10}} = 2.66$ (SD = 1.02) for adverbs, and $M_{\text{Log10}} = 2.82$ (SD = 0.65) for nouns[7]. NP2 familiarity did not vary across conditions ($F2$(2, 28) = 2.82, $p = .076$, $\eta_p^2 = .17$, $\varepsilon = .51$)

In addition to the 45 experimental sentences, distractor sentences were comprised of (i) 32 syntactically and semantically good sentences (e.g., *The receptionist greeted the lawyer enthusiastically*); (ii) 31 syntactically and or semantically anomalous sentences

[5] The instrument character of the NP2 associated with the verbs *sell* and *read* is, however, unclear, i.e. *the store sells*, and *the firm trades* (Fagan, 1988, 1992; Rapoport, 1999).

[6] In this research, all word frequency counts were obtained from Subtlex(US) (Brysbeart & New, 2009).

[7] Collocation frequency counts for N–V, and V–Adv, had an averaged 30% empty return in the COCA database.

(e.g., *Charles met for the interpreter; The bishop baptized the dentist next summer);* and (iii) 32 also grammatical genitive sentences which were part of an unrelated study. The latter were structured as NP-V-NP-of the-NP, exhibiting the norman variant of the English genitive (e.g., *The chef broke the lid of the pan*).

Three booklet versions were created. Each participant was given one of the versions. Besides 15 experimental sentences, each booklet comprised 30 normal English sentences, 31 anomalous sentences, and 11 genitive sentences. Experimental materials were distributed following a Latin square design, i.e., one third (or the equivalent of five sentences) in each condition, never repeating verbs. Items order was different for each booklet version, as it was set pseudo randomly. Items were first arranged according to a Microsoft Excel *rand* function and later re-arranged so that no sentence of the same condition would immediately follow each other.

Booklets had an instructions page followed by 10 other pages, containing the test materials. Participants were instructed to rate how acceptable the sentences were in a scale from 1 (*not acceptable*) to 5 (*acceptable*) placed next to each sentence. An acceptable sentence, for the purpose of the experiment, was defined as being grammatical and plausible in a real world situation. Participants were further instructed not to dwell on any of the sentences, and to answer according to their first instinct.

Results – Normative Study.

Ratings for the 15 sentences in each condition were averaged according to subjects and items analyses. Means were all above the threshold of 4, with a slightly higher acceptance level for Transitive Inanimate ($M = 4.26$; $SD = 0.60$), followed by

Transitive Animate (M = 4.22; SD = 0.65), then by Middle (M = 4.13; SD = 0.63) (as per participants' analysis).

It is important to note that one sentence, initially placed under the middle condition, with the verb *cook*, pertained in reality to the active inanimate condition. Therefore, in all the analyses that verb was treated as such. Ratings for the *cook* sentence, initially in the MID condition, were integrated into those of the *cook* sentence in the active inanimate condition. Following Tabachnick and Fidell (2007), in relation to the missing value in the middle condition, in the items analysis, a Missing Value Analysis was run in SPSS 19, having all the variables inserted in the model. Little's MCAR test suggested that the value was missing at random, $\chi^2(2, 45) = 0.68$, $p = .711$. The resulting estimated mean (EM) values were then inputted into the middle condition, thus replacing the missing values for the *cook* sentence at all levels of the rating variable.

Analysis of variance. A 3 (sentence type: Transitive Animate, Transitive Inanimate, Middle) X 1 (ratings) repeated-measures ANOVA, and pairwise comparisons confirmed no statistically significant difference in ratings between conditions. Even though sentence type accounted for around 36% and 34% of the variance in the participants' and items' analyses, respectively, the models failed to reach statistical significance (respectively: $p = .344$, and $p = .614$). These results show that Middle sentences were rated as acceptable as their Transitive counterparts, and that sentences in all conditions were overall found to be acceptable by the participants.

Correlational analyses. In order to investigate whether word familiarity played a role in acceptability rates, multiple regression analyses were carried out. Word familiarity has been demonstrated to predict on line measures of lexical processing, such as lexical

recognition judgment tasks, and reading times (Brysbaert & New, 2009; and Brysbaert et al., 2011; Rayner, 2009). The purposes of the analysis are (i) to exclude familiarity as a confound affecting materials acceptability, and (ii) to investigate whether off-line sentence comprehension is related to word familiarity in general, or whether such relation, if any, is specifically associated with a given syntactic category, such as N, V(verb), or Adv (adverb).

Mean ratings for each of the 15 items in each condition (Transitive Animate, Transitive Inanimate, Middle) were inserted as the dependent variable in a three-steps hierarchical regression analysis, (i) having verb frequency as predictor in the first model, (ii) having verb and adverb frequencies as predictors in the second model, and (iii) having verb, adverb and noun[8] frequencies altogether as predictors in the third model. Results indicated that none of three models showed a statistically significant change in the proportion of variance accounted for in each condition's mean ratings. In fact, none of the models were strong nor statistically significant in accounting for the variance in any of the ratings variables (Models in ascending order for Middle: R^2_{adj} = -.07, $F(1, 13)$ = 0.11, p = .745; R^2_{adj} = -.16, $F(2, 12)$ = 0.06, p = . 940; R^2_{adj} = -.11, $F(3, 11)$ = 0.55, p = .676. Models in ascending order for Tr. Ina.: R^2_{adj} = -.07, $F(1, 13)$ = 0.07, p = .801; R^2_{adj} = -.16, $F(2, 12)$ = 0.06, p = .945; R^2_{adj} = -.22, $F(3, 11)$ = 0.17, p = .916. Models in ascending order for Tr. Ani.: R^2_{adj} = -.02, $F(1, 13)$ = 0.72, p = . 411; R^2_{adj} = -.09, $F(2, 12)$ = 0.40, p = .681; R^2_{adj} = -.18, $F(3, 11)$ = 0.29, p = .830). Results also indicated that neither verb, adverb or noun familiarity were indeed good predictors of acceptability ratings for any of the sentence types. See Table 1 for detailed results.

[8] Missing value for the NP in the *cook* sentence was replaced by the variable mean.

Frequency analysis. Because the values entered into the ANOVA analyses consisted of collapsed means, we proceeded to investigate whether the concentration of raw ratings across the scale was independent form sentence type. Accordingly, rates 1 through 5 were transformed into a categorical variable and all 450 counts were tabulated into a frequency table. A 3 (sentence type) x 5 (rate) Chi Square analysis revealed no statistically significant association between those variables, $\chi^2(8, 450) = 4.77$, $p = .782$, $\Phi^2 = .01$. In fact, percentages of rates 1 to 5 showed that around half of the totality of scores fell on the higher extreme of the scale (rate 5), and that the concentration of ratings 5, and 4 together accounted for more than 70% of the scores. See Table 2 for percentages in the participants' analysis.

Table 1

Hierarchical Multiple Regression Analyses Predicting Sentences Ratings from Word Frequency.

	Sentence Type					
	Middle		*Active inanimate*		*Active animate*	
Predictor	ΔR^2	β	ΔR^2	β	ΔR^2	β
Step 1	.01		.01		.05	
Verb frequency		.09		-.07		-.23
Step 2	.00		.00		.01	
Verb frequency		.08		-.18		-.26
Adverb frequency		-.05		.23		-.10
Step 3	.11		.03		.01	
Verb frequency		.05		-.22		-.29
Adverb frequency		-.11		.15		-.15
Noun frequency		.34		-.63		-.12

Note: All *p-values were nonsignificant.*

Table 2

Percentages of Rates across Sentence Type (by participants).

	Ratings				
Sentence Type	1[a]	2	3	4	5[b]
Middle	2.1	3.6	17.9	34.3	42.1
Transitive Inanimate	2.0	2.0	15.0	30. 7	50.3
Transitive Animate	2.7	2.0	17.3	26. 7	51.3

Note: [a] 1 = *not acceptable;* [b] *5 = not acceptable*

Discussion – Normative Study.

By embedding the target sentences in a larger coordinated structure, we expected to alleviate participant's rejection towards the sentences in the Middle conditions. Results from the normative study were consistent in showing that middles and transitive sentences were all judged as equally acceptable sentences. Hence, these results corroborate the hypothesis that contextual information would alleviate the higher rejection levels previously reported for middle sentences (Di Sciullo et al., 2007; Maia et al., submitted). Word familiarity, regardless of their syntactic category (V, Adv, and N) was found not to significantly predict sentence acceptability ratings, thus refuting word frequency of occurrence as a confound variable to be captured by such off-line tasks. Moreover, the investigation of the concentration of ratings across the acceptability scale allowed assessing with greater detail participants' behaviour on judging how acceptable the sentences in each condition were, as well as the internal validity of the range of the scale. For instance, polarized rating tendencies, that is 1 and 5, and conservative rating tendencies, that is, 3, would possibly reflect relevant information regarding participant's attitude towards the sentences being tested. One could make the case that high concentration of extreme values in the middle condition, as opposed to a more disperse distribution of ratings for the active conditions, could (i) reflect participants' inability to grasp semantic-pragmatic nuances derived from the middles syntactic configuration, (ii) indicate that participants had difficulties retrieving or formulating the non-canonical semantic relations of middle structures; or (ii) reveal the pattern that participants opted for checking 3, when in doubt. For reasons earlier described, such subtleties in ratings pattern were likely to be left under-detected in the ANOVA. Chi Square results suggested

participants implemented unchanged strategies while rating the sentences of all three types. This analysis also revealed that participants understood and made appropriate use of the full rating range, thus rendering these materials a valid instrument for measuring overall acceptability.

It is pertinent to note, nonetheless, that booklets acceptability ratings, which ranged from 1 (*not acceptable*) to 5 (*acceptable*), combine, by definition, judgments of plausibility and grammaticality. Therefore, they do not permit identifying whether participants based their responses solely on grammatically, solely on plausibility, or to which degree they balanced out each criterion while judging the sentences.

Gruber and Gibson (2004) developed a questionnaire methodology, which allows dissociating plausibility from other "goodness" ratings. We implement such methodology in Experiment 2. It is also important to note that we acknowledge the possibility that fillers may have influenced the results, even though fillers were counterbalanced (50 % good, grammatical sentences, and 50% semantically, and or syntactically anomalous sentences) with the very purpose of dissuading any possible influence of filler acceptability over experimental sentences ratings.

In summary, results from the normative study suggested that (i) sentences were overall rated high in the scale (all means were above the threshold of 4), (ii) neither sentence type or word frequency interfered with ratings rank or distribution, and that (iii) the scale of measurement was appropriate. Assured that sentences in each condition were well accepted, we proceed to the self-paced reading task.

Self-paced Reading Study

Method.

Participants. One hundred and fourteen undergraduate students from the Department of Psychology at Concordia University participated in the experiment in return for course credit. They were 99 females and 11 males, with ages ranging from 18 to 48 yeas (*M*=23.35; *SD=5.85*). Sixty-three of them had corrected-to-normal vision. They were all native speakers of Canadian English, having learned the language before the age of three. Twelve were English monolinguals, and 98 spoke one or more languages other than English.

Materials and apparatus.

Materials. Three lists of materials were drawn from the sentences used in the normative study. Each list contained 15 experimental sentences, and 30 fillers. Experimental sentences were the same ones used in the normative study, and were also distributed following a Latin square design, with one third of the verbs in each condition. Fillers were comprised of 20 short grammatical English sentences composed of 4 to 6 words, with varied structures such as NP-V-NP/AdjP (e.g., *The receptionist greeted the lawyer*, and *The clients hated the new logo)*, and 30 long sentences, 10 per list, composed of 12 to 13 words. Long sentences were part of an unrelated study and were structured as NP-V-NP-of the-NP-Relative Clause (RC) (e.g., *The chef broke the lid of the pan that was on the counter*).

All items, including practice ones, were followed by a comprehension question. The correct answer for the short filler questions was set to a ratio of 50% "yes", and 50% "no". For instance, for the sentences *Laura bought new winter boots,* and *The crazy man*

hit the wall, the questions were, respectively, *Did Laura buy new boots?,* and *Was the man sane?*. The questions following the long filler sentences were devised so that the RC was linked to the first NP in the genitive construction (e.g., for the sentence *The chef broke the lid of the pan that was on the counter,* the question was *Was the lid on the counter?).*

In the case of experimental sentences, to ensure participants would read the whole sentence, the questions were construed so as to draw information from the first and second conjuncts. Questions were devised in such a way that the modifier in the first clause was either (i) linked to the subject of the first clause, which would prompt an "yes" answer, or (ii) linked to the subject of the second clause, which would prompt a "no" answer (e.g., for the middle sentence *That croissant is flaky, but this bread cuts neatly,* the question *Is this bread flaky?* followed the type (ii) pattern). "Yes" and "no" answers were balanced within experimental conditions, so that each participant saw a 2 by 3 ratio of type (i)/type (ii) questions per condition. This means that, out of the five sentences in each condition per list, 2 had a "yes" for correct answer and 3 had "no" for correct answer.

Items order was set pseudo randomly, according to the same procedures used in the normative study.

Apparatus. Experimental sessions took place in a quiet room equipped with a 20" Viewsonic monitor, a CMU response box, and a pair of Bose noise cancelling headphones. All of which were connected to a Macintosh 3G computer, running on a Mac OS9 platform. Stimuli were presented at the center of the screen, single lined, on Courier font, using PsyScope version 9 software (Cohen, MacWhinney, Flat, & Provost,

1993). Due to the fact that each list was already pseudo-randomized, PsyScope was to set present trials according to a fixed order.

Procedure. The task consisted of a self-paced reading moving-window paradigm. Each trial started with a row of hash marks (#), with each mark corresponding to a letter. Each time the middle button on the response box was pressed, one word was revealed and the previous one turned back to rows of harsh marks. This way, participants read the sentences word by word. After pressing the middle button after the last word of the sentence had been revealed, a comprehension question would appear at the center of the new screen. To provide their answer, participants would press the red (left) button for "no" or, the green (right) button for "yes". After the response, another sentence with letters replaced by hash signs would appear on the center of the screen. This sequence would repeat itself until participants saw on the screen a message indicating the end of the experimental session. See Figure 1.

Screen 1:

"#### ######### ## ###### ### #### ##### #### ######."

Screen 2:

"That ######### ## ###### ### #### ##### #### ######."

Screen 3:

"#### croissant ## ###### ### #### ##### #### ######."

Screen 4:

"#### ######### is ###### ### #### ##### #### ######."

Screen 5:

"#### ######### ## flaky, ### #### ##### #### ######."

Screen 6:

"#### ######### ## ###### but #### ##### #### ######."

Screen 7:

"#### ######### ## ###### ### this ##### #### ######."

Screen 8:

"#### ######### ## ###### ### #### bread #### ######."

Screen 9:

"#### ######### ## ###### ### #### ##### cuts ######."

Screen 10:

"#### ######### ## ###### ### #### ##### #### neatly."

Screen 11:

"Is this bread flaky?"

Figure 1. Stimuli presentation sequence.

Participants were instructed to have their hands at all times on the button box unit. They were also instructed to wear noise-cancelling headphones during the experimental session to avoid potential distractions. Participants were then explained the task and prompted to read the instructions presented on the screen. Before the actual experimental trials were presented, participants had a practice session comprised of seven trials. An instructions reinforcement screen followed the practice section. Each participant saw only one of the three materials lists. Experimental sessions lasted about 30 minutes.

Analyses. In this self-paced reading study, we compare the online processing patterns of transitives and Middle conditions. Sentence in the Middle condition are expected to be more difficult to process than their Transitive Animate and Transitive Inanimate counterparts. In the sentence processing literature, middle constructions can be entertained under the umbrella of object-shifted sentences, that is sentences displaying the non-canonical object-verb-(subject) (OV[S]) order, such as it is also the case of passive constructions (e.g., *The ball was kicked [by John].* Object-shifted sentences have been found to display an associated extra processing cost when compared to sentences that follow the canonical subject-verb-object (SVO) order (Sekerina, 2003). However, results from different languages, obtained with different techniques, are not consistent in pinpointing the locus of the revision. Some studies present greater processing cost at the verb region (e.g., Pickering & Barry, 1991); other studies at the NP object in subject position (e.g., Miyamoto & Takahashi, 2000); and others at the post-verbal region (e.g., Sekerina, 2003). In the present study, we analyzed reading times (RTs) for these three sentence regions: (i) the NP2 in subject position (region V-1); (ii) the verb (region V); and (iii) and the adverb in post-verbal position (region V+1). Latencies for the

comprehension questions as well as percentages of correct answers were also analyzed across the three sentence types (Transitive Animate, Transitive Inanimate, and Middle). See Table 3 for sentence regions.

Table 3

Sentence regions V-1 (NP2), V (verb), V+1 (adverb) per sentence types.

	Sentence regions			
	Beginning of the sentences	Critical regions		
Sentence Type		V-1*	V*	V+1*
Middle	That croissant is flaky, but this	bread	cuts	neatly
Transitive Inanimate	That blade is dull, but this	knife	cuts	neatly
Transitive Animate	That cook is messy, but this	chef	cuts	neatly

Note: * Experimental regions.

Results and Discussion – Self-paced Reading Study. Because no time constraints were set in the experiment, raw data had to be trimmed. After plotting RTs for each sentence region for each sentence type, a ceiling of 1500 ms. determined to be a suitable one. Adoption of such ceiling allowed us to clean out extreme values due either to anticipatory (too fast reading presses) or to excessive delays in moving forward (distractions). This preserved the online properties of the measure, as well as the characteristics of the distribution at each intersection. This cut-off procedure was implemented at all levels of the region variable, except for one. The distribution of RTs for the V+1 region indicated the need for a dilated value to be implemented in order to prevent unnecessary loss of relevant data. At this position, greater RTs were expected to be registered due to an experiment artefact. Besides encompassing processing costs associated with parsing the adverb, RTs would also be liable to reflect (i) wrap-up-effects, typically found at end-of-sentence regions (e.g., Waters & Caplan, 2004) (ii) potential revision considering that the adverb was substituting the verb's direct object in the transitive conditions. Thus, specifically for the V+1 region, the duration ceiling value was set at 2000 ms. Latencies for the answers were also plotted for each sentence type, and following the same rationale, a ceiling of 10 sec. was established. Only the answers corresponding to the remaining questions were retained. "Yes" and "no" answers were then transformed into percentages of correctness. These procedures accounted for the eliminations of 231 data points, or around 3% of the raw data.

