A multi-methodological investigation of the processing and interpretation of coordinate sentences involving verb phrase anaphors

Callahan, Sarah M.

2009

UC San Diego Electronic Theses and Dissertations

Ph. D., Psychology UC San Diego

http://escholarship.org/uc/item/42g752wv

b6636470

The studies in this dissertation investigated the processing and interpretation of coordinate sentences involving verb phrase (VP) anaphors by obtaining converging evidence from off-line measures sensitive to conscious preference and on-line measures sensitive to real-time processing. The studies in Chapters 2, 3, and 4 investigated the effects of anaphoric form by comparing sentences involving a null VP-anaphoric phrase (e.g. did too) with those involving an overt VP-anaphoric phrase (e.g. did it too). The results indicated that participants were highly accurate in comprehending both forms and interpret both forms similarly; nonetheless, anaphoric form did affect processing, suggesting a number of conclusions. First, in contexts where multiple anaphoric forms are grammatical and there is a single highly-accessible antecedent, comprehenders expected the least informative anaphoric form possible, as reflected in higher naturalness ratings for the null VP-anaphoric condition. Second, violations of this expectation elicited additional processing at the anaphor, as reflected in increased initial processing for the overt VP-anaphoric condition. Finally, variation in when the anaphoric relationship can be detected resulted in variation in processing, as reflected in a reversal pattern whereby the null VP-anaphoric condition was later associated with increased processing. Inspired by the natural linking between anaphora and parallelism, the study in Chapter 5 then investigated the processes underlying parallelism effects in coordinate sentences involving VP anaphors. This study used a cross-modal priming methodology to evaluate the activation of a verb from the first conjunct throughout the two conjuncts. The results indicated that activation related to the initial presentation of the verb decayed relatively quickly. Then, reactivation occurred immediately following a coordinating conjunction (i.e. and) and was sustained throughout the second conjunct. Since this reactivation renders the first conjunct highly-accessible, the results of this study suggested that effects related
to parallelism and anaphoric form may be generated by similar underlying processes. By way of conclusion, the findings presented in this dissertation were interpreted in the context of a new model of the processing of coordinate sentences involving VP anaphors that highlights the interplay of prior experience and the immediate context in generating expectations that affect processing and interpretation.

Copyright Information:
All rights reserved unless otherwise indicated. Contact the author or original publisher for any necessary permissions. eScholarship is not the copyright owner for deposited works. Learn more at http://www.escholarship.org/help_copyright.html#reuse
A multi-methodological investigation of the processing and interpretation of coordinate sentences involving verb phrase anaphors

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Psychology

by

Sarah M. Callahan

Committee in charge:

Professor Tracy Love, Chair
Professor Seana Coulson
Professor Victor Ferreira
Professor John Moore
Professor John Wixted

2009
Copyright
Sarah M. Callahan, 2009
All rights reserved.
The Dissertation of Sarah M. Callahan is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_________________________________________________________________
_________________________________________________________________
_________________________________________________________________
_________________________________________________________________

Chair

University of California, San Diego

2009
DEDICATION

To Mom, who taught me perseverance. You were always there for me and I couldn’t have done it without you.

To Dave, who set me on the path. I miss our talks.

To Alex, who has stuck with me through the worst. I promise the best is yet to come.

To Tira, who is jealous of all the attention the laptop gets, but loves me anyway.
Ignorance has no beginning, but it has an end. There is a beginning but no end to knowledge.

B.K.S. Iyengar
# TABLE OF CONTENTS

Signature page ........................................................................................................ iii
Dedication ................................................................................................................ iv
Epigraph .................................................................................................................. v
Table of Contents .................................................................................................... vi
List of Tables ........................................................................................................... viii
List of Figures ......................................................................................................... ix
Acknowledgements ................................................................................................. x
Curriculum Vitae ..................................................................................................... xii
Abstract .................................................................................................................. xvi
Chapter 1: Introduction ......................................................................................... 1
  References ............................................................................................................. 15
Chapter 2: Effects of variation in anaphoric form on the processing and interpretation of VP-anaphoric constructions: An eye-tracking study .......... 20
  Abstract .............................................................................................................. 21
  Methods, Experiment 1 ...................................................................................... 29
  Results, Experiment 1 ....................................................................................... 32
  Discussion, Experiment 1 .................................................................................. 33
  Methods, Experiment 2 ...................................................................................... 34
  Results, Experiment 2 ....................................................................................... 36
  General Discussion .............................................................................................. 42
  References .......................................................................................................... 48
  Appendix ............................................................................................................. 51
Chapter 3: An ERP investigation of the effects of anaphoric form: Processing null and overt verb phrase anaphors ................................................................. 61
  Abstract .............................................................................................................. 62
  Methods .............................................................................................................. 75
  Results ................................................................................................................. 81
  Discussion .......................................................................................................... 86
  References .......................................................................................................... 95
LIST OF TABLES