As in the normative study, RTs for the Middles *cook* sentence, were integrated into those of the Transitive Inanimate *cook* sentence before any manipulation was made on the raw data. Concerning the missing value in the Middle condition, in the items

analysis, a Missing Value Analysis was run in SPSS 19, having all the variables inserted in the model. Little's MCAR test suggested that the value was missing at random, $\chi^2(12, 264) = 12.63$, $p = .396$. The resulting EM values were then inputted into the middle condition, replacing the missing values for the *cook* sentence in all the variables here analyzed.

Subsequently, RTs for each region as well as answers' latencies were analyzed separately according to sentence type. Following Van Selst and Joliceour's (1994) suggested marker for outlier identification in a distribution with $N > 100$, all the data points falling beyond 2.5 *SD*s in the upper tails were characterized as outliers and therefore were replaced by the value of each and respectively newly calculated restricted mean $M + 2.5$ SD. This accounted for the replacement of also around 3% of the data. All further analyses were based on the resultant data. Furthermore, RTs and question-answering latencies were averaged per sentence condition according to subjects and items analyses.

Analysis of variance of the online measures. A 3 (sentence type: Transitive Animate, Transitive Inanimate, and Middle) X 3 (sentence regions: V-1, V, V+1) revealed only a statistically significant main effect of region in both participants' and items' analyses, having it accounting respectively for around 74% and 96% of the variance in the models ($F1(2, 226) = 313.17$, $p < .001$, $\eta_p^2 = .74$, $\varepsilon = 1$; $F2(2, 28) = 316.66$, $p < .001$, $\eta_p^2 = .96$, $\varepsilon = 1$). RTs were slightly greater for Tr. Ani. (M =611.90 ms., SE = 14.13), followed by Mid. (M =610.68 ms., SE = 13.24), and then by Tr. Ina. (M =600.90 ms., SE = 13.86). Region means revealed greater RTs for V+1 (M =802.98 ms.,

SE = 20.95), followed by V (*M* =523.03 ms., *SE* = 10.23), then by V-1 (*M* =497.42 ms., *SE* = 10.33).

Even though sentence type main effect was found not to be statistically significant, we proceeded to examine its simple effects at each of the sentence region levels, as these were planned comparisons. More informative to our study, results for simple effects of sentence type at each level of the region variable revealed all not to be statically significant, with one exception. Sentence type had a statistically significant effect, however negligible in its magnitude, at the V-1 region, and restricted to the participants' analysis ($F1(2, 226)= 4.67,\ p = .010,\ \eta_p^2 = .04,\ \varepsilon = .78$) [9]. In fact, pairwise comparisons[10] at the V-1 revealed a statistically significant difference between the Middle condition compared against each of the Transitive conditions (Mid. against Tr. Ani.: $t1(113) = -2.82$, $p = .017$, $d = -0.20$, $t2(14) = -2.38$, $p = .032$, $d = -0.91$; Mid. against Tr. Ina. ($t1(113) = -2.62$, $p = .030$, $d = -0.19$). No statistically significant difference was found between the Transitive conditions. See Table 4 for means and standard deviations.

[9] Albeit the effect Middle sentence type had on RTs at the V-1 region was of small magnitude ($\eta_p^2 = .04$), we reject the null hypothesis considering that if large sample size had influenced probability values, this effect would have been noticed across all levels of the region variable. However, we assume the effect at the subject position to be negligible.

[10] In each pairwise comparison conducted in each Experiment of this study, family wise alpha was corrected using Bonferroni adjustment for multiple pairwise comparisons.

Table 4

Mean RTs for regions, and for question-answering latencies, and percent of correct answers according to sentence type.

	Sentence Type					
	Middle		Transitive Inanimate		Transitive Animate	
Region	M	SD	M	SD	M	SD
V-1	482.05	113.50	504.59	129.59	505.60	124.39
V	528.71	123.86	516.00	121.71	524.367	120.28
V+1	821.26	254.83	781.95	259.79	805.738	276.29
Question	2534.57	762.02	2231.22	507.03	2275.371	631.50
Correct Answer	92.24%	13.86	91.75%	13.77	89.21%	16.54

Results from the online measures, taken in isolation, do not corroborate Di Sciullo et al.'s (2007) findings, as our results indicate no extra processing load associated with the online processing of middle sentences. Moreover, the large effect found between Middle and Transitive Animate at the V-1 region go against a broad body of the evidence showing that inanimate nouns in subject position tend to generate greater processing costs (e.g., Kaiser & Trueswell, 2004; Sekerina, 2003), thus refuting animacy effects as a viable explanation to our results. The data seem to evidence an effect of typicality of the noun occupying subject region, but in the opposite direction from that we had anticipated. The less typical *theme* subject in the Middle condition was read faster than the more typical *agent*, and *instrument* subjects respectively in the Transitive Animate and in the Transitive Inanimate conditions. The fact that NP2 mean frequency of occurrence in immediate pre-verbal subject position did not vary across conditions further supports the validity of our claim (Mid: M_{log10}= 3.08, *SD* =0.64; Tr. Ina.: M_{log10} = 2.80, *SD* = 0.48; Tr. Ani. :M_{log10} = 2.86, *SD* = 0.64; $F2(2, 28) = 1.00$, $p = .381$, $\eta_p^2 = .07$, $\varepsilon = 0.21$, t-tests were all n.s.).

Frequencies of occurrence of the word as immediate pre-verbal subject were collected from the Corpus of Contemporary American English (COCA) database (Davies, 2009). The frequencies collected were controlled for homonymy (e.g., turkey: the food, the animal, and the country). Moreover, N-V and V-Adv collocation frequency counts were searched on COCA under the parameters of V + and - two positions, respectively. Query results were zero for around one third of the pairs, thus rendering the analysis unfeasible.

Correlational analyses for the online measures. Multiple regression analyses were carried out in order to answer the question as to whether the variance in RTs could have been accounted for by word familiarity—due to frequency of use[11, 12]. Because each sentence type had a different subset of nouns in the V-1 region, we calculated simple linear regressions between RTs and noun familiarity separately for each condition. Concerning V and V+1 regions, in which the words used were kept constant across sentence type (e.g., *sells steadily*), familiarity values for each verb and adverb subsets were inserted as the dependent variable in two separate three-steps hierarchical regression analyses, respectively, with (i) Mid. RTs as predictor in the first step, with (ii) Mid. and Tr. Ina. RTs as predictors in the second step, and with (iii) Mid, Tr. Ina., and Tr. Ani. RTs altogether as predictors in the third step.

V-1 region. For the Transitive conditions, results showed that familiarity significantly accounted for about 65% of the variance in RTs for the V-1 position in the Tr. Ani condition (R^2_{adj} = .37, $F(1, 13) = 9.31$, p = .009, β = -.65), and for about 77% of the variance in the Tr. Ina. (R^2_{adj} = .56, $F(1, 13) = 18.78$, p = .001; β = -.77). For the Middle condition, even though noun familiarity accounted for about 44% of the variance, the model lacked statistical significance (R^2_{adj} = .13, $F(1, 13) = 3.16$, p = .099; β = -.44, p =.090). The negative correlation between familiarity and RTs in the transitive conditions is congruent with what's been found in the literature (e.g., Rayner, 2009). Unexpected was the nonsignificant correlation found for the Middle condition. Taken together, these

[11] Word frequencies used in the models were the same ones described and used in the analyses of the data in the normative study. In fact, those same frequencies are used for all further analyses.

[12] Even though collocation frequencies for N–V, and V–Adv were important to the analyses, on average, 30% of those pairs were null in the COCA database.

results might suggest that the parser employed dissimilar strategies while processing the subject in the middle constructions. We elaborate on this topic in the remainder of this section.

V region. For the V region, results indicated that none of the models were strong nor statistically significant ((i) step 1: R^2_{adj} = -.05, $F(1, 13) = 0.30$, $p = .592$; (ii) step 2: R^2_{adj} = -.03, $F(2, 12) = 0.78$, $p = .480$; (iii) step 3: R^2_{adj} = -.12, $F(3, 11) = 0.51$, $p = .682$). In fact, none of the steps displayed a statistically significant change in the proportion of variance accounted for in models. See Table 5 for detailed results. These results suggested that familiarity was not a factor in accounting for the variance in verb RTs for any of three sentences types. Taking into consideration results from the analysis of variance at the V region in which no differences were found between types, the present results might be interpreted as an indication that verbs are indeed more complex lexical items, in a way that more salient information for their processing could be derived from sentence-internal relations, such as argument relations, more than from familiarity with its lexical form. This being so, the parser would rely less on word familiarity while processing verbs, congruently with our results.

V+1 region. For the V+1 region, even though none of the models were found statistically significant ((i) step 1: R^2_{adj} = -.04, $F(1, 13) = 0.51$, $p = .487$; (ii) step 2: R^2_{adj} = -.11, $F(2, 12) = 0.32$, $p = .731$; (iii) step 3: R^2_{adj} = .16, $F(3, 11) = 1.89$, $p = .189$), results from step 3 showed that the insertion of Tr. Ani. RTs produced a statistically significant change in the model, having Tr. Ani. adverb RTs and familiarity sharing around 54% of variance (r_{12}= -.54). See Table 5 for detailed results.

Table 5

Results from Hierarchical Multiple Regression Analyses between Familiarity and V (verb), and V+1 (adverb) RTs.

	Word frequency					
	Step 1		Step 2		Step 3	
Sentence Type	ΔR^2	β	ΔR^2	β	ΔR^2	β
V						
Middle	.02	.15	.09	.34	.01	.31
Transitive Inanimate				-.36		-.39
Transitive Animate						.10
V+1						
Middle	.04	-.20	.01	-.26	.29**	-.31
Transitive Inanimate				.13		.14
Transitive Animate						-.54*

Note. * $p = .05$, ** ΔF (1,11)= 4.828, $p = .05$

Familiarity only significantly accounted for the variance in RT at the V+1 Transitive Animate intersection. Such inconsistency does not come as a surprise because RTs for the V+1 region are possibly encompassing effects other than the processing of the adverb alone. As noted earlier, it is at this region where the expected internal object, normally present in transitive constructions, is replaced by the adverb in the Transitive conditions. Effects derived from such mismatch could have been captured by the RTs.

Analyses of off-line measures. Question-answering latencies means for each sentence type showed that participants took slightly longer to answer the questions following Middle sentences (see Table 4 for means and standard deviations). A 3 (sentence type: Middle, Transitive Animate, Transitive Inanimate) X 1 (latency) repeated measures ANOVA revealed a statistically significant effect of sentence type, restricted to the participants analysis, but accounting for only around 12% of the variance ($F1(2, 226)= 14.81$, $p = 0.00$, $\eta_p^2 = .12$, $\varepsilon = 1$). Pairwise comparisons confirmed that, also in the participants analysis, it took longer to answer questions in the Middle condition compared against each of the Transitive conditions (Mid. against Tr. Ani. : $t1(113) = 3.81$, $p = .001$, $d = 0.37$; Mid. against Tr. Ina.: $t1(113) = 4.79$, $p < .001$, $d = 0.47$). No statistically significant difference was found between the Transitive conditions.

Concerning the percentage of correct answers, results indicated that around 90% of the questions were answered correctly, regardless of sentence type. See Table 4 for means and standard deviations. Indeed, results from a 3 X 1 repeated measures ANOVA between sentence type and percent correct responses, as well as further pairwise comparisons did not reach statistical significance ($F1(2, 226)= 1.82$, $p = .167$, $\eta_p^2 = .02$, $\varepsilon = .10$; $F1(2, 28)= 0.37$, $p = .531$, $\eta_p^2 = .05$, $\varepsilon = .15$)

Results from the off-line measures seem to indicate a trade-off effect between speed and accuracy, but restricted to the Middle condition. The fact that participants took longer to correctly answer the comprehension questions associated with middle sentences could be interpreted as evidence for greater interpretational cost associated with middle's non-canonical configuration. But we cannot dismiss plausibility as a possible confound. Additionally, we also cannot exclude the possibility that latency measures captured a spill over effect derived from the greater integration cost linked to the processing of the sentences in the Middle condition. This argument for the capture of a spill-over effect becomes more salient when we consider that the V+1 critical region was at the end-of-sentence, a region which is traditionally associated a wrap-up effect.

Comparison to Di Sciullo et al. (2007). The main purpose of Experiment 1 was to investigate whether the findings from Di Sciullo et al. (2007)'s online task, which were interpreted as evidence that middle constructions are more complex than their transitive counterparts, would be replicated once the rejection rates of middles construction were reduced. We failed to replicate their findings, as our results revealed no extra processing load in the online processing of the middle construction. The data showed a statically significant facilitatory effect over the processing of the subject in the middle constructions. Off-line measures, on the other hand, corroborate the longer question-answering latencies previously found for middles (Di Sciullo et al., 2007; Maia et al., submitted). However, we do not interpret results as evidence against greater structural complexity of middle constructions. It would seem that our attempt to rule out high rejection rates as a possible confound generated another variable we were not expecting: the priming for an unaccusative structure.

In contrast with the broadly established phenomenon that object-shifted constructions are more complex to process (Kaiser & Trueswell, 2004; Miyamoto & Takahashi, 2000; Pickering & Barry, 1991; Sekerina, 2003), our results were consistent in showing no effects of processing complexity for middle sentences. We argue that the addition of adjectival predicative clauses as the first conjunct facilitated the processing of sentences in the Middle condition. In the remainder of this section, we discuss some evidence supporting this argument,

First, there is evidence suggesting that contextual information can modulate parsing expectations. Altmann and Steedman (1988), Arai, van Gompel, and Scheepers (2007), Kaiser and Trueswell (2004), and Sekerina (2003) found that difficulties associated with the processing of non-canonical object-shifted sentences can be alleviated given the supportive context. Kaiser and Trueswell (2004), for instance, examined the processing of object-initial sentences in Finish, constructions which are more frequently used when the object is a given term in the discourse (*rheme*), as opposed to a when it is a new term (*theme*). Their results indicated that the processing of the least canonical OS-new was facilitated given the supportive context. Furthermore, there is emerging evidence indicating that parsing strategies can, in fact, be more directly influenced by preceding structures, showing that preceding semantic and syntactic information facilitate the processing of subsequent similar semantic items, and similar syntactic structures (Frazier, Munn, & Clifton, 2000; Knoeferle & Crocker, 2009; Ledoux; Traxler, & Swaab, 2007; Tooley, Traxler, & Swaab, 2009). The phenomena of semantic and syntactic priming are discussed in greater depth in the General Discussion section.

A second type of evidence for the facilitation of our middle constructions is that adjective predicates and middles constructions both exhibit an *experiencer* or *theme*, in subject position. Also, both constructions have been proposed to convey a predicative statement concerning the entity in subject region (Ackema & Schoorlemmer, 1994; Fagan, 1992). But it appears that these constructions are similar at a deeper level. According to Hale and Keyser (2002), middle verbs are object-*experiencer* verbs. As such, those verbs' argument structure has the *experiencer*, or *theme*, NP raised to sentence subject position born in VP-internal Spec position. In Alexiadou, Haegeman, and Stavrou (2007, p. 68, 290), we see that the copula verb *be* is an unaccusative verb, which selects an AdjP as its only complement. It is the NP appearing in Spec position on the AdjP which then raises to surface subject position. Notwithstanding potential theory-internal discrepancies, relevant here is the notion that both constructions are ultimately formed via argument movement (move-α operation (Chomsky, 1981)), having the argument raised from the right periphery of the verb. If so, then our results could be taken to support the pre-syntactic derivation of middle constructions, as we entertain that the semantic and syntactic characteristics of the preceding adjectival conjunct seem to have dissuaded any extra processing cost expected to be found in such object-shifted sentences. Furthermore, we argue that the facilitated parsing of middle constructions to be due to more than just the semantic similarities between NP1 and NP2. If we had only captured an effect of semantic priming, then no sentence differences would have been found at V-1 region, since both NP1 and NP2 were semantically related across sentence types. It is also possible that expectations for a parallel structure could have been elicited by the coordinative conjunction *but* (see Knoeferle & Crocker, 2009). Once expectations

were ratified by the semantic characteristics of the Middle NP2, that of being a *theme*, the same syntactic representation was then reapplied. The indication that the middle subject was read faster than the also inanimate Transitive Inanimate subject can be explained by the greater conformity of a *theme* (e.g., *bread*), as opposed to an *instrument* (e.g., *knife*), to be the complement of the verb in a unaccusative construction. Thus, results at V-1 region ultimately support our hypothesis for typicality effects over the processing of the nouns in subject position. In other words, at the same time that an *agent* subject is more congruent with a transitive active construction, a *theme* subject seems to be more congruent with a middle construction (Keiser & Trueswell, 2004; Sekerina, 2003).

In sum, data are consistent in showing that priming effects can strongly alter parsing heuristics, and that argument structure information should be available during on line processing. This is mostly evidenced by the results from the V-1 region in the Middle condition, which showed (i) faster RTs as well as (ii) a nonsignificant correlation with noun familiarity for the Middle NP2. More importantly, our results can be taken as evidence for middle verbs to reach syntax as unaccusatives, in agreement with A-shift.

However, the effects captured in Experiment 1 might have been constrained by the nature of the technique implemented. For instance, an experimental artefact attributable to greater integration costs derive from the moving-window task, which only allow participants to read on word at a time. This could have affected the RTs observed at the end-of-sentence position, as well as the question-answering response latencies. Another possible confound in the present experiment derives from the fact that the critical V+1 region is also the end-of sentence. We correct that in subsequent experiments by adding an end-of-sentence PP to our materials.