Chapter 2
Table 1. Sample Sentences ................................................................. 56
Table 2. Means (ms) and Standard Errors for the NP-Anaphoric Conditions ................................................................. 57
Table 3. Means (ms) and Standard Errors for the VP-Anaphoric Conditions ................................................................. 58

Chapter 3
Table 1. Sample Sentences ................................................................. 101
Table 2. Means and Standard Deviations in terms of Percent Correct on the Sensicality Judgment Task ................................................................. 102

Chapter 4
Table 1. Sample Sentences ................................................................. 134
Table 2. Mean Accuracy (in %) and Standard Deviations for the Unambiguous Conditions ................................................................. 135
Table 3. Mean Naturalness Ratings and Standard Deviations ............... 136

Chapter 5
Table 1. Means and Standard Deviations in terms of Response Time (ms) for Experiments 1 and 2 ................................................................. 172
LIST OF FIGURES

Chapter 2

Figure 1. First-pass (top) and total (bottom) reading times for the NP-anaphoric conditions ................................................................. 59

Figure 2. First-pass (top) and total (bottom) reading times for the VP-anaphoric conditions ........................................................................ 60

Chapter 3

Figure 1. Grand average ERPs from the onset of the second subject for Plausible (dotted lines) and Implausible items (solid lines). Negative is plotted upwards. PLAUS = Plausible; IMPLAUS = Implausible .......... 103

Figure 2. Grand average ERPs from the onset of “too” for items involving Null VP anaphors (dotted lines) and items involving Overt VP anaphors (solid lines). Negative is plotted upwards. P-NVP = Plausible Null Verb Phrase; P-OVP = Plausible Overt Verb Phrase ....................... 104

Chapter 4

Figure 1. Sample question from the questionnaire illustrating the Likert scale ........................................................................................................ 137

Chapter 6

Figure 1. A model of the processing of coordinate sentences involving verb phrase anaphors ........................................................................ 184
ACKNOWLEDGEMENTS

I am forever indebted to my advisor for the last three years, Tracy Love, for her kindness, advice, and support. Had she not generously agreed to supervise me after my original advisor passed away, I would never have been able to complete this dissertation. I don’t know how she held it together when everything was falling apart, but I truly cannot thank her enough.

I am also extremely grateful for the three years I spent working with my first advisor, Dave Swinney. Although he did not get a chance to see the completion of this work, he was integral to the design and development of the studies in Chapters 2 and 3 of this dissertation. Even more, he was a true mentor in every sense of the word and the most generous spirit I have ever known. Though he is gone, I continue to follow his advice and his example not only in academia but also in life.

I would particularly like to thank my collaborator on the studies in Chapters 2 and 3, Janet Nicol, for her support and guidance over the years. Her comments are always tremendously insightful and her editing skills are second to none.

I would also like to thank my collaborator on the studies in Chapters 4 and 5, Lew Shapiro, for many helpful discussions. Our good-natured debates always clarify my thinking tremendously.

I am also very grateful to Barbara Lust, my mentor during college, who inspired and encouraged me to pursue research on language. She quietly nudged me into a new world where I think I always belonged.
I would also like to thank Matt Walenski for his endless patient assistance and guidance. Without his expertise in ERPs, statistics, and academia, I might not have survived.

Although they are too numerous to name individually, I would like to thank the many research assistants who assisted with the data collection for the studies contained herein. I do not delegate well, but I truly could not have done it alone. In particular, I would like to thank Jeffrey Witzel for his meticulous work collecting the eye-tracking data in Chapter 2.

Finally, I would like to thank the members of my committee for many helpful comments. Their suggestions have certainly made this dissertation much stronger.

Chapter 2, in full, has been submitted for publication as it may appear in *Journal of Experimental Psychology: Learning, Memory, & Cognition*, Callahan, S.M., Nicol, J., & Love, T. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, has been submitted for publication as it may appear in *Brain and Language*, Callahan, S.M., Nicol, J., Love, T., & Swinney, D. The dissertation author was the primary investigator and author of this paper.