The purposes of Experiment 2 are to (i) serve as normative for the new materials, which are also used in Experiment 3, and (ii) probe for differences in plausibility and in syntactic complexity between sentence types. The latter measures would help clarify whether greater question-answer latencies found for the middles condition would be reflecting plausibility oddity, or whether they could constitute an spill-over effect, as discussed above.

Experiment 2

Gruber and Gibson (2004) presents a questionnaire paradigm designed to scrutinize different phenomena which can potentially affect results in sentence processing studies, such as plausibility and structural complexity. In this paradigm, sentences are presented in pairs, with first and second sentences paraphrasing one another. Participants are to judge the pairs regarding the degree at which (i) sentences *a* and *b* describe the same situation, and (ii) the described situation is a plausible one. They are also required to judge each sentence separately regarding the degree at which (iii) the construction is an easy, natural, or simple way to convey that information. Gruber and Gibson (2004) found that, once plausibility is partialled out, the ratings in (iii) capture effects of syntactic complexity more accurately than would traditional goodness scales. For instance, they observed that whereas paraphrased passive and transitive active sentences pairs (*e.g.*, respectively: *The house was built by the architect,* and *The architect built the house*) have similar meaning, as suggested by scale (i) ratings, active constructions constitute a significantly easier way to convey the message, as suggested by scale (iii) ratings. These findings were interpreted as evidence that such protocol is in fact sensitive to structural complexity, given that passives are to be syntactically derived via move-α (Rohde &

Gibson, 2003; Gruber & Gibson 2004). We make use of this experimental paradigm to further assess the syntactic and semantic complexity of our materials.

Method

Participants. Forty-five undergraduate students from the Department of Psychology at Concordia University, Montreal, participated in the experiment in return for course credits. They were 39 females and 6 males, with ages ranging from 18 to 48 (*M*=25.36; *SD=7.22*). Twenty-nine of them had corrected-to-normal vision. They were all native speakers of Canadian English, having learned the language before the age of 3. Only four were English monolinguals, and 41 spoke one or more languages other than English.

Materials. Three booklets were created. Each booklet contained 30 filler pairs, and 15 experimental pairs. Filler pairs were comprised of 50 % good paraphrases (*e.g.,* a. *John saw the pen's cap that was on the book*, b. *John saw the cap of the pen that was on the book*), and 50% bad paraphrases (*e.g.,* a. *The doctor healed the patient,* b. *The patient sent for the doctor*). Experimental sentences were comprised of those used in Experiment l added of an end of sentence PP. The resulting structure was NP1-be-AdjP, *but* NP2-V-AdvP-PP. See (16), (17), and (18). The full set of materials is provided in Appendix B. Materials in the Transitive Animate condition, as shown in (16a), and in the Transitive Inanimate, as shown in (17a), were paired with a paraphrase containing a complex adjectival predicative construction: respectively, a Predicative Animate construction, as shown in (16b); and a Predicative Inanimate constructions, as shown in (17b). Materials in the Middle condition were paired with a paraphrase containing a Passive construction, as shown respectively in (18 a, b).

(16) Animate Pair:

a. Transitive Animate: That cook is messy, but this chef cuts neatly with the knife.

b. Predicative Animate: That cook is messy, but this chef is a neat cutter with the knife.

(17) Inanimate Pair:

a. Transitive Inanimate: That blade is dull, but this knife cuts neatly through any meat.

b. Predicative Inanimate: That blade is dull, but this knife is a neat cutter through any meat.

(18) Middle Pair:

a. Middle: That croissant is flaky, but this bread cuts neatly into small slices.

b. Passive: That croissant is flaky, but this bread was cut neatly into small slices.

As in Experiment 1, (i) the verb and its associated adverb were constant across sentence types, and (ii) NP1 and NP2 were semantically related, but varied according to sentence type: Transitive Animate and Predicative Animate had animate *agent* nouns (*e.g., cook/chef*); Transitive Inanimate and Predicative Inanimate, had inanimate *instrument* nouns (*e.g.*, blade*/knife*); and Middle and Passive, had inanimate *theme* nouns (*e.g.*, *croissant/bread*). First conjuncts were kept constant within the pairs. Word length varied from 4 to 10 letters ($M = 5.88$; $SD = 1.60$) for the verb, from 4 to 9 letters ($M = 6.80$; $SD = 1.37$) for the adverb, and from 4 to 9 letters ($M = 5.62$; $SD = 1.45$) for the NP2. NP2 length did not vary across conditions ($F2(2, 28) = 0.616$, $p = .547$, $\eta_p^2 = .04$, $\varepsilon = .14$). As expressed by frequency of occurrence, word familiarity results were ($M_{Log10} = 2.13$; $SD = 0.59$) for verbs, ($M_{Log10} = 2.66$, $SD = 1.02$) for adverbs, and ($M_{Log10} = 2.82$,

SD =0.65) for NP2s[13]. NP2 familiarity also did not vary across conditions ($F2(2, 28) = 3.12$, $p = .060$, $\eta_p^2 = .18$, $\varepsilon = .55$). Moreover, NP2 frequency as the immediate pre-verbal subject did not vary across conditions (Mid.: $M_{log10} = 3.07$, $SD = 0.65$; Ina.: $M_{log10} = 2.80$, $SD = .48$; Ani.: $M_{log10} = 2.86$, $SD = .64$; $F2(2, 28) = 0.99$, $p = .386$, $\eta_p^2 = .07$, $\varepsilon = .20$).

Experimental sentences were distributed following a Latin square design, with one third of the verbs in each condition, without repeating verbs. Items order was different for each booklet version, as it was determined pseudo-randomly following the same procedure described in Experiment 1. Experimental materials were counterbalanced regarding the order of presentation (position *a* or *b)* within and across lists. Because experimental materials were in odd numbers (5 per condition, 15 in total) sentences in two of the conditions (e.g., Transitive Animate and Transitive Inanimate.) appeared following the ratio 3 in position *a* by 2 in position *b*; and sentences in the third condition (e.g., Middle), followed the ratio 2 in position *a* by 3 in position *b*. In each booklet version, a different condition followed the 2*a*/3*b* ratio: in booklet version 1, it was the Middle condition; in booklet version 2, Transitive Animate, and in booklet version 3, Transitive Inanimate.

Procedure. Participants were instructed to rate the sentences according to four categories: (I) similarity of meaning, as in (19a), from 1(*similar*) to 7 (*different*); (II) plausibility of the situation described, as in (19b), from 1 (*plausible*) to 7 (*implausible*) pairs; and (III) simplicity of the structure, as in (19c, d), from 1 (*ease, natural, simple*) to

[13] Collocation frequency counts for N–V, and V–Adv, had an averaged 30% empty return in the COCA database.

7 (*not easy, nor natural, or simple*). The lower the ratings, the more similar and plausible the pairs are; and the less complex are the structures.

(19) a. The degree to which the two sentences describe the same situation.

b. The degree to which the situation described is a plausible or natural situation, i.e., likely to occur.

c. The degree to which sentence (a) is an easy, simple or natural way to express that situation.

d. The degree to which sentence (b) is an easy, simple or natural way to express that situation.

Participants were further instructed not to dwell on any of the sentences, and to answer according to their intuition.

Analyses. With this paradigm we assessed (i) whether plausibility is a confound in our materials, (ii) whether middles, passives, as well as adjectival predicatives, are perceived as more complex structures than transitive active sentences, and (iii) whether adjectival predicatives are good paraphrases of transitive constructions—the latter was used as a way of investigating whether the first conjunct (also predicative construction) can also provide a supportive context for the transitive conditions. We analyzed mean ratings, as well as the frequency of rate distribution across (i) the three pair types (Animate, Inanimate, and Middle) in the Similarity, and in the Plausibility scales, and across (ii) the six sentence types (Transitive Animate, Predicative Animate, Transitive Inanimate, Predicative Inanimate, Middle, and Passive) in the Simplicity scale. Ratings for the 15 sentences in each pair type, as well as in each sentence type, were averaged according to participants' and items' analyses, per scale.

Results and Discussion

Similarity Scale.

Analysis of variance. In the Similarity scale, overall means for the three pairs fell at the lower half of the scale (1 to 3), which indicated that participants perceived target and paired sentences as conveying the same meaning (Gruber & Gibson, 2004). See Table 6 for means and standard deviations. Nevertheless, a 3 (pair type: Animate, Inanimate, and Middle) X 1(ratings) repeated measure ANOVA revealed that, restricted to the items analyses, the middle-passive pair conveyed less similar meanings, than did the transitive-predicative pairs in both Animate and Inanimate conditions ($F1(2, 88)= 42.29$, $p < .001$, $\eta_p^2 = .49$, $\varepsilon = 1$, $F2(2, 28)= 25.27$, $p < .001$, $\eta_p^2 = .64$, $\varepsilon = 1$). Pairwise comparisons confirmed a statistically significant strong difference between the Middle pair and each transitive pair (Mid. pair vs. Ani. pair: $t1(44) = 7.07$, $p < .001$, $d = 1.22$; $t2(14) = 5.28$, $p < .001$, $d = 2.03$; Mid. pair vs. Ina. pair ($t1(44) = 6.73$, $p < .001$, $d = 1.17$; $t2(14) = 6.10$, $p < .001$, $d = 2.15$). The difference between the transitive pairs was found not to be statistically significant.

Frequency analysis. Because the values entered into the ANOVA were collapsed means, we ran a 3 (pair type) X 7 (rate) Chi Square analysis using the raw data, having ratings transformed into a categorical variable. In the comparison between the Middle pair against each of the transitive pairs, the analyses confirmed an effect of pair type on the rating pattern employed by participants (Mid. pair vs. Ani. pair: $\chi^2(6, 450) = 67.65$, $p < .001$, $\Phi^2 = .39$; Mid. pair *vs*. Ina. pair: $\chi^2(6, 450) = 95.85$, $p < .001$, $\Phi^2 = .46$). Results also show that the differences in frequencies between the transitive pairs across the rate

variable were, in fact, not due to chance (Ani. pair vs. Ina. pair: $\chi^2(6, 450) = 14.97$, $p = .020$, $\Phi^2 = .18$). See Table 7 for percentages of rate frequencies across pair types.

Table 6

Means and Standard Deviations for Sentence Types across Rating Categories

	Scale					
	I- Similarity [a]		*II-Plausibility* [b]		*III-Simplicity* [c]	
Sentence Type	*M*	*SD*	*M*	*SD*	*M*	*SD*
Middle	3.12	1.45	1.77	0.82	2.96	1.20
Transitive Inanimate	1.72	0.92	1.56	0.76	1.79	0.77
Transitive Animate	1.71	0.82	1.48	0.65	2.09	0.91
Passive					2.54	1.24
Predicate Inanimate					4.05	1.45
Predicate Animate					3.70	1.40

Note: [a, b] Means and standard deviations for sentence pairs ratings. [c] Means and standard deviations for individual sentences ratings. [a] 1 = describe same situation, 7 = describe different situations. [b] 1 = plausible, 7 = implausible. [c] 1 = easy, natural, simple, 7 = not easy, natural, simple.

Table 7

Percentages of Similarity, and Plausibility Rates across Pair Types.

	Scale						
	I- Similarity						
Pair Type	1	2	3	4	5	6	7
Middle	35.6	16.0	12.0	3.6	14.2	7.1	11.6
Inanimate	76.4	11.1	0.9	4.4	1.8	1.8	3.6
Animate	68.4	14.2	7.1	4.0	3.1	0.9	2.2
	II-Plausibility						
Middle*	65.3	14.7	8.4	4.0	4.0	1.8	1.3
Inanimate	72.9	12.0	8.0	3.1	2.2	1.3	0.4
Animate	73.8	13.8	6.2	3.1	2.7	0.4	0.0

Note: *0.44% of void data.

We entertain that ratings in the Similarity variable may have captured aspectual differences derived from the different verb tense used in the middle and in the passive verbs. This factor, unique to the Middle pair, could be responsible for the differences in similarity found between the Middle and each of the transitive pairs in both ANOVA and Chi Square analysis. Concerning the difference in the frequency of rates found between the transitive pairs, we speculate that such a small effect could be possibly related to the agentivity of the NP2, although we cannot rule out an effect of animacy. Moreover, results from the Similarity scale show that transitive and predicative constructions convey similar meanings, thus suggesting that the first conjunct can also provide supportive context for the Transitive Animate, and the Transitive Inanimate sentence types. Although at a lesser degree, Middles and Passives were also found to display similar meaning, given that (i) mean ratings were at the threshold of 3, and that (ii) the greater concentration of ratings fell on the lower end of the scale, about 63%. Our design does not allow for a direct comparison between Middle and Predicative sentences in this scale.

Plausibility scale.

Analysis of variance. In the Plausibility scale, overall means for the three pairs also fell at the lower end of the scale, which indicated that the situations conveyed by the materials were plausible ones (Gruber & Gibson, 2004). See Table 6 for means and standard deviation. A 3 (pair type) X 1 (ratings) repeated-measures ANOVA revealed a small effect, which accounted for only around 10% of the variance in the participants analysis ($F1(2, 88)= 4.77,\ p = .011,\ \eta_p^2 = .10,\ \varepsilon = 78$). Pairwise comparisons showed only a statically significant difference—almost reaching moderate magnitude— between

Middle and Animate pairs, also at the participants analysis ($t1(44) = 2.94$, $p = .016$, $d = 0.38$).

Frequency analysis. The analysis of the frequency of counts in the rate variable confirmed that the concentration of rates was not significantly different across pair types (Mid. pair vs. Ani. pair: $\chi^2(6, 449) = 7.62$, $p = .267$, $\Phi^2 = .13$; Mid. pair vs. Ina. pair: $\chi^2(6, 449) = 4.09$, $p = .665$, $\Phi^2 = .10$; and Ani. pair vs. Ina. pair: $\chi^2(6, 450) = 2.88$, $p = .824$, $\Phi^2 = .08$). See Table 7 for the rates frequencies across pair types.

Taking into consideration results from Experiment 1, we could interpret the present results as possibly reflecting a typicality effect of the noun in subject position. Concerning the pairwise comparisons, the lack of statistically significant difference between Middle and Inanimate pairs, as opposed to the statistically significant difference between Middle and Animate pairs, could be explained in terms of the predicatives paired with the transitive sentences. Note that the compound occupying the complex AdjP (e.g., *neat cutter*) in both Predicative Animate and Predicative Inanimate seems to be more compatible with an *agent* NP (e.g., *chef*) than with an *instrument* NP (e.g., *knife*). The animate agentive subject *chef* in the Animate pair is more plausibly a *neat cutter*, than the inanimate instrument subject *knife* in the Inanimate pair is. Such aspect might have been captured by the Plausibility scale. Moreover, due to the small magnitude of the effect found in the ANOVA, $\eta_p^2 = .10$, added to the fact that the distribution of rates (1 to 7) frequencies was found not to be dependent on pair type, we rule out plausibility as a possible confound affecting our study.

Simplicity scale.

Scope of the Simplicity scale. Gruber and Gibson (2004) also examined whether results from McKoon and Macfarland (2000), which suggested that externally caused change of state verbs (e.g., *break*) were more complex than internally caused change of state verbs (e.g., *bloom*). Gruber and Gibson tested those same verbs according to a 2 (verb type: externally, and internally caused change of state verbs) X 2 (sentence type: transitive and intransitive constructions) (e.g., (i) *The missile exploded*, and (ii) *The scientists exploded the nuclear device*). Results did not show plausibility to be a confound, but the analysis of the Simplicity scale showed that, restricted to the transitive condition (ii), externally caused verbs were rated as more complex than internally caused verbs. Relevant to us is the fact that some of the verbs used in both externally and internally caused conditions were in fact also non-alternating ergatives, such as *bloom*, and *splinter*. It is important to note that no analysis comparing transitives against intransitives counterparts was presented by Gruber and Gibson (2004). Also, their study did not control for ergativity. Noting that it remains unclear whether alternating and non-alternating ergatives are syntactically derived the same way (Friedmann et al., 2008), we can only speculate on the possibility that the differences in simplicity ratings between verb types found exclusively at the transitive level could be reflecting canonicity effects, given that non-alternating ergatives are canonically found in intransitive constructions.

For the Simplicity scale, we analyzed both target and paired sentence types: Middle, Transitive Animate, Transitive Inanimate, Passive, Predicative Animate, and Predicative Inanimate respectively. This was done for two reasons. First, because results from the Experiment 1 suggested an effect derived from the first conjunct of the

coordination—which is occupied by an adjectival predicative construction (e.g., *This cook is messy*,)—:we wanted to investigate whether the proposed structural complexity of predicative constructions, possibly formed via a move-α operation, would be reflected in the scale. Second, to investigate whether the paired sentences could have had influenced the ratings of the target sentences.

It was expected that the transitive sentence types receive the lowest scores in this scale. As opposed to the other constructions tested, the syntactic structure of these transitive constructions does not postulate argument displacement, such as move-α operations. Transitive active constructions also conform with the most canonical linear order of arguments—that is external argument in subject position, and internal argument in object position. Moreover, both sentence types—the Transitive Animate to a greater degree than the Transitive Inanimate—better conform with θ-role hierarchy, as they present an agent (e.g. *chef*), and an instrument (e.g., *knife*) in subject position.

It was also expected that Middle and Passive sentence types receive greater scores than the transitives, as the syntactic structure of middle and passive constructions would entail argument displacement. The passive derivation may preclude a more complex syntactic structure than that of middles (Di Sciullo, 2005). However, passive constructions could, in fact, present a processing advantage over middles. The passive morphology would hint the processor as to which syntactic structure should be applied. As a consequence, participants could perceive passives as less complex than middles. Moreover, no differences due to animacy, or to θ-role preference can be postulated between these sentence types, since the subject in both sentence types is the same *theme* inanimate NP(e.g., *bread*).

Predicative constructions should display the greater scores in the simplicity scale. Not only the copula verb *be* deploys an unaccusative syntactic structure—which precludes argument displacement—but the complement of the verb is a complex AdjP, which accounts for the compound *neat cutter* (Alexiadou et al., 2007).

Analysis of variance. With the exception of Predicative Animate and Predicative Inanimate, overall means fell in the lower half of the scale (below 4), which would indicate that sentences in the Middle, Passive, and both the transitive conditions were judged as being an easy, simple, or natural way to express a situation (see Gruber & Gibson, 2004). See Table 6 for means and standard deviations. A 6 (sentence type: Middle, Transitive Animate, Transitive Inanimate, Passive, Predicative Animate, and Predicative Inanimate) X 1 (rating) repeated-measures ANOVA analysis indicated a main effect of sentence type, accounting for around 50% and 63% of the variance in the participants' and items' analyses respectively ($F1(5, 220)= 44.90,\ p < .001,\ \eta_p^2 = .51,\ \varepsilon = 1$, $F2(5, 70)= 24.30,\ p < .001,\ \eta_p^2 = .63,\ \varepsilon = 1$). Pairwise comparisons revealed that, consistent across participants' and items' analyses, were the statistically nonsignificant differences in complexity between (i) Middles and Passives ($p = 1$, and $p = 1$, respectively), (ii) between the transitive conditions ($p = .074$, and $p = 1$, respectively), and (iii) between the predicatives ($p = .265$, and $p = 1$, respectively).

However, pairwise comparisons did reveal that middle constructions were perceived to be more complex than transitive ones (Mid. vs. Tr. Ani.: $t1(44) = 4.85$, $p < .001$, $d = 0.82$, $t2(14) = 5.20$, p = $.002$, $d = 1.04$; Mid. vs. Tr. Ina. : $t1(44) = 6.20$, $p < .001$, $d = 1.16$; $t2(14) = -5.85$, $p = .001$, $d = 1.48$), but to be less complex than predicatives ones (Mid. vs. Pr. Ani.: $t1(44) = -3.51$, $p = .016$, $d = -0.56$; Mid vs. Pr. Ina.: $t1(44) = -5.77$, p

$< .001$, $d = -0.81$). A similar pattern was found for the passive condition, with the exception of the Passive and Transitive Animate comparison, which did not reach statistical significance at neither participants' nor items' analyses ($p = .075$; $p = .130$, respectively) (Pass. vs. Tr. Ina.: $t1(44) = 5.00$, $p < .001$, $d = 0.73$, $t2(14) = 6.24$, $p < .001$, $d = 1.22$; Pass. vs. Pr. Ani.: $t1(44) = -6.33$, $p < .001$, $d = -0.87$, $t2(14) = -3.70$, $p = .036$, $d = -1.28$; Pass. vs. Pr. Ina.: $t1(44) = -7.17$, $p < .001$, $d = -1.11$; $t2(14) = -5.18$, $p = .002$, $d = -2$). Regarding the predicatives, they were also rated as more complex than the transitive constructions (Pr. Ani. vs. Tr. Ani.: $t1(44) = 8.59$, $p < .001$, $d = 1.36$, $t2(14) = 5.00$, $p = .003$, $d = 1.70$; Pr. Ani. vs. Tr. Ina.: $t1(44) = 9.47$, $p < .001$, $d = 1.68$, $t2(14) = 8.22$, $p < .001$, $d = 2.13$; Pr. Ina. vs. Tr. Ani.: $t1(44) = 9.24$, $p < .001$, $d = 1.61$, $t2(14) = 6.99$, $p < .001$, $d = 2.44$; Pr. Ina. vs. Tr. Ina.: $t1(44) = 10.55$, $p < .001$, $d = 1.93$, $t2(14) = 10.84$, $p < .001$, $d = 3.03$).

Results from the Simplicity scale clearly suggest that transitive constructions were the least complex of the sentence types entered in the model. Results also showed that the level of simplicity for middle and passive constructions was the same, and that both sentence types were more complex than transitives. These results are in line with the proposal that both middle and passive constructions are derived via the operationalization of a voice morpheme (Alexiadou & Doron, 2012; Bowers, 2002). However, given the off-line nature of the measure, such results are better interpreted as supporting evidence for Experiment 1's findings. In other words, despite middles being more complex structures to process than transitive constructions, the processing of middles can be facilitated granted the appropriate context.

Predicative sentences were found to be the most complex constructions in the model (means were around 4). This was congruent with our expectations. Concerning the differences found in the analysis of variance between the sentences in the Middle pair (Middle and Passive) and the predicatives (Animate and Inanimate), we suggest that the Simplicity scale might have indeed captured the extra structural complexity associated with the post-verbal complex AdjP (e.g., the compound *neat cutter*) in the predicative conditions, as opposed to the simplex post verbal AdvP (e.g., neatly) in the Middle and Passive conditions (see Alexiadou et al., 2007).

Frequency analysis. Rates 1 through 7 were transformed into a categorical variable and were tabulated into frequencies tables. A closer look to the distribution of ratings revealed an increase of counts at the higher end of the Simplicity scale, that is, ratings from 5 to 7, for the structurally complex sentence types (Middle, Passive, Predicative Animate, and Predicative Inanimate). See Table 8 for percentages of rate frequencies across the Simplicity scale. The analysis revealed that the frequencies of rates (1 to 7) were indeed associated with sentence type in the comparisons between Middle and predicatives (Mid. and Pr. Ani.: $\chi^2(6, 449) = 18.65$, $p = .005$, $\Phi^2 = .20$; Mid. and Pr. Ina.: $\chi^2(6, 449) = 34.28$, $p < .001$, $\Phi^2 = .28$), but also between middle and passive conditions (Mid. and Pass.: $\chi^2(6, 448) = 15.41$, $p = .017$, $\Phi^2 = .19$). Sentence type was also found to affect the frequency of rates through the Simplicity scale in the comparisons between Passive and predicatives (Pass. and Pr. Ani.: $\chi^2(6, 449) = 40.68$, $p < .001$, $\Phi^2 = .30$; Pass. and Pr. Ina.: $\chi^2(6, 449) = 62.58$, $p < .001$, $\Phi^2 = .37$).

Regarding animacy, whereas differences in frequency of rates between the predicative conditions were found to be due to chance (Pr. Ani. and Pr. Ina.: $\chi^2(6, 450) = 4.94$, $p = .552$, $\Phi^2 = .11$), the comparison between the transitive conditions did reveal an association of sentence type (Tr. Ani. and Tr. Ina.: $\chi^2(6, 450) = 14.97$, $p = .020$, $\Phi^2 = .18$).

Table 8

Percentage of Simplicity Rates across Sentence Types.

Sentence Type	Simplicity Scale Rating						
	1	2	3	4	5	6	7
Middle*	38.22	13.78	16.00	5.33	8.89	5.78	11.56
Transitive Inanimate	60.44	22.22	8.44	1.78	2.67	2.67	1.78
Transitive Animate	53.78	18.23	12.44	4.44	5.33	3.56	2.22
Passive*	40.44	18.22	16.89	7.11	7.11	7.11	2.67
Predicative Inanimate	17.33	12.00	15.56	6.67	17.78	13.33	17.33
Predicative Animate	24.89	10.22	15.56	7.56	16.44	12.00	13.33

Note: *0.44% of void data.

The Simplicity scale captured the different degrees of structural complexity attributed to the different sentence types in the study: in crescent order from simple to complex are: (i) Transitive Animate, Transitive Inanimate, (ii) Middle, Passives; and (iii) Predicative Animate, Predicative Inanimate. The Chi Square analyses revealed a further dimension of the effects captured by the scale. At the same time that Chi Square results corroborated the differences in complexity found between (i), (ii), and (iii), they also showed internal differences between (i) and (ii), but not between (iii).

Concerning the differences found between transitive conditions, we argue that these results are consistent with the notion that the more typical *agent* subject in Transitive Animate more strongly favours the canonical NP-V-NP transitive construction. Accordingly, because both transitive sentences had an adverb in the immediate post-verbal position, the absence of the direct object could explain the higher percentage of 5 to 7 ratings (11%), and the lower percentage of 1 to 3 ratings (84%) observed in the Transitive Animate condition, when compared to the Transitive Inanimate (7%, and 95% respectively). Such argument is partially in agreement with our results, in the sense that less canonical object-shifted constructions—that is, middles and passives—were found to be more complex than the canonical transitive constructions in Transitive Animate, and Transitive Inanimate, especially once all the verbs were repeated across these sentence types. But canonicity in itself does not explain the greater complexity rates for the predicative constructions (Predicative Animate and Predicative Inanimate). If indeed the simplicity measure captures semantic complexity, then it is possible that the greater complexity found for the predicates could be attributed to the semantic oddity. See de

Almeida, Manouilidou, Roncero & Riven (2011) for evidence suggesting that novel semantic relations render metaphoric constructions more complex.

Concerning the differences found between middles and passives, an explanation based on effects of typicality of the noun in subject position becomes invalid since both constructions display the same demoted inanimate *theme* NP2 in pre-verbal position. However, even though we cannot provide further corroborating data, we argue it would be hard to dispute the claim that between these two object-shifted constructions, passives not only seem to be more frequently used than middles, but they also do display a morphological marker, that is the auxiliary verb *to be*, which unambiguously indicates that such sentences follow a passive construction (see Maia et al., submitted for a discussion on morphological clues on sentence comprehension). Moreover, we cannot discard an effect of participants' lower familiarity with middle constructions accounting to the higher percentage of 5 to 7 rates observed for Middle.

Concerning the results found for the complex predicative conditions, which showed a lack of statistically significant difference in ratings pattern between PA and PI, we argue that, given the limitations of the scale, any potential effects of typicality of the noun in subject position would have been masked by their structural complexity. Furthermore, the analysis of rate frequencies in the Simplicity scale suggests that participants indeed made use of the whole scale, and that it was done independently from sentence pairing. This weakens the validity of the argument that the greater complexity found for the predicative conditions could have been intensified because of their pairing with the simplest transitive constructions, that is the sentences in the Transitive Animate and Transitive Inanimate conditions.

With the present experiment, we were also able to replicate Gruber and Gibson's (2004) findings that passive constructions are perceived as more complex than transitive active ones. Our results also corroborated the validity of this paradigm in capturing distinctive characteristics—such as semantics, structural complexity, and plausibility—for a varied set of construction types. Results suggested that the questionnaire was sensitive enough to reflect differences derived from verb tense. This was evidenced by the results from the Similarity scale which showed that sentences in the Middle pair were less similar than the sentences in the other pair were. Results also indicated that the Simplicity scale was sensitive enough to capture the increased structural complexity attributed to passive, middle and the complex predicative constructions, which are believed to be derived via a move-α operation. Our study indeed presented evidence that the paradigm was also capable of dissociating plausibility from structural complexity, given that results from the Plausibility scale were found to be independent from those from the Simplicity scale.

We observed that the conduction of both analyses of frequencies and of variance was crucial in bringing to light more detailed information regarding the materials being tested. Both analyses combined were also informative in assessing the scope and validity of the measures being collected. For instance, Chi Square results from the comparison between Transitive Animate and Transitive Inanimate indicated that the Simplicity scale, initially proposed to capture structural complexity, is also sensitive to canonicity effects.

In summary, findings from Experiment 2 are in line with those from Di Sciullo et al.'s (2007), and Maia et al.'s (submitted) off-line tasks, in that we also found evidence that middle constructions are more complex than their transitive counterparts. Although

we acknowledge the limitations imposed by the off-line character of the present task, our findings are also in agreement with those from Di Sciullo at al.'s (2007) self-paced reading task, in the sense that results from both tasks suggest that middle constructions are indeed different from passives.

In Experiment 3, we proceed to investigate the online processing of middle constructions. In comparison with the self-paced reading task in Experiment 1, the following study utilizes a more sensitive technique (eye-tracking), as well as improved materials (sentence in the Transitive Animate, Transitive Inanimate, and Middle conditions of the present experiment).

Experiment 3

Method

Participants. One hundred and fourteen undergraduate students at Concordia University participated in the present experiment in return for course credits. They were 92 females and 22 males with ages ranging from 18 to 39 (M = 22.06; SD = 3.50). Twenty-nine of them had corrected-to-normal vision. They were all native speakers of Canadian English, having learned the language before the age of 3. Only thirteen were English monolinguals.

Materials and apparatus

Materials. Materials consisted of 3 lists with 107 items. Each list consisted of 15 experimental sentences, and 92 fillers. Experimental sentences were the same improved ones used in the Transitive Animate, the Transitive Inanimate, and the Middle conditions in Experiment 2, and were distributed in three lists according to a Latin square design. Fillers were comprised of 92 sentences belonging to three unrelated studies (e.g., *The*

tailor felt the elastic during the adjustment, *The host thickened the soup before the dinner*). About 50% of the fillers were odd sentences (*e.g.*, *The moratorium on mining will not twice much longer)*. Items order for each list, separately, was set pseudo randomly so that no sentences of the same condition would immediately follow each other. About 30% of the sentences in each list were followed by a comprehension question. Correctness rate was of about 87%. Because the comprehension questions were used as a mean to keep participants engaged in effectively reading the materials, question-answering latency does not constitute a variable in the present study.

Apparatus. The experiment was programmed using Experiment Builder software (SR Research, 2009). Stimuli were presented in a Viewsonic 19" screen (CRT Graphic series G90fb, 1024 x 768 pixel resolution, 100 Hz refresh rate). Data were collected using an Eyelink II head-mounted eye-tracking device (SR Research Ltd.), set to record monocular pupil movement patterns of the right eye, with sampling rate of 500 Hz per second (for pupil only measurements, average error is of less than 0.5°, and a spatial resolution is of less than 0.01°, or 40" of arc). Participants were sat at 57 cm from the screen, which allowed for 1^0 of visual angle, subtending approximately 4.0 characters.

Procedure. Trials consisted of whole sentences presented-single lined at the center of the screen, using font Courier New. Participants used a gamepad-like button box to go forward to the next trial, and also to answer "yes" or "no" to the eventual comprehension questions. In case participants did not trigger the next trial, a time out of 10 seconds embedded in the program would trigger the next trial to appear.

We chose this paradigm over a moving-window one for a set of reasons. One, we have already explored a moving-window paradigm in Experiment 1. Two, the paradigm

adopted in Experiment 3 allows for a more naturalistic linguistic processing pattern to take place, as it also reduces any potentially unnecessary usage of other cognitive resources, such as attention and working memory. More substantially in an eye-tracking moving-window paradigm—even those which allow for regressions—participants would be required to keep active in their memory not only the linguistic information being processed, but also the spatial coordinates associated with different sentence regions, in the case they want to regress to those areas. Concerning the potential risk of having parafoveal preview as a confound, there seems to be evidence in the eye-tracking literature dedicated to language processing that such phenomenon does not significantly impair the processing of linguistic structural complexity (Clifton et al., 2003)

Each session began with a 9-point grid calibration type. And before each trial was presented, a drift correction was performed. Tracking was recalibrated if necessary. Accuracy in the validations was equal or better than 0.5° of visual angle. Also before each trial, but after drift correction, a fixation cross appeared at the utmost left center coordinates of the screen for 1.5s. After the instruction screen was presented, participants ran seven practice trials. Before the actual experimental trials began, an instructions reinforcement screen was presented. Participants were instructed to read the sentences at a normal pace, and also to pay attention the sentences in order to answer the comprehension questions appearing after some trials (about 30% of the trials). Experimental sessions lasted around 45 minutes. Information regarding fixations, and buttons pressed were recorded for each participant in a separate EDF file. All data files were later manipulated using DataViewer software (2009).

Analyses. We analyzed the same regions as in Experiment 1 (V-1, V, and V+1) across the three sentence types (Transitive Animate, Transitive Inanimate, and Middle), according to measures of initial parse: first fixation duration, , fist pass time, including those measures indicative of reanalysis, that is, go-past time, re-reading duration, and percent regressions (see Rayner, 2009). We also analyzed total fixation duration, which encompasses first and second pass fixations at a region. A fixation occurs when the eye gaze stays still over a period of time (Holmqvist et al., 2011). And the underlying hypothesis is that processing takes place when the eye fixates on a word. Therefore the longer a region needs to be fixated at (regressively or progressively) and the more regressive patterns it elicits, the more difficult to process that region is assumed to be (*e.g.*, Rayner, 2009; Staub & Rayner, 2007). First fixation duration indicates the time span of the first fixation in a region. First pass time is the sum off all fixations in a region before the eye first moves to another region, be it rightwards or leftwards. These measures are assumed to reflect the time course of first parsing processes, and have been found to capture effects of initial processing difficulties, associated with lexical access, and with syntactic complexity (Rayner, 2009; Staub, 2007, 2010; Staub & Rayner, 2007; Traxler & Pickering, 1996; but see Frazier & Rayner, 1982). If priming effects from the first conjunct over the processing of the experimental second conjunct are to be replicated, then we expect to find no delays associated with these measures of initial parsing for the Middle condition. In fact, reduced means for the first fixation durations would indicate whether such priming effects are to occur as early as the stage of lexical recognition.

Moreover, go-past time indicates the amount of time it takes the eye to first move past a region in a progressive manner. This measure is associated with processing difficulties, as it encompasses the total amount of time spent fixating in the region plus any amount of time spent regressing to previous regions. Re-reading duration is the difference between the first pass time and the go-past duration. Hence, this measure indicates the amount of time required to revise previous regions before moving forward to the next region, during first pass. Percent regressions reflects the portion in which the region was regressed into during first pass. These measures allow us to establish (i) which regions elicited the most need for revision, as indicated by greater go-past times, (ii) the complexity of the revision, as indicated by greater re-reading times, and (iii) which regions required most revision, as indicated by percent regressions (Rayner, 2009; Staub & Rayner, 2007). If priming is to occur, then we expect to find smaller go-past times, and re-reading durations, as well as fewer regressions towards regions V-1, and V in the Middles condition.