Chapter 5, in full, has been submitted for publication as it may appear in *Journal of Psycholinguistic Research*, Callahan, S.M., Shapiro, L., & Love, T. The dissertation author was the primary investigator and author of this paper.
CURRICULUM VITAE

2009  Doctor of Philosophy in Psychology, University of California, San Diego

2004  Master of Arts in Psychology, University of California, San Diego

2001  Bachelor of Arts in Psychology, Cornell University, Ithaca, NY

FELLOWSHIPS/AWARDS

UCSD Department of Psychology Travel Award, 2004
Dean’s Scholarship, UCSD, 2003-2008
National Merit Scholarship, 1997-2001
Cornell Tradition Fellowship, Cornell University, 1997-2001
Dean’s List, Cornell University, 1997-2001
Human Ecology Dean’s Scholarship, Cornell University, 1997-1998

PAPERS


POSTERS


INVITED TALKS/LECTURES


“Processing different types of anaphors,” Center for Research in Language, May, 2007

“All pronouns are not created equal: The processing and interpretation of null and overt pronouns in Spanish,” Center for Research in Language, November, 2006

“Bilingualism: How do two languages share one brain?” UCSD Dept. of Linguistics, Psycholinguistics, November, 2006

“Bilingualism: How do two languages share one brain?” UCSD Dept. of Psychology, Psychology of Language, October, 2006
“All pronouns are not created equal: The processing and interpretation of null and overt pronouns in Spanish,” UCSD Dept. of Psychology Cognitive Brownbag, October, 2006

“All pronouns are not created equal: Anaphoric processing in speakers of Spanish and English,” UCSD Dept. of Psychology, June, 2004

RESEARCH EXPERIENCE

2003-2009  University of California, San Diego, Department of Psychology, Language and Cognitive Neuroscience Laboratory. Working with David Swinney, Ph.D. until he passed away in 2006 and currently working with Tracy Love, Ph.D., researching real-time anaphoric processing in English monolinguals, Spanish monolinguals, and Spanish-English bilinguals, using a variety of methodologies including behavioral techniques, electrophysiological techniques, and eye-tracking techniques (in collaboration with Janet Nicol, Ph.D., Univ. of AZ, Tucson). Designed and programmed experiments, tested participants, analyzed behavioral and electrophysiological data, presented findings at conferences, and prepared manuscripts for publication.

2001-2003  Columbia University, Department of Radiology, fMRI Research Center. Working with Joy Hirsch, Ph.D. as the Neuroscience Research Coordinator, assisted with the design, running, and analysis of fMRI experiments on a variety of topics. With Barbara Lust, Ph.D., Suzanne Flynn, Ph.D., Gita Martohardjono, Ph.D., Valerie Shafer, Ph.D., and Kathy Kessler, Ph.D., conducted collaborative research on sentence processing in Spanish-English bilinguals.

2000-2001  Cornell University, Department of Human Development, Cornell Language Acquisition Laboratory. Working with Barbara Lust, Ph.D. as a research assistant, assisted with the transcription of interviews with Spanish-speaking children.

TEACHING EXPERIENCE

2008  Lecturer, Introduction to Statistics. University of California, San Diego. Taught large undergraduate lecture course. Supervised graduate and undergraduate teaching assistants, prepared lectures, homework, and exams.
2003-2009  **Supervisor, Independent Study**, University of California, San Diego. Upper division undergraduate elective. Mentored undergraduate students, including training in experimental design, testing participants, behavioral and electrophysiological methods, and data analysis.

2007-2008  **Teaching Assistant, Introduction to Statistics**, taught by various instructors, University of California, San Diego. Large undergraduate lecture course. Taught discussion sections, guest lectured, graded exams.

2004-2006  **Teaching Assistant, Psychology of Language**, taught by David Swinney, Ph.D., University of California, San Diego. Upper division undergraduate lecture course. Taught discussion sections, guest lectured, graded exams.

2000  **Teaching Assistant, Language Acquisition**, taught by Barbara Lust, Ph.D., Cornell University. Upper division undergraduate lecture course. Undergraduate teaching assistant and grader.