Results and Discussion

Data cleaning was conducted using Dataviewer software (ver.2009, SR. Research). Regions of interest were set to be automatically recognized, with segmented space set to 7 pixels. Data filter was pre-set to exclude fixations below .50 ms., as fixations with a duration span below such threshold would not reflect cognitive processing (Holmqvist et al., 2011).The fixations trim spanned option was selected. All fixations inferior to 80 ms. but falling in the 1 character margin of a neighbouring fixation were merged into that longer fixation. All fixations inferior to 40 ms., but falling within a 3 characters margin of a neighbouring fixation were merged into that longer

fixation. All subsequent fixations below the 80 ms. threshold were excluded from the analysis.

Even though Holmqvist et al. (2011) suggests a minimal threshold of 50 ms. before any meaningful recognition can take place, we take the theoretical stand that on the top of mere visual recognition; subsequent linguistics processes are to take effect. Thus, following Rayner and Pollatsek (1989), Staub (2007), and Staub, Clifton, and Frazier (2006), we implemented an 80 ms. value for minimal fixation duration in what it is assumed to be a minimal threshold encompassing online language processing. In accordance, we established a maximum fixation duration value of 800 ms. We assume that fixations above that value would reflect cognitive processes hierarchically superior than those pertaining to online language processing alone.

It is worth noting, for consistency's sake, that the different ceiling cut-offs adopted in Experiment 1 are a reflection of the nature of the task. In the self-paced reading task, the index of linguistic processing is the reaction time variable. It is inherent to RT's the fact that they do encompass reading times added to participant's motor responses, i.e. the time devoted to the participant's decision on which button to press, and to command and execute that appropriate response. On the other hand, in an eye-tracking task, eye fixations, as the dependent variable reflecting linguistic processing, are registered with a much reduced margin of error. Thus, appropriately, we set a different, lower maximum cut-off for the eye-tracking data.

Moreover, all sentences with irreparable track loss were excluded from analyses. This accounted for about 10% of the total number of trials. Fixations occurring outside the regions of interest but within a range of 30 pixels above and below on the Y axis were

drift-corrected (A. Johnson, personal communication, March, 2011). This accounted for about 7 % of the remaining trials.

Main effects. A 3 (sentence type: Transitive Animate, Transitive Inanimate, Middle) X 3 (region: V-1, V, V+1) repeated measures ANOVA for the dependent variables was conducted. Table 9 shows means and standard deviations. The analysis revealed that main effects of region all reached statistical significance for most of the first parse measures, as well as for total fixation duration, in both participants' and items' analyses (first fixation duration: $F2(2, 226) = 37.61$, $p < .001$, $\eta_p^2 = .73$, $\varepsilon = 1$; first pass time: $F2(2, 226) = 17.29$, $p < .001$, $\eta_p^2 = .55$, $\varepsilon = 1$; go-past time: $F2(2, 226) = 6.54$, $p = .005$, $\eta_p^2 = .32$, $\varepsilon = .88$; and total fixation duration: $F2(2, 226) = 6.57$, $p = .005$, $\eta_p^2 = .32$, $\varepsilon = .88$). Table 10 shows results from the participants' analysis. For the two variables which directly reflect revision patterns (re-reading duration and percent regress), results either did not reach statistical significance or exhibited small effect size. Results for re-reading duration were only significant in the participants' analysis, but with a small effect size ($\eta_p^2 = .03$; $F2(2, 226) = 1.78$, $p = .187$, $\eta_p^2 = .11$, $\varepsilon = .34$); and percent regress effects were not significantly different across regions ($F2(2, 226) = 2.78$, $p = .079$, $\eta_p^2 = .17$, $\varepsilon = .50$), thus suggesting that revision patterns were independent of region. Results also indicated that the three syntactic categories noun, the verb, and the adverb, corresponding respectively to the lexical items in the regions V-1, V, and V+1, overall triggered different processing loads, and that each respective processing complexity was computed at the first pass, as evidenced by the null effect of region over the two revision measures in contrast with the significant effects found over the other variables.

Table 9

Experiment 3 Mean RTs in milliseconds for each Dependent Variable across Sentence Regions per Sentence Type (by participants)

	Region					
	V-1		V		V+1	
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*
First Fixation duration						
Middle	171.37	56.62	228.17	59.80	230.13	57.38
Transitive Inanimate	180.78	60.33	234.92	45.76	231.73	62.94
Transitive Animate	184.53	59.55	242.08	55.80	237.67	65.68
Total[a]	178.89	4.07[b]	235.06	4.18[b]	233.18	4.80[b]
First pass time						
Middle	193.55	75.63	268.08	88.74	281.93	92.41
Transitive Inanimate	219.17	95.91	283.51	77.23	271.10	83.30
Transitive Animate	206.65	72.58	278.05	78.76	281.59	93.02
Total[a]	206.46	5.68[b]	276.55	6.44[b]	278.21	6.84[b]
Go-past time						
Middle	240.71	128.01	303.70	117.87	329.01	117.25
Transitive Inanimate	276.90	158.66	327.08	113.82	319.48	126.62
Transitive Animate	280.65	153.37	316.87	106.88	319.06	117.25
Total[a]	266.09	9.48[b]	315.88	7.94[b]	322.52	8.75[b]

	Region					
	V-1		V		V+1	
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*
Re-reading duration						
Middle	11.10	27.12	12.54	25.87	18.74	40.10
Transitive Inanimate	11.40	26.44	17.81	32.08	21.37	39.08
Transitive Animate	13.17	29.01	17.79	40.78	14.66	27.22
Total[a]	11.89	1.75 [b]	16.05	1.80 [b]	18.25	2.13 [b]
Percent regress [c]						
Middle	27.59	27.35	21.69	23.00	20.96	24.27
Transitive Inanimate	24.89	25.97	25.32	25.32	24.21	28.51
Transitive Animate	26.92	25.17	20.51	23.74	24.14	29.97
Total[a]	26.47	1.74 [b]	22.51	1.63 [b]	23.10	1.80 [b]
Total fixation duration						
Middle	305.35	161.95	358.60	126.83	374.83	121.98
Transitive Inanimate	313.69	128.69	382.85	122.03	375.24	123.27
Transitive Animate	313.12	118.65	367.03	109.65	377.87	150.26
Total[a]	310.72	10.24 [b]	369.49	8.88[b]	375.97	9.35 [b]

Note. [a]Means per region, collapsed across sentence types. [b] Standard Error. [c] Mean Percentages.

Table 10

Analysis of Variance Results for Main Effects of Region, Sentence Type, and Interactions for each Dependent Variable in Experiment 3 (by participants).

Variable	df	*F*	*p*	η_p^2	ε
Region					
First fixation duration	(2, 226)	123.95	.000	.52	1
First pass time	(2, 226)	95.42	.000	.46	1
Go-past time	(2, 226)	25.07	.000	.18	1
Re-reading duration	(2, 226)	3.51	.032	.03	.65
Percent regress	(2, 226)	2.39	.094	.02	.48
Total fixation duration	(2, 226)	30.48	.000	.21	1
Sentence type					
First fixation duration	(2, 226)	5.73	.004	.05	.86
First pass time	(2, 226)	2.77	.065	.02	.54
Go-past time	(2, 226)	2.82	.062	.02	.55
Re-reading duration	(2, 226)	0.73	.484	.01	.17
Percent regress	(2, 226)	0.33	.697	.00	.10
Total duration	(2, 226)	0.90	.407	.01	.21
Interaction					
First fixation duration	(4, 452)	0.32	.862	.00	.12
First pass time	(4, 452)	2.52	.041	.02	.72
Go-past time	(4, 452)	1.81	.126	.02	.55
Re-reading duration	(4, 452)	0.92	.453	.01	.29
Percent regress	(4, 452)	1.08	.367	.01	.34
Total duration	(4, 452)	0.53	.699	.05	.169

Main effects for sentence type and for interactions were all statistically nonsignificant for both participants' and items' analyses, with only two exceptions. Even though statistically significant in the participants analysis, effects of sentence type over first fixation duration accounted for only about 5% of the variance (*F1* η_p^2 = .05; *F2*(2, 226) = 2.78, *p = .079,* η_p^2 = .17, ε = .50), and the interaction at the first pass time variable accounted for merely 2% of the variance (*F1* η_p^2 = .02; *F2*(2, 226) = .94, *p = .404,* η_p^2 = .06, ε = .20). These results suggest that overall processing patterns were independent of sentence type, and that no meaningful interaction between region and sentence type was found at any of the dependent variables.

Analyses of sentence regions. The analysis of the differences between means, collapsed across sentence type, for each of the sentence regions was conducted to assess possible effects derived from semantic priming between the semantically related NP1 (e.g., *cook*) and NP2 (e.g., *chef*), as well as those derived from syntactic priming. Pairwise comparisons revealed that V (verb) and V+1 (adverb) exhibited similar processing patterns, having none of the pairwise comparisons reaching statistical significance for any of the dependant variables at both items' and participants' analyses. The comparisons also revealed that main effects of region were probably influenced by the V-1 region means. Results showed that the NP2 exhibited overall faster (i) first fixation duration (V-1 vs. V: *t2*(14)= -8.38, *p* < *.001*, *SE* = 6.69, V-1 vs. V+1: *t2*(14)= -6.14, *p* < *.001*, *SE* = 8.73), (ii) first pass time (V-1 vs. V: *t2*(14)= -4.74, *p = .001*, *SE* = 14.86, V-1 vs. V+1: *t2*(14)= -4.84, p = *.001*, *SE* = 14.50), (ii) go-past time (V-1 vs. V: *t2*(14)= -2.80, *p = .043*, *SE* = 17.86, V-1 vs. V+1: *t2*(14)= -3.36, *p =.014*, *SE* = 16.75), and (iv) total fixation duration (V-1 vs. V: *t2*(14)= -3.00, *p = .029*, *SE* = 19.34, V-1 vs.

V+1: *t2*(14)= -2.88, *p* =*.036*, *SE* = 20.775) than verbs and adverbs. Table 11 shows results for the pairwise comparisons in the participants' analyses.

Table 11

Comparisons of Regions Means Collapsed across Sentence Type (by participants)

Comparison	df	*t*	*p*	SE
First fixation duration				
V-1 (NP2) X V (verb)	113	-13.94	.000	4.03
V-1 (NP2) X V+1 (adv)	113	-12.04	.000	4.51
V (verb) X V+1 (adv)	113	0.53	1	3.56
First pass time				
V-1 (NP2) X V (verb)	113	-12.03	.000	5.83
V-1 (NP2) X V+1 (adv)	113	-11.38	.000	6.30
V (verb) X V+1 (adv)	113	-0.29	1	5.64
Go-past time				
V-1 (NP2) X V (verb)	113	-5.61	.000	8.87
V-1 (NP2) X V+1 (adv)	113	-5.98	.000	9.43
V (verb) X V+1 (adv)	113	-0.86	1	7.75
Re-reading duration				
V-1 (NP2) X V (verb)	113	-1.80	.222	2.31
V-1 (NP2) X V+1 (adv)	113	-2.45	.047	2.60
V (verb) X V+1 (adv)	113	-0.92	1	2.41
Percent regress				
V-1 (NP2) X V (verb)	113	2.15	1	1.84
V-1 (NP2) X V+1 (adv)	113	1.68	.102	2.01
V (verb) X V+1 (adv)	113	-0.30	.289	2

Comparison	df	*t*	*p*	SE
Total fixation duration				
V-1 (NP2) X V (verb)	113	-7.08	.000	8.30
V-1 (NP2) X V+1 (adv)	113	-6.15	.000	10.60
V (verb) X V+1 (adv)	113	-0.76	1	8.55

Moreover, pairwise comparisons revealed that (i) re-reading times were about the same for all regions, with only the comparison V-1 (NP2) vs. V+1 (adverb) reaching statistical significance at the participants analysis (p = .049), and that (ii) all regions displayed the same probability of being regressed into, also with the comparison V-1 (NP2) vs. V+1 (adverb) marginally significant at the items' analysis ($t2(14)$= 2.78, p =.*044*, SE = 1.67).

In summary, results of the region variable indicated that (i) NP2 was overall read faster than the verb, and faster than the adverb, and that (2) NP2 did not elicit greater reanalysis nor was it regressed into more often than verbs and adverbs. Due to the inexistence of a control variable in which the nouns were not semantically related, we cannot confidently assert these results to be congruent with a priming effect between the semantically related NP1 and NP2. Nevertheless, we assume that results from the regions analyses also do not refute effects of semantic priming, and also that they are congruent with the assumption that verbs are more demanding to process than nouns. In the present Experiment 3, the differences found between the NP2 and the adverb can no longer be credited to an end-of-sentence wrap-up effect because the V+1(adverb) region was removed from a sentence final position. It could be the case that the absence of the post-verbal internal complement in the Transitive Animate, and Transitive Inanimate conditions inflated this region's means. Furthermore, these tests also indicated that, disregarding the greater numerical value for the mean percentage of regressions towards the NP2, no overall significant difference was found for that variable across regions. Again, it is possible that nonsignificant results for percent regress could be masking an effect derived from sentence type.

Analyses of sentence type per sentence region. Pairwise comparisons between the three sentence types at each region were conducted for all the dependent variables. Results were consistent in demonstrating a similar reading pattern for the three sentence types. See Table 12 for results in the participants' analysis.

At the V-1 (NP2) region, the only t-test reaching statistical significance was the comparison between Middle and Transitive Inanimate first pass reading times (p = .021). Even though results indicated that the *instrument* NP2 in Transitive Inanimate (e.g., *knife*) took longer to be first parsed than the *theme* NP2 in Middle (e.g., *bread*), the effect was small (d = -0.30), and restricted to the participants' analysis.

At the V region, the difference between Middle and Transitive Inanimate first pass times reached statistical significance in the items' analysis (the comparison was marginally significant in the participants' analysis, p = .079), with the verb in Transitive Inanimate being read significantly longer than the verb in the Middle condition ($t2$(14) = -2.72, p =.049, d = -0.44). The comparisons also indicated that Middle first fixation durations were shorter than Transitive Animate's, but the effect was of a small magnitude and restricted to the participants' analysis and (p =.013, d = -0.24).

At the V+1 (adverb) position, no differences between sentence types were found.

Table 12

Pairwise Comparisons between Sentence Types across Sentence Regions per Dependent Variable (by participants)

Comparison	V-1				V				V+1			
	df	*t*	*p*	*SE*	df	*t*	*p*	*SE*	df	*t*	*p*	*SE*
First fixation duration												
Mid. X Tr. Ani.	113	-1.97	.154	6.68	113	-2.90	.013	4.80	113	-1.46	.441	5.17
Mid. X Tr. Ina.	113	-1.44	.456	6.53	113	-1.32	.349	5.12	113	-0.27	1	6.01
Tr. Ani. X Tr. Ina.	113	0.61	1	6.12	113	1.45	.446	4.92	113	1.02	.937	5.85
First pass time												
Mid. X Tr. Ani.	113	-1.55	.371	8.45	113	-1.29	.601	7.73	113	0.04	1	8.01
Mid. X Tr. Ina.	113	-2.75	.021	9.31	113	-2.27	.076	6.81	113	1.19	.711	9.11
Tr. Ani. X Tr. Ina.	113	-1.38	.512	9.08	113	-0.79	1	6.95	113	1.28	.609	8.19
Go-past time												
Mid. X Tr. Ani.	113	-2.20	.089	18.15	113	-1.07	.856	12.27	113	0.76	1	13.06
Mid. X Tr. Ina.	113	-2.13	.105	16.97	113	-2.04	.131	11.46	113	0.66	1	14.39
Tr. Ani. X Tr. Ina.	113	0.22	1	16.95	113	-0.81	1	12.57	113	-0.03	1	12.75

Comparison	V-1				V				V+1			
	df	*t*	*p*	*SE*	df	*t*	*p*	*SE*	df	*t*	*p*	*SE*
Re-reading duration												
Mid. X Tr. Ani.	113	-0.56	1	3.69	113	-1.13	.788	4.66	113	0.91	1	4.48
Mid. X Tr. Ina.	113	-0.10	1	2.93	113	-1.40	.495	3.77	113	-0.56	1	4.71
Tr. Ani. X Tr. Ina.	113	0.55	1	3.21	113	-0.01	1	4.85	113	-1.55	.373	4.34
Percent regress												
Mid. X Tr. Ani.	113	0.22	1	3.09	113	0.41	1	2.86	113	-1.10	.817	2.88
Mid. X Tr. Ina.	113	0.92	1	2.95	113	-1.23	.665	2.95	113	-0.99	.970	3.28
Tr. Ani. X Tr. Ina.	113	0.69	1	2.96	113	-1.64	.314	2.94	113	-0.02	1	3.53
Total fixation duration												
Mid. X Tr. Ani.	113	-0.52	1	14.89	113	-0.72	1	11.67	113	-0.22	1	13.80
Mid. X Tr. Ina.	113	-0.63	1	13.14	113	-2.01	.141	12.07	113	-0.03	1	13.74
Tr. Ani. X Tr. Ina.	113	-0.05	1	12.58	113	-1.34	.549	11.81	113	0.18	1	14.83

In summary, results showed that the Middle condition does not display any greater processing difficulties compared against the other two transitive conditions. This overall finding supports our argument for a facilitatory semantic-syntactic priming effect over the processing of putatively more complex middle constructions, effect which we claim derived not only from the semantically related items in subject position (e.g., *croissant*, *bread*), but also from the similar unaccusative syntactic structure presented by the adjectival predicative construction in the first conjunct (e.g., *This croissant is flaky*). Because we were able to substantiate such priming effects, our results could be entertained as evidence that middle constructions are syntactically represented as unaccusatives.

Experiment 3 was designed to assess whether different reading pattern between transitives and Middle conditions would emerge with the implementation of the more sensitive eye-tracking technique. Essentially, in the present experiment, (i) participants were given the chance to revise previous regions of the experimental sentences, and (ii) materials were corrected for a possible confound affecting the data collected from the critical post-verbal region (V+1), which was removed from an end-of-sentence position. Results indicated a tendency for participants to process the Middle sentences with greater ease, with small to medium facilitatory effects at the NP2 and at the verb. The fact that those differences were not captured by measures that directly indicate revision (re-reading duration and percent regress), as well as go-past time, suggests that any potential difficulties associated with the processing of the transitive conditions were immediately and locally resolved.

Once again, our findings contrast with those from Di Sciullo at al. (2007), which found greater complexity associated with the online processing of middle constructions, when compared against their transitive counterparts. In fact, Experiment 3 corroborates the interpretation offered for the results in Experiment 1, in which we pose that the unaccusative adjectival predicative construction in the first conjunct facilitated the processing of the middle constructions. In Experiment 1, results showed that the NP2 (e.g., *bread*) in the Middle condition was read faster than the NP2 in each Transitive Animate (e.g., *chef*), and Transitive Inanimate (e.g., *knife*) conditions. In the present experiment, the differences related to the processing of the NP2 were consistent in both participants' and items' analyses, but they were restricted to the comparison between the Middle and the Transitive Inanimate first pass reading times, measure which encompasses the sum of all fixations in that area before the eye moves forwards or backwards in the sentence. Results from Experiment 3 showed a small effect at the verb region (effect which was not captured in Experiment 1), revealing that participants' first fixation on the verb took longer in the Transitive Animate condition than in the Middle condition.

Although it is not clear in the literature whether first fixation duration captures strictly immediate lexical phenomena, or whether it is also sensitive to syntactic complexity (see Staub & Rayner, 2007), we entertain the possibility that the greater first fixation durations found at the verb region in the Transitive Animate condition could be accounted for by an *in situ* readjustment of the verb's argument structure, from the primed and less canonical unaccusative ($[NP_i[_{VP}\ V\ t_i]]$) to the more canonical transitive structure ($[NP[_{VP}\ [V\ NP]]]$), given the agentive nature of Transitive Animate's NP2 (e.g.,

chef). The fact that there were no differences between the more canonical NP2 *agent* in Transitive Animate (e.g., *chef*) and the less canonical NP2 *theme* in Middle (e.g., *bread*) could be in part explained by the semantic priming effect derived from semantically related NP1 in the first conjunct (*e.g., cook* in Transitive Animate, and *croissant* in Middle). However, we cannot discard the possibility that the greater first fixation durations found at the Transitive Animate verb region could be reflecting a spill-over effect derived from the cost of reconciling the *agent* NP2, canonically the verb's actual external argument, with the primed unaccusative structure, which displays a complement in subject position. In this case, readjustment would have taken place already at encountering the agentive animate NP2. Either way, both accounts entail that the parser would have to have immediate access to not only semantic information, but also to information pertaining a verb argument structure.

Results also show a tendency for a greater processing cost associated with Transitive Inanimate against Middle, also at an early stage of parsing. The greater first pass durations found at the NP2 (*e.g., knife*), and at the verb (e.g., *cuts*) in the Transitive Inanimate condition could also be explained by the need to readjust the primed structure. We argue that because Transitive Animate's animate *agent* NP2 (*e.g., chef)* constitutes the most canonical external argument (Spec) for the verb (Keiser & Trueswell, 2004; Sekerina, 2003), the parser would have had sufficient information to shift to a transitive structure. Concerning Middle's inanimate *theme* NP2 (*e.g., bread*), we argue it to be canonically associated with the complement position of an unaccusative structure (Keiser & Trueswell, 2004; Sekerina, 2003). This would have facilitated the processing of the subject in Middle sentences. On the other hand, we argue that the inanimate *instrument*

NP2 in Transitive Inanimate (e.g., *knife*) can be more freely associated with both specifier and complement positions. As a consequence, it is possible that the parser was unable to commit with either a transitive or an unaccusative argument structure, carrying the ambiguity up to verb region where the impasse was ultimately resolved.

Another possibility is that the parser resolves the ambiguity at the NP2 region, and that the Transitive Inanimate's greater reading times at the verb region are due to a spill-over effect. This account is more congruent with syntax first models (e.g., Fodor J.D. & Frazier, 1980; Frazier & Clifton, 1989, Frazier & Fodor J.D., 1978). Furthermore, these results are also in agreement with the proposal that the parser is to be able to compute information pertaining to lexical items' semantics as well as information pertaining to verb argument structure.

Typicality of the noun in subject position. Experiment 3 results are suggestive of a preference for a theme in subject position. Typicality effects of the noun in subject position have been found to correlate to animacy features, in the sense that inanimate nouns, contrarily to animate nouns, would trigger the expectation for a more complex structure (Clifton et al. 2003; Frazier et al. 2000; Philipp, Bornkessel-Schlesewsky, Bisang, & Schlesewsky, 2008; Traxler, Morris, & Seely, 2002). We propose that those effects extend beyond animacy, and, consequently, that more detailed semantic information is to be available to the parser during online processing. If typicality effects were only a matter of animacy, then no differences would have been found between Middle and Transitive Inanimate conditions at the V-1 region (NP2). Furthermore, there is the fact that animate nouns can pose as the argument of middles, as demonstrated by

Keyser and Roeper's (1984) famous example of middle construction *Bureaucrats bribe easily*.

Indeed, evidence for typicality effects related to the kind of argument structure a verb takes have been detected by measures as early as first fixation durations. This is the case of Traxler and Pickering's (1996) study, which compared typical against non-typical NPs as the subjects of sentences such as *That's the **pistol/garage** with which the heartless killer shot*…. They found greater processing load, as evidenced by greater first fixation and first pass durations at the region containing the verb (*shot*) when the NP was typical (*pistol*) than when it was the less typical (*garage*) (see also Pickering & Traxler, 2003; Staub, 2007). These findings showing that semantic information influences parsing strategies further support our account given of the different reading patterns associated with the typicality of the NP in subject position as obtained in the present experiment.

Argument structure information during online processing. There is evidence suggesting that argument structure information can influence online processing (van Gompel, Pickering, & Traxler (2001), but Adams, Clifton, , & Mitchell, 1998). Staub (2007), for instance, compared three verb types: ergative verbs (e.g., *depart*), and transitive verbs of NP- (e.g., *call*), and PP-preference (e.g., *spoke*) as their internal argument, across two sentence types: object-shifted structures with embedded relative clauses (e.g., *The gadget that the manager called occasionally about after the accident still didn't work*, and *The props that the actor spoke briefly about to the acting coach were ugly*), and their non-shifted counterparts (e.g., *The manager called occasionally*

about the gadget, and *The tor spoke briefly about the props to the acting coach*)[14]. The temporal ambiguity relied on the fact that, in the object-shifted condition, the NP subject of the relative clause (e.g., *manager*), would be initially interpreted as NP object of the main clause. That is to say that the initial interpretation of such constructions would be similar to that in the non-shifted condition: the manager called the gadget. Most important to us, are the results from the ergative condition, which showed no extra processing load associated with the object-shifted condition (e.g., *The airport that the ambassador departed rapidly from during the unrest was closed to most traffic)*, as evidenced by the nonsignificant difference between sentence types in the ergative verb region, in comparison to the greater first fixation durations at the verb region which were found in the object-shifted condition for both NP-, and PP-preference transitive verb types. The evidence that in the ergative condition the parser did not build the relative NP subject as the main clause NP object (that is, an initial interpretation such as *the ambassador departed the airport* was never built) suggests that the parser indeed was not expecting to find an argument in post verbal position, in agreement with an unaccusative argument structure. Such findings corroborated the hypothesis that subcategorization information is available during online processing.

It is important to note that our first and second conjuncts were not composed by subordination, but by coordination, formation which constitutes a less complex structure (see Frazier et al., 2000). Also noteworthy is the fact that, separating the first conjunct

[14] Example of material in the ergative sentence type condition:

(i) a. Object-shifted condition: The airport that the ambassador departed rapidly from during the unrest was closed to most traffic.

b. Non-shifted condition: The ambassador departed rapidly from during the unrest.

from the first target region (NP2), there were the conjunction (*but*) and the determiner (*this*). We assume that these words acted as a buffer, which would have absorbed any spill-over effects derived from wrapping up the first conjunct, as well as any possible effects of interclausal integration (see Staub, 2007).

At the same time that our account of the results does not preclude a null effect at this adverb region, it also does not exclude one. We speculate that any potential complexity derived from the absence of the complement in the transitive conditions could have been manifested at a later point in the sentence. Such notion becomes particularly feasible if we entertain that adverbs can stand in between the verb and its complement, as exemplified by the grammatical sentence *John paints easily this wall,* or in *"After the dog scratched pathetically the veterinarian…"* (Adams et al., 1998). Moreover, because word length and familiarity in all three regions (V-1, V, V+1) did not vary across sentence type, we do not take the null results found at the V+1 region as a liability to our interpretation of the data.

In view of the fact that our results corroborate the notion that not only semantic information but also information pertaining to verb argument structure is to be computed in real time, together with syntactic processes, we suggest that the present experiment (as well as Experiment 1) is evidence that the parser is open to argument structure information at early stages of parsing. Moreover, if argument structure information is operationalized in real time, then the greater processing times found at the verb region in the transitive conditions could be interpreted as evidence of an A-shift, and not simply due to revisions on the syntactic structure. Hence, it makes it more plausible the proposal that argument structures are to be flexible, and that the argument structure attributed to a

verb can be ultimately dependent on a given association between the verb and the arguments presented in the utterance (see Di Sciullo, 2005).

General Discussion

In the present study, we set out to investigate whether the evidence presented in the psycholinguistic literature regarding the comprehension of middle constructions, that is, the intransitive counterpart of a diathesis alternation undergone by canonically transitive verbs, would hold upon more criterious scrutinization. Di Sciullo et al. (2007) and Maia et al. (submitted) provided indication that middle constructions are more complex to process than their transitive counterpart. But, as we have found, both studies failed to provide supporting results. Results from our three experiments were consistent with the conclusion that the processing of middles, when compared against their transitive animate and inanimate counterparts, incurred no extra processing costs.

Our materials consisted of two clauses placed in a contrastive coordinative relation, with the first and second conjuncts, respectively an adjectival predicative clause (e.g., *That croissant is flaky*) and the experimental sentences (e.g., *this bread cuts neatly*), linked by the connector *but*. We argue that semantic-syntactic information derived from the first conjunct influenced the parsing strategies employed on the processing of the second conjunct. Such effect seemed to have overridden canonicity effects, which in turn would have favoured the processing of the transitive conditions. We claim that the unaccusative structures in the first conjunct ultimately facilitated the processing of middles. The present findings touch on three distinct but interrelated research topics: linguistic approaches on middle formation, semantic-syntactic priming, and argument

structure computation during online processing. We discuss our findings in the light of these three topics.

Syntactic Derivation of Middles

The present psycholinguistic study on middle constructions was motivated by the long standing debate in the linguistic literature concerning the properties and the ontogenesis of middles (Ackema & Schoorlemmer, 1994, 1995, 2003; Bowers, 2002; Di Sciullo, 2005; Fagan, 1989, 1992; Hale & Keyser, 2002; Hoekestra & Roberts, 1993; Kemmer, 1993; Keyser & Roeper, 1984; Levin, 1993; Lekakou, 2005; Marelj, 2004; Rapoport, 1999; Steinbach, 2002; Stroik, 1992, 1995, 1999, 2006; Zribi-Hertz, 1993; among others). As we have a shown, there has been no consensus in linguistics, with many proposals stumbling into internal inconsistencies, external incoherencies, or failing to survive evolutions in the linguistic theory. However, it is possible to establish two major strands: (i) one proposing that middles project to syntax a transitive structure, and (ii) one proposing that middles project to syntax an unaccusative structure. If middle verbs project to syntax an unaccusative structure, then their processing pattern should be similar to that of ergatives (e.g., *The vase broke*). Friedmann et al. (2008) found supporting evidence for the unaccusative hypothesis (Burzio, 1986, Perlmutter, 1978). This hypothesis states that ergative verbs only select a complement and that these verbs are unaccusatives (i.e., they fail to assign accusative case to their complement). Friedmann and colleagues assessed whether intransitives constructions with alternating ergative verbs (e.g., *The vase broke)*, non-alternating ergative verbs (e.g., *John fell)*, and unergatives verbs [e.g., *John cried*) are processed the same way. They found priming effects in post-verbal positions for words semantically related to the NP subject, but only

in ergative constructions. Assuming that priming would indicate reactivation of the *trace* left by the moved argument, Friedmann and colleagues interpreted the results as supporting evidence for the unaccusative hypothesis, given that results were coherent with the proposal that ergative constructions are syntactically derived via the demotion of the argument from complement to subject Spec position, as predicted by the hypothesis.

The copula verb *be,* present in the adjectival predicatives in our materials' first conjunct, is also argued to be unaccusative, projecting to syntax a one-place argument structure, where its internal argument is moved to Spec position (Alexiadou et al., 2007). Our online and off-line results consistently showed that the middle condition, contrary to what previous studies on object-shifted constructions have found (e.g., Miyamoto & Takahashi, 2000; Pickering & Barry, 1991; Sekerina, 2003), did not display any associated extra processing cost. In fact, our results showed that middles were easier to process than their transitive counterparts. We claim such facilitation in the Middle condition took place because both conjuncts, besides displaying semantically related NPs in subject position, shared the same syntactic structure (we discuss semantic-syntactic priming effects below). As such, we suggest our findings can serve as evidence that middle verbs are syntactically derived as unaccusatives.

The notion that middle verbs can behave syntactically as unaccusatives is especially supported by Experiment 3 results at the subject and verb regions, which showed greater difficulty associated with the processing of the transitive conditions. We claim that such greater processing cost is associated with the need to shift the primed unaccusative structure being built to a transitive one. However, we cannot rule out the possibility that middle verbs do project to syntax a transitive structure. It is also possible

that the unaccusative structure (object-verb [OV]) in the first conjunct facilitated the processing of an OVS (object-verb-subject) transitive structure. The fact that no effects were found at the adverb could be interpreted as evidence for a greater processing cost spread across all three conditions: whereas it was due to the missing internal argument in the transitive conditions, it was due to the missing external argument in the middle condition. On the other hand, because we did not analyze any regions past the adverb, we also cannot rule out the possibility of a spill-over effect associated with the missing internal argument in the transitive conditions, scenario which strengthens our claim. In addition, as shown in Experiment 3 results section, null effects at this region have also been found in the literature contrasting unaccusatives and transitive constructions while still pointing to the reality of their different argument structures (e.g., Staub, 2007).

Friedmann et al. (2008) related a spin off derived from the alternating ergative condition, in which priming effects were not consistently found across all the verbs in this condition. Friedmann and colleagues entertain that because those verbs, much like middle verbs, are of ambiguous transitivity (i.e., they can be realized in both transitive and intransitive constructions) the parser, in the absence of cues as to whether the NP in subject position is to be the internal argument, would initially parse the structure according to the more canonical transitive counterpart (see Shapiro, Zurif, & Grimshaw (1987). We entertain that elements which would cue for verb transitivity can derive, for instance, from the language's morphology (e.g., the verbal infixes in Karajá [–*i*– for unaccusative, and –a– for transitive]), or from clausal context, as shown in the present study. But based on previous evidence suggesting that the parser is opaque to lexical-semantic information when building the syntax of a sentence (Fodor, 1983; Hickok et al.,

1992; Swinney, 1991), Friedmann et al. pose that what they call "sentence proper" (i.e., semantic appropriateness or plausibility of lexical items to fit into a certain verb structure) should not play a role in influencing parsing strategies. (We address below the proposal that online parsing is to ignore information other than syntactic). However, a closer look at their materials revealed that the NPs used as subject in the alternating ergative condition varied in animacy and in agentivity (e.g., *kid*, *racket*, and *pie*, which are typically *agent*, *instrument*, and *theme*). It is possible that lexical-semantic information deriving from the NP in subject position could have influenced the parsing strategies used to process their ergative sentences, and consequently rendered the results for that condition inconsistent.

Rapoport (1999) proposes that the differences between middle and ergative constructions rely solely on the instrument/manner (I/M) component of the verb participating in such constructions. This proposal can be appropriated under the notion that differences between verb types, such as the I/M component contrast between middles and ergatives, might be independent from the type of argument structure verbs employ (Di Sciullo, 2005). Such notion can be further supported by our findings, which suggest that both middle and adjectival predicative constructions employ an unaccusative argument structure, even though middle verbs are inherently different from the predicate copula verb *be.*

Moreover, future research should focus on the direct comparison between middles and ergatives. If utilizing the same paradigm here implemented, the analysis of the regions V-1, V, and V+1 of the adjectival predicative clause in the first conjunct would serve as a comparable baseline against effects of priming. Additionally, investigation of

regions subsequent to V+1 in the second conjunct would provide further insight concerning the derivation of middles.

In summary, our study sheds light on the debate regarding the nature of middle formation. Even though our findings do not permit any assertion concerning whether or how information related to the verb's alternating transitivity is ultimately represented in our language system, we presented evidence suggesting that transitive verbs undergoing middle construction can employ an unaccusative argument structure. Moreover, our study also presented evidence that argument structure information is articulated along with syntactic and lexical-semantic information in real time processing, thus opening the possibility that argument structures can be adjusted in real time.

Semantic-Syntactic Priming and the Online Processing of Argument Structure Information

Recent studies have investigated the phenomenon of syntactic priming in comprehension. For instance, Frazier et al. (2000) found evidence that the processing of coordinated items within the same clausal structure were facilitated once they belonged to the same category as the elements presented in the preceding conjunct. Their results showed that, not only adverbs would reduce reading times for adverbs in construction such as *John walked **slowly** and **carefully***, but also that a previous NP-AdjP would facilitate the processing of a subsequent NP-AdjP in construction such as *Hilda noticed a **strange man** and a **tall woman***…. These effects extrapolate the barrier of simple lexical priming, given that they are also manifested over phrase internal complexity. This suggests that the phenomenon is related to the parallelism inherent of coordinated items, thus justifying the denomination of syntactic parallelism. Frazier et al. (2000) elaborate

that the information dictating that coordination is to assume listing of alike items is derived from normative rules, and as such should be treated as extraneous from the grammar. Moreover, the paper argues against the idea of a priming effect associate with the previously used syntactic template, and suggests that the facilitation is simply derived from the fact predicable structures, as it is the case of coordinated items, are generally read with greater ease (p. 360-361).

Although syntactic parallelism effects were proposed to be restricted to coordination because they traditionally imply repetition of alike items, similar facilitatory effects were later observed in subordination (e.g., *A boss* ***who was demanding*** *said that a worker* ***who was lazy*** *did not do the job properly;* Sturt, Keller, & Dubey, 2010 [see also Knoeferle & Crocker, 2009]). Indeed, in an ERP study, Ledoux et al. (2007) examined the effects of syntactic parallelism by contrasting temporally ambiguous sentences with reduced RC (relative clauses) (e.g., *The manager* ***proposed*** *by the directors …*) against simplex transitive constructions (e.g., *The speaker* ***proposed*** *the solution…*). The ambiguity in the RC condition relied on the fact that the deverbal adjective (e.g., *proposed*) is homonymic to the participle from of the verb (e.g., *proposed*). The ambiguity was to be sustained up until the homonymic participle, which would be parsed preferentially as a verb. This would render the initial parsing of the sentence in the RC condition identical to that of the simplex condition up until the homonymic verb region (*proposed*). Ledoux et al. (2007) controlled for whether prime and target sentences were of the same condition, that is, whether, for instance, the target RC would be primed by another sentence in the RC or by a sentence in the simplex condition. Results showed a facilitatory effect in the RC condition, but not in the simplex condition, by means of a

lower positivity (P600) at the post-verbal region (region where revision would be elicited in the RC condition) when sentences in both conditions were primed by a sentence in the RC condition. On the other hand, no such effect was found when RC sentences were preceded by sentences in the simplex condition, despite the repetition of the homonymic verb. Reduction in the P600 (measure which is positively associated with syntactic processing cost) was interpreted as evidence for syntactic parallelism derived from the prime sentence, which in turn had been presented at least two distractors before the target sentence[15]. Evidence for syntactic parallelism effects outside the realm of coordination raises the question as to whether the phenomena is in fact an instance of priming, as proposed by Sturt et al. (2010).

Besides syntactic priming, semantically related lexical items have been found to produce a facilitatory effect. For instance, Knoeferle and Crocker (2009), and Frazier et al. (2000) observed that semantically primed items would activate syntactic priming effects (for instance: object *fencer*, object *wrestler;* NP *man* and NP *woman*). However, whether syntactic priming is dependent of verb repetition, it is still debatable (see: Arai et al., 2007; but Knoeferle & Crocker, 2009; Sturt et al. 2010).

Indeed, the interplay between semantic and syntactic priming is unclear. For instance, Ledoux et al. (2007) utilized verb repetition. They observed that a reduction of the negativity associated with semantic processing complexity (N400) was found at the primed verb when prime and target sentences were dissimilar, thus suggesting a facilitation of lexical repetition. However, this effect was restricted to the simplex

[15] Results such as these lead Knoeferle & Crocker (2009) to posit a difference between priming and parallelism. Parallelism is bound to coordination; priming does not obey the same restriction. On the other hand, Sturt et al. (2010) assume both phenomena as instances of priming: "parallelism preference is an instance of structural priming.".(p. 348).

condition. Instead, when prime and target sentences were in the RC condition, results at the verb region of the target RC showed a decreased P600-800, a measure posited to be related with memory retrieval of lexical items. Ledoux et al. (2007) propose that the P600-800 facilitatory effect could be interpreted as an indication that syntactic priming is in fact due to verb lexical-syntactic information being kept salient in the working memory.

If syntactic priming is associated with reapplication of the whole verb argument structure, then an account of how parsing strategies incorporate such information needs to be offered. Concerning parsing models, object-shifted constructions (object-verb [-subject], OV[S]); such as the passive *John was interrupted [by Mary]*, or the middle *This wall paints smoothly*) can be approached under two contrasting views. On one hand, the processing of object-shifted constructions is entertained by constraint-based models as a matter of canonicity, that is, statistical frequencies would determine canonicity, and canonicity would constrain parsing preferences (Mitchell, Cuetos, Corley, & Brysbaert, 1995; MacDonald, Pearlmutter, & Seidenberg ,1994; Trueswell, 1996; Trueswell Tanenhaus & Garnsey, 1994). Because those models generally assume parallel processing, that is, possible parallel structures to be active during online processing, OV sentences would be costlier to parse because they require a switch from the default SOV canonical structure to the less salient OV(S) competitor. On the other hand, according to syntax-first parsing models, such as that proposed by the garden-path theory (Frazier & Fodor J.D., 1978), the notion that parsing should be fast and efficient is formalized by principles such as the Minimal Attachment (MA), or DeVincenzi's (1991) Minimal Chain Principle (MCP). These principles postulate that nearby elements should be

preferentially attached locally, and that no extra chains, or syntactic nodes, should be initially postulated, given that the simplest syntactic structure should be the default. Because those principles assume a serial incremental processing, generally in terms of a two-stages parser (syntax first, semantics later), OV sentences would be costlier because the NP in subject position needs to be reattached, this time as the verb complement. This revision of the structure is prompted by the mismatch between the current semantic information and the default simplest syntactic structure already being built (e.g., Rayner, Carlson, & Frazier, 1983). The notion that argument structure is visible during online processing would be more congruent with constraint-based models in the sense that they more overtly assume memory constraints over parsing operations, as well as they entertain that possible connections associated with verbal knowledge are to be activated in real time. However evidence for argument structure information to permeate parsing operations does not necessarily dismiss the alternative serial/modular account, as we discuss below.

At the same time that Ledoux et al.'s (2007) interpretation of their results supports frequency-based parsing models, which conceive that the verb matrix is to remain active during the whole course of processing (Ferretti, McRae, Hatherell, 2001; McRae, Ferretti, & Amyote, 1997; Pickering & Frisson, 2001), their results do not necessarily exclude syntax-first models. In fact, in an earlier study, Frazier et al. (1984) suggests the possibility that expediency constraints over parsing operations could propel the reapplication of the recently build structure, once the parser has indications that the subsequent structure is congruent with the preceding one. The syntactic prediction locality theory (Gibson, 1998) indeed entertains the influence of computational resources

over syntactic processing. In the case of object-shifted structures, such as middles, a greater memory load would be required in order sustain the postulation of a *trace* to be inserted at a later point in the structure (see Just & Carpenter (1992), and Waters & Caplan (2004, 2005) for a debate on working memory constraints over online processing). Even though Frazier at al. (1984)'s proposal does not necessarily explain Ledoux et al.'s (2007) long distance priming effects, it does predict our findings, in the sense that the conjunction *but* would have triggered the expectation for a subsequent unaccusative structure.

If priming is in effect by the maintenance of the whole structure, lexical items included, as Ledoux et al. (2007) suggest, then it would had been expected that we found indications of revision at the verb region in all three sentence types, especially because the verb in the second conjunct—which was kept constant across conditions (e.g., *chef cuts*, *knife cuts*, and *brad cuts*) — was different than the copula verb *be* used in the first conjunct. Our results do not corroborate such prediction. We do assume, however, the role of semantic priming in our results. Given the priming for an unaccusative structure, a grater processing cost at pre-verbal region in the transitive conditions would be elicited because the NP2 *agent* (*chef*) and the NP2 *instrument* (*knife*) are less congruent with a noun complement than the NP2 *theme* (*bread*) in the Middle condition. This greater processing cost could be attributable to the shift from the primed unaccusative construction to a transitive one. It is a plausible interpretation that priming between the semantically related NP1 and NP2 alleviated the processing cost at the NP2 in the transitive conditions, thus the small to medium effects found at this position in both

Experiment 1 and 3. Ultimately, our findings are congruent with the notion that the phenomenon of syntactic priming can be independent from lexical repetition.

Our study presents suggestive evidence that syntactic priming effects can be observed with distinct verbs, at least to the extent that target and prime verbs can be realized over the same argument structure. We observed that semantic and syntactic priming can occur concomitantly, and that such interaction proved to be strong enough to counteract effects of canonicity, and to counterpose the parser's preference for building the simplest structure.

However, faced with the evidence, we suggest that, beyond what has been called syntactic priming, the phenomenon here captured is more congruent with the proposal of an argument structure priming effect. If, as proposed by Di Sciullo (2005), argument-structures are elements of the grammar, which are to be present at different domains of the linguistic knowledge, then they also have to be operationalized at the syntax. Once argument-structures become part of syntax, the hierarchical precedence of syntax is maintained, and it allows the parser to process this kind of information in real time. Such approach does not dismiss influences of other cognitive resources associated with language processing, such as memory, at the same time it does preserve the domain specificity of linguistic knowledge.

References

Ackema, P., & Schoorlemmer, M. (1994). The middle construction and the syntax-semantics interface. *Language*, *93*, 59–90.

Ackema, P., & Schoorlemmer, M. (1995). Middles and nonmovement. *Linguistic Inquiry*, *26*, 173–197.

Ackema, P., & Schoorlemmer, M. (2003). Review of Markus Steinbach, the middle voice. *Journal of Germanic Linguistics*, *15*, 372–384.

Adams, B. A., Clifton, C., Jr., & Mitchell, D. C. (1998). Lexical guidance in sentence processing? *Psychonomic Bulletin and Review*, *5*, 265–270.

Alboiu, G., Barrie, M., & Frigeni, C. (2004). SE and the unaccusative/unergative paradox. In Coene, M., de Cuyper, G., & D'Hulst, Y. (Eds.), *Antwerp working papers in linguistics*, 107, 109–139. Universiteit Antwerp.

Alexiadou, A., & Doron, E. (2012). The syntactic construction of two non-active voices: passive and middle. *Journal of Linguistics, 48*, 1–34

Alexiadou, A., Haegeman, L., & Starvou, M. (2007). Noun phrase in the generative perspective. Berlin: Mouton de Gruyter.

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

Arai, M., van Gompel, R. P. G., & Scheepers, C. (2007). Priming ditransitive structures in comprehension. *Cognitive Psychology*, *54*, 218–250.

Baayen, R. H., Piepenbrock, R., & Van Rijn, H. (1993). *The CELEX lexical database [CD-ROM]*. Philadelphia: Linguistic Data Consortium.

Baker, M. (1983). Objects, themes, and lexical rules in Italian. In Levin, L., Rappaport-Hovav, M., & Zaenen, A. (Eds.), *Papers in lexical-functional grammar*, 1–46. Bloomington: Indiana University Linguistics Club.

Bowers, J. (2002). Transitivity. *Linguistic Inquiry*, 2-33, 183–224

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41*(4), 977–990.

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: a review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology, 58,* 412–424.

Burzio, L. (1986). *Italian syntax: A government-binding approach*. Dordrecht, The Netherlands: Reidel.

Chierchia, G. (2004). A semantics for unaccusatives and its syntactic consequences. In Alexiadou, A., E. Anagnostopoulou, E., & Everaert, M. (Eds.), *The unaccusativity puzzle: Explorations of the syntax-lexicon interface* (pp. 22–59). Oxford, UK: Oxford University Press.

Chomsky, N. (1981). Lectures on government and binding. Dordrecht, The Netherlands: Foris.

Chomsky, N. (1995). The minimalist program. Cambridge, MA: The MIT Press.

Clifton, C. Jr., Traxler, M. J., Mohamed, M. T., Williams, R. S., Morris, R. K, & Rayner, K. (2003). The use of thematic role information in parsing: Syntactic processing autonomy revisited. *Journal of Memory and Language*, *49*, 317–334.

Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive Graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, and Computers, 25,* 257–271.

Correa, P. (2007). Expressao da mudanca de estado na interlingua de brasileiros aprendizes de espanhol. (Doctoral dissertation). Federal University of Rio de Janeiro, Rio de Janeiro.

Davies, M. (2009). The 385 Million Word Corpus of Contemporary American English (1990-2008): Design, Architecture, and Linguistic Insights. *International Journal of Corpus Linguistics, 14,* 159-90.

de Almeida, R. G., Manouilidou, C., Roncero, C., Riven, L. (2010, submitted, under review). Three tales of semantic decomposition: Causatives, coercion, and metaphors.

De Vincenzi, M. (1991). *Syntactic parsing strategies in Italian.* Dordrecht, The Netherlands: Kluwer.

Di Sciullo, A. M. (1993). The complement domain of a head at morphological form. Probus. *International Journal of Latin and Romance Linguistics*, 5(*2*), 95-125.

Di Sciullo, A. M. (1995). A modular approach to the selectional properties of derivational affixes. *El Coloquio de Grammatica Generativa*. University of Vitoria-Gasteiz, 67-91.

Di Sciullo, A. M. (1997). Prefixed verbs and adjunct identification. In Di Sciullo, A. M. (Ed.), *Projections and Interface Conditions: Essays on Modularity* (pp. 52–74). New York: Oxford University Press.

Di Sciullo, A. M. (2005). *Asymmetry in morphology.* Cambridge, MA: The MIT Press.

Di Sciullo, A. M., de Almeida, R. G., Manouilidou, C., & Dwivedi, V. (2007, August). *This poster reads clearly: Processing English middle constructions*. Poster presented at the 13th Annual Conference on Architectures and Mechanisms of Language Processing Conference. Turku, Finland.

Doron, E. (2003). Agency and voice: the semantics of Semitic templates. *Natural Language Semantics*, *11*, 1–67.

Dowty, D. R. (1991). Thematic proto roles and argument selection. *Language*, *67*, 547–619.

Embick, D. (2004). Unaccusative syntax and verbal alternations. In Alexiadou, A., E. Anagnostopoulou, E., & Everaert, M. (Eds.), *The unaccusativity puzzle: Explorations of the syntax-lexicon interface* (pp. 137–158). Oxford, UK: Oxford University Press.

Fagan, S. (1988). The English middle. *Linguistic Inquiry*. *19*, 181–203.

Fagan, S. (1992). *The syntax and semantics of middle constructions. A study with special reference to german*. Cambridge: Cambridge University Press.

Ferretti, T. R., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language, 44***,** 516–547.

Fodor, J. A. (1983). *Modularity of mind: An essay on faculty psychology*. Cambridge, MA: The MIT Press.

Fodor, J. D., & Frazier, L. (1980).Is the human sentence processing mechanism an ATN? *Cognition, 8,* 417–459.

Frazier, L (1989). Against lexical generation of syntax. In Marslen-Wilson, W. (Ed.), *Lexical representation and process* (pp. 505–528). Cambridge, MA: The MIT Press.

Frazier, L., & Fodor, J. D. (1978). The sausage machine: a new two-stage parsing model. *Cognition, 6*, 291–325.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14*, 178–210.

Frazier, L., Taft, L., Roeper, T., & Clifton, C. Jr. (1984). Parallel structure: a source of facilitation in sentence comprehension. *Memory and Cognition, 12*, 421–430.

Frazier, L., & Clifton, C., Jr. (1989) Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes, 4,* 93–126.

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

Frazier, L., Munn, A., & Clifton, C., Jr. (2000). Processing coordinate structures. *Journal of Psycholinguistic Research, 29*, 343–371.

Friedmann, N., Taranto, G., Shapiro, L.P., & Swinney, D. (2008). The leaf fell (the leaf): The online processing of unaccusatives. *Linguistic Inquiry*, *39*, 355–377

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition, 68*, 1–76.

Gibson, E. A. F., Pearlmutter, N., Canseco-Gonzalez, E., & Hickok, G. (1996). Recency preference in the human sentence processing mechanism. *Cognition, 59,* 23–59.

Grimshaw, J. (1990). *Argument Structure*. Cambridge, MA: The MIT Press .

Gruber, J. (1965). *Studies in lexical relations*. Bloomington, IN: Indiana University Linguistics Club.

Gruber, J., & Gibson, E. (2004). Measuring linguistic complexity independent of plausibility. *Language, 80,* 583–590.

Haegeman, L. (1994). Introduction to Government and Binding Theory.

Hale, K., & Keyser, S. J. (1993). On argument structure and the lexical expression of syntactic relations. *The View from Building 20*. Cambridge, MA: MIT Press.

Hale, K., & Keyser, S. J. (2002). *Prolegomenon to a theory of argument structure*. Cambridge,Mass.: MIT Press.

Harley, H. (2006). On the causative construction. In Miyagawa, S., & Saito M. (Eds.), *Handbook of Japanese linguistics*. Oxford, UK: Oxford University Press. (http://ling.auf.net/lingBuzz/000433)

Hickok, G., Canseco-Gonzalez, E., Zurif, E., & Grimshaw, J. (1992). Modularity in locating wh-gaps. *Journal of Psycholinguistic Research, 21*, 547–563.

Hoekstra, T & Roberts, I. (1993). Middle Construction in Dutch and English. In Reuland, E. & Abraham, W. (eds.), *Knowledge and Language* (pp. 183–220). Dordrecht: Kluwer Academic Publishers.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H. & Van de Weijer, J. (2011). Eye-tracking: A comprehensive guide to methods and measures. Oxford, UK: Oxford University Press.

Jackendoff, R. (1990). *Semantic structures*. Cambridge, MA: MIT Press.

Jackendoff, R. (1989). What is a concept, that a person may grasp it? *Mind and Language, 4*, 68–102.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological Review, 99*, 122–149.

Kaiser, E., & Trueswell, J. C. (2004). The role of discourse context in the processing of a flexible word-order language. *Cognition*, 94, 113–147.

Kayne, R. S. (1994). *The Antisymmetry of Syntax*. The MIT Press: Cambridge.

Kemmer, S. (1993). *The middle voice*. Amsterdam: John Benjamins,

Keyser, S. J., & Roeper, T. (1984). On the middle and ergative construction in English. *Linguistic Inquiry, 15*, 381–416.

Knoeferle, P., & Crocker, M. W. (2009). Constituent order and semantic parallelism in online comprehension: Eye-tracking evidence from German, *The Quarterly Journal of Experimental Psychology, 62*, 2338–2371.

Kučera, H., & Francis, W.N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Laka, I. (2000). Theta blind case: Burzio's generalization and its image in the mirror. In Reuland, E. (Ed). *Arguments and case: explaining Burzio's generalization (pp.* 103–129). Amsterdam: John Benjamins.

Ledoux, K., Traxler, M., & Swaab, T.Y. (2007). Syntactic priming in comprehension: Evidence from event-related potentials. *Psychological Science*, *18*, 135–143.

Lekakou, M. (2005). In the middle, somewhat elevated. The semantics of middles and its cross-linguistics realizations (Doctoral dissertation). University College London, UK.

Levin, B. (1993). *English verb classes and alternations. A preliminary investigation*. Chicago: The University of Chicago Press.

Levin, B., & Rappaport-Hovav, M. (2005). *Argument realization*. Cambridge, UK: Cambridge University Press.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The Lexical Nature of Syntactic Ambiguity Resolution. *Psychological Review, 101*, 676–703.

Maia, M.A.R., Oliveira, R., & Santos, S. (submitted). Este título leria mais claramente em Karajá do que em Xavante, Português (ou Inglês): Um estudo comparativo sobre o processamento da alternância causativa. In Storto, L. (Ed). *Estrutura Argumental em Línguas Indígenas*.

Manouilidou, C., de Almeida, R. G., Schwartz, G., & Nair, N. P. V. (2009). Thematic roles in Alzheimer's disease: Hierarchy violations in psychological predicates. *Journal of Neurolinguistics, 22, 167*–186.

Marantz, A. (1997). No escape from syntax: Don´t try a morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers in Linguistics 4 (2)*, 201–225.

Marelj, M. (2004). *Middles and argument structure across languages* (Doctoral dissertation). The Netherlands: University of Utrecht.

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

McKoon, G., & Macfarland, T. (2002). Event templates in the lexical representations of verbs. *Cognitive Psychology, 45,* 1–44.

McRae, K., Ferretti, T.J., & Amyote, L. (1997). Thematic roles as verb-specific concepts. *Language and Cognitive Processes, 12*, 137–176.

Miyamoto, E. T., & Takahashi, S. (2000). The processing of VP-internal scrambling in Japanese. Talk presented at the Japanese Language Processing Workshop. Makuhari, Japan.

Mitchell, D. C., Cuetos, F., Corley, M. M. B., & Brysbaert, M. (1995). Exposure-based models of human parsing: evidence for the use of coarse-grained (nonlexical) statistical records. *Journal of Psycholinguistic Research, 24,* 469–488.

Perlmutter, D. (1978). Impersonal passives and the unaccusative hypothesis. In. Jaeger J. J. et al. (Eds.). *Proceedings of the Fourth Annual Meeting of the Berkeley Linguistics Society* (pp. 157–189). Berkeley, CA: University of California, Berkeley Linguistics Society.

Pesetsky, D. (1995) *Zero Syntax: Experiencer and cascades*. Cambridge, MA: The MIT Press.

Philipp, M., Bornkessel-Schlesewsky, I., Bisang, W., & Schlesewsky, M. (2008). The role of animacy in the real time comprehension of Mandarin Chinese: Evidence from auditory event-related brain potentials. *Brain and Language 105*, 112–133.

Pickering, M. J., & Barry, G. (1991). Sentence processing without empty categories. *Language and Cognitive Processes*, 6, 229–259.

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

Pickering, M.J., & Traxler, M.J. (2003). Evidence against the use of subcategorisation information in the processing of unbounded dependencies. *Language and Cognitive Processes, 18*, 469–503.

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

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quartely Journal of Experimental Psychology. 62*, 457–506.

Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior, 22*, 358–374.

Rayner, K., & Pollatsek, A. (1989). *The psychology of Reading.* Englewood Cliffs, NJ: Prentice Hall.

Reinhart, T., & Reuland, E. (1983). Reflexivity. *Linguistic Inquiry, 24,* 657–720.

Reinhart, T., & Siloni, T. (2004). Against the unaccusative analysis of reflexives. In Alexiadou, A., E. Anagnostopoulou, E., & Everaert, M. (Eds.), *The unaccusativity puzzle: Explorations of the syntax-lexicon interface* (pp. 159–180). Oxford, UK: Oxford University Press.

Reinhart, T., & Siloni, T. (2005). The lexicon-syntax parameter: Reflexivization and other arity operations. *Linguistic Inquiry, 36,* 389–436.

Rizzi, L. (1986). Null objects in Italian and the theory of *pro*. *Linguistic Inquiry*. *17,* 501–557.

Rohde, D, & Gibson, E. (2003). The on-line processing of active and passive structures in English. Poster presented at the 16th Annual CUNY Conference on Human Sentence Processing, MIT. Cambridge, MA.

Sekerina, I. (2003). Scrambling and processing: Dependencies, complexity and constraints. In S. Karimi, S. (Ed.), *Scrambling and word order* (pp. 301–324). Malden, MA: Blackwell.

Shapiro, L. P.; Zurif, E. & Grimshaw, J. (1987). Sentence processing and the mental representation of verbs. *Cognition, 27*, 219-

SR Research. (2009). EyeLink-II head-mounted user manual. Mississauga, Ontario, Canada: SR Research.

Staub, A. (2007). The parser doesn't ignore intransitivity, after all. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *33*, 550–569

Staub, A. (2010). Eye movements and processing difficulty in object relative clauses. *Cognition, 116,* 71–86.

Staub, A., Clifton, C., Jr., & Frazier, L. (2006). Heavy NNP shift is the parser's last resort: evidence from eye movements. *Journal of Memory and Language*, *54,* 389–406.

Staub, A., & Rayner, K. (2007). Eye movements and on-line comprehension processes. In Gaskell, G. (Ed.), *The Oxford Handbook of Psycholinguistics* (pp. 327–342). Oxford, UK: Oxford University Press.

Steinbach, M. (2002). Middle voice: A comparative study in the syntax-semantics interface of German. *Linguistics Today*. North America, PA: John Benjamins.

Stroik, T. (1992). Middles and movement. *Linguistic Inquiry, 23*, 127–137.

Stroik, T. (1995). On middle formation: A reply to Zribi-Hertz. *Linguistic inquiry, 26*, 165–171.

Stroik, T. (1999). Middles and reflexivity. *Linguistic Inquiry.* 30, 119–131

Stroik, T. (2006). Arguments in middles. In Lyngfelt, B. & T. Solstad T. (Eds.), D*emoting the agent: passive, middle and other voice phenomena* (pp. 301–326). Linguistics Today, 96. Amsterdam: John Benjamins.

Sturt, P., Keller, F. & Dubey, A. (2010). Syntactic priming in comprehension: Parallelism effects with and without coordination. *Journal of Memory and Language*, *62*, 333–351

Swinney, D. (1991). The resolution of indeterminacy during language comprehension: evidence from lexical and structural processing. In Simpson, G. (Ed.), *Understanding Word and Sentence* (Advances in Psychology Series). North-Holland, Amsterdam.

Swinney, D., Onifer, W., Prathe, P., & Hirshkowitz, M. (1979). Semantic facilitation across sensory modalities in the processing of individual words and sentences. *Memory and Cognition 7*: 159–165.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics.* Boston, MA: Pearson Education Inc. (Eds.).

Tooley, K. M., Traxler, M. J. Swaab, T.Y. (2009). Electrophysiological and behavioral evidence of syntactic priming in sentence comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 19–45.

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

Traxler, M. J., Morris, R. K.& Seely, R. (2002). Processing subject and object relative clauses: evidence from eye movements. *Journal of Memory and Language* 47, 69–90.

Trueswell, J. C. (1996). The Role of Lexical Frequency in Syntactic Ambiguity Resolution. *Journal of Memory and Language*. *35*, 566–585

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, *33*, 285–318.

Van Gompel, R. P. G., Pickering, M. J., & Traxler, M. J. (2001). Making and revising syntactic commitments: Evidence against current constraint-based and two-stage models. *Journal of Memory and Language, 45*, 283–307.

Van Selst, M., & Jolicoeur, P. (1994). A solution to the effect of sample size on outlier elimination. *The Quarterly Journal of Experimental Psychology*, *47a*, 631–650.

Waters, G. S., & Caplan, D. (2004). Individual differences in working memory capacity and on-line syntactic processing: Evidence from self-paced listening. *Quarterly Journal of Experimental Psychology*, *57*, 129–163.

Waters, G. S., & Caplan, D. (2005). The relationship between age, processing speed, working memory capacity and language comprehension. *Memory*, *13*, 403–413.

Wenzhong, H. (2005). Defining the Middle Construction. *Foreign Language education*, *4*. doi: cnki: ISSN:1000-5544.0.2005-04-002

Woolford, E. (2003). Burzio's generalization, markedness, and locality constraints on nominative objects. In *Brandner, E. & Zinsmeister, H. (Eds.), New Perspectives on Case Theory (pp.301–329).* Stanford, CA: CSLI.

Zribi-Hertz, A. (1993). On Stroik's analysis of English middle constructions. *Linguistic Inquiry. 24,* 583–589.

Appendix A

Below are the materials used in Experiment 1.

Sentences in the Middle Condition

That batter is perfect, but this dough bakes poorly.
That roast is tender, but this turkey carves badly.
That carpet is unwashable, but this floor cleans perfectly.
That pan burns everything, but this crockpot cooks slowly.
That croissant is flaky, but this bread cuts neatly.
That stucco is rough, but this wall paints smoothly.
That carrot is mushy, but this potato peels properly.
That violin is cumbersome, but this guitar plays easily.
That soup is chunky, but the sauce pours smoothly.
That article is confusing, but this paper reads clearly.
That novel is unpopular, but this book sells steadily.
That roast is stringy, but this meat slices evenly.
Those shares are down, but this stock trades well.
That chronicle is unclear, but this text translates quickly.
Those artifacts are fragile, but these valuables transport reliably.

Sentences in the Transitive Inanimate Condition

That grill is great, but this oven bakes poorly.
That knife is sharp, but this blade carves badly.
That detergent is awful, but this soap cleans perfectly.
That microwave is brand new, but this stove cooks slowly.
That blade is dull, but this knife cuts neatly.
That roller is clumpy, but this brush paints smoothly.
That peeler is rusted, but this utensil peels properly.
That equipment is confusing, but this turntable plays easily.
That kettle is leaking, but this pitcher pours smoothly.
That machine is broken, but this scanner reads clearly.

That shop is going under, but this store sells steadily.

That electric knife is broken, but this blade slices evenly.

That consortium is tainted, but this firm trades well.

That program is buggy, but this software translates quickly.

Those vans are too old, but these trucks transport reliably.

Sentences in the Transitive Animate Condition

That man is a chef, but this woman bakes poorly.

That artist is dexterous, but this sculptor carves badly.

That housekeeper is lousy, but this maid cleans perfectly.

That maid is skilled, but this nanny cooks slowly.

That cook is messy, but this chef cuts neatly.

That contractor is sloppy, but this worker paints smoothly.

That cook is careless, but this baker peels properly.

That musician is hesitant, but this pianist plays easily.

That barmaid is clumsy, but this bartender pours smoothly.

That teacher stutters, but this tutor reads clearly.

That employee is lazy, but this clerk sells steadily.

That assistant is inept, but this cook slices evenly.

That agent is incompetent, but this broker trades well.

That interpreter is slow, but this teacher translates quickly.

Those movers are reckless, but these drivers transport reliably.

Appendix B

Below are the materials used in Experiments 2 and 3.

Sentences in the Middle Condition (Experiments 2 and 3)

That batter is perfect, but this dough bakes poorly in the oven.
That roast is tender, but this turkey carves badly with a knife.
That carpet is unwashable, but this floor cleans perfectly with this mop.
That chicken always burns, but this fillet cooks slowly on the stove.
That croissant is flaky, but this bread cuts neatly in small slices.
That stucco is rough, but this wall paints smoothly with a brush.
That carrot is mushy, but this potato peels properly with a knife.
That symphony is beautiful, but this song plays easily on the radio.
That soup is chunky, but this sauce pours smoothly on the plate.
That article is confusing, but this paper reads clearly with the students.
That novel is unpopular, but this book sells steadily in all stores.
That roast is dry, but this meat slices evenly with any knife.
Those shares are down, but these stocks trade well in the market.
That chronicle is unclear, but this text translates quickly with our software.
Those artefacts are fragile, but these valuables transport reliably in the truck.

Sentences in the Transitive Inanimate Condition (Experiments 2 and 3)

That grill is great, but this oven bakes poorly at high heat.
That knife is sharp, but this blade carves badly at all times.
That detergent is awful, but this soap cleans perfectly in the machine.
That microwave is brand new, but this stove cooks slowly on high heat.
That blade is dull, but this knife cuts neatly through any meat.
That roller is clumpy, but this brush paints smoothly on all surfaces.
That peeler is rusting, but this utensil peels properly under water.
That violin is cumbersome, but this guitar plays easily on the streets.
That kettle is leaking, but this pitcher pours smoothly into the cups.
That machine is broken, but this scanner reads clearly under all conditions.

That shop is going under, but this store sells steadily all year long.

That electric knife is broken, but this blade slices evenly at any time.

That consortium is tainted, but this firm trades well in the market.

That program is buggy, but this software translates quickly into Chinese.

Those vans are too old, but these trucks transport reliably in bad weather.

Sentences in the Transitive Animate Condition (Experiments 2 and 3)

That man is a chef, but this woman bakes poorly at the hotel.

That artist is dexterous, but this sculptor carves badly onto the wood.

That housekeeper is lousy, but this maid cleans perfectly around the stove.

That maid is skilled, but this nanny cooks slowly at all times.

That cook is messy, but this chef cuts neatly with the knife.

That contractor is sloppy, but this worker paints smoothly with any tool.

That cook is careless, but this baker peels properly for fruit tarts.

That musician is hesitant, but this pianist plays easily at concerts.

That barmaid is clumsy, but this bartender pours smoothly into the pitchers.

That teacher stutters, but this tutor reads clearly on any occasion.

That employee is lazy, but this clerk sells steadily at the men's section.

That assistant is inept, but this cook slices evenly through any roast.

That agent is incompetent, but this broker trades well under pressure.

That interpreter is slow, but this teacher translates quickly at conferences.

Those movers are reckless, but these drivers transport reliably at all times.

Sentences in the Passive Condition (Experiment 2)

That batter is perfect, but this dough was baked poorly in the oven.

That roast is tender, but this turkey was carved badly with a knife.

That carpet is unwashable, but this floor was cleaned perfectly with this mop.

That chicken always burns, but this fillet was cooked slowly on the stove.

That croissant is flaky, but this bread was cut neatly in small slices.

That stucco is rough, but this wall was painted smoothly with a brush.

That carrot is mushy, but this potato was peeled properly with a knife.

That symphony is beautiful, but this song was played easily on the radio.

That soup is chunky, but this sauce was poured smoothly on the plate.

That article is confusing, but this paper was read clearly with the students.

That novel is unpopular, but this book was sold steadily in all stores.

That roast is dry, but this meat was sliced evenly with any knife.

Those shares are down, but these stocks were traded well in the market.

That chronicle is unclear, but this text was translated quickly with our software.

Those artefacts are fragile, but these valuables were transported reliably in the truck.

Sentences in the Predicative Inanimate Condition (Experiment 2)

That grill is great, but this oven is a poor baker at high heat.

That knife is sharp, but this blade is a bad carver at all times.

That detergent is awful, but this soap is a perfect cleaner in the machine.

That microwave is brand new, but this stove is a slow cooker on high heat.

That blade is dull, but this knife is a neat cutter through any meat.

That roller is clumpy, but this brush is a smooth painter on all surfaces.

That peeler is rusting, but this utensil is a proper peeler under water.

That violin is cumbersome, but this guitar is an easy player on the streets.

That kettle is leaking, but this pitcher is a smooth pourer into the cups.

That machine is broken, but this scanner is a clear reader under all conditions.

That shop is going under, but this store is a steady seller all year long.

That electric knife is broken, but this blade is an even slicer at any time.

That consortium is tainted, but this firm is a good trader well in the market.

That program is buggy, but this software is a quick translator into Chinese.

Those vans are too old, but these trucks are reliable transporters in bad weather.

Sentences in the Predicative Animate Condition (Experiment 2)

That man is a chef, but this woman is a poor baker at the hotel.

That artist is dexterous, but this sculptor is a bad carver onto the wood.

That housekeeper is lousy, but this maid is a perfect cleaner around the stove.

That maid is skilled, but this nanny is a slow cooker at all times.

That cook is messy, but this chef is a neat cutter with the knife.

That contractor is sloppy, but this worker is a smooth painter with any tool.

That cook is careless, but this baker is a proper peeler for fruit tarts.

That musician is hesitant, but this pianist is an easy player at concerts.

That barmaid is clumsy, but this bartender is a smooth pourer into the pitchers.

That teacher stutters, but this tutor is a clear reader on any occasion.

That employee is lazy, but this clerk is a steady seller at the men's section.

That assistant is inept, but this cook is an even slicer through any roast.

That agent is incompetent, but this broker is a good trader under pressure.

That interpreter is slow, but this teacher is a quick translator at conferences.

Those movers are reckless, but these drivers are reliable transporters at all times.