**ACADEMIC SERVICE**

Faculty Recruitment Representative, University of California, San Diego, 2005-2006
Colloquium Representative, University of California, San Diego, 2004-2005
ABSTRACT OF THE DISSERTATION

A multi-methodological investigation of the processing and interpretation of coordinate sentences involving verb phrase anaphors

by

Sarah M. Callahan

Doctor of Philosophy in Psychology

University of California, San Diego, 2009

Professor Tracy Love, Chair

The studies in this dissertation investigated the processing and interpretation of coordinate sentences involving verb phrase (VP) anaphors by obtaining converging evidence from off-line measures sensitive to conscious preference and on-line measures sensitive to real-time processing. The studies in Chapters 2, 3, and 4 investigated the effects of anaphoric form by comparing sentences involving a null VP-anaphoric phrase (e.g. did too) with those involving an overt VP-anaphoric phrase (e.g. did it too). The results indicated that participants were highly accurate in comprehending both forms and interpret both forms similarly; nonetheless, anaphoric form did affect processing, suggesting a number of conclusions. First, in contexts
where multiple anaphoric forms are grammatical and there is a single highly-accessible antecedent, comprehenders expected the least informative anaphoric form possible, as reflected in higher naturalness ratings for the null VP-anaphoric condition. Second, violations of this expectation elicited additional processing at the anaphor, as reflected in increased initial processing for the overt VP-anaphoric condition. Finally, variation in when the anaphoric relationship can be detected resulted in variation in processing, as reflected in a reversal pattern whereby the null VP-anaphoric condition was later associated with increased processing.

Inspired by the natural linking between anaphora and parallelism, the study in Chapter 5 then investigated the processes underlying parallelism effects in coordinate sentences involving VP anaphors. This study used a cross-modal priming methodology to evaluate the activation of a verb from the first conjunct throughout the two conjuncts. The results indicated that activation related to the initial presentation of the verb decayed relatively quickly. Then, reactivation occurred immediately following a coordinating conjunction (i.e. *and*) and was sustained throughout the second conjunct. Since this reactivation renders the first conjunct highly-accessible, the results of this study suggested that effects related to parallelism and anaphoric form may be generated by similar underlying processes. By way of conclusion, the findings presented in this dissertation were interpreted in the context of a new model of the processing of coordinate sentences involving VP anaphors that highlights the interplay of prior experience and the immediate context in generating expectations that affect processing and interpretation.
Chapter 1

Introduction
The term, *anaphora*, describes the relationship between two phrases that refer to the same real-world entity or concept (i.e. the antecedent and anaphor, indicated by italics in examples). In context, an anaphor receives its interpretation via a link to a preceding antecedent phrase. Anaphora occurs regularly in natural language and appears to be universal across languages. Hence, linguistic theories of grammar (e.g. Chomsky, 1981, 1986) and psycholinguistic models of sentence processing (e.g. Frazier, 1978, 1987; Frazier & Clifton, 1996; MacDonald, Pearlmutter, & Seidenberg, 1994) have incorporated accounts of anaphoric phenomena. For quite some time, linguists and psycholinguists have also studied anaphoric constructions, yielding a large literature of introspective and behavioral data (see Chomsky, 1981, 1986; Fodor, 1989; Garnham, 2001; Garrod & Sanford, 1994; Nicol & Swinney, 2002; Trueswell & Tanenhaus, 2005 for reviews). In the last fifteen years, researchers have also begun to study anaphoric constructions using neurophysiological methodologies. These investigations have led to the development of a number of linguistic theories of anaphora (e.g. Ariel, 1990, 1991; Chomsky, 1981; Givón, 1983; Murphy, 1985a; Reinhart & Reuland, 1993) and psycholinguistic models of anaphoric processing (e.g. Almor, 1998; Garrod & Sanford, 1994; Garrod & Terras, 2000; Gordon & Hendrick, 1998; Sag & Hankamer, 1984). In all of these theories, anaphoric form plays a critical role.
Anaphoric Form

Comprehenders use information provided by the anaphor to identify the antecedent. Anaphoric form plays an important role in this process because the information provided by the anaphor varies according to form, ranging from forms that carry detailed semantic information (e.g. a repeated noun phrase) to those that carry only featural information (e.g. an overt pronoun marked for gender and number) to those that are phonetically unpronounced and carry no information (e.g. a null pronoun). Moreover, it is well-known that comprehenders expect anaphoric forms that provide enough information to identify the antecedent, but not additional unnecessary information (i.e. the maxim of quantity, cf. Grice, 1975).

In addition to taking a variety of more or less informative forms, anaphors can also refer to antecedents of a variety of phrasal categories such as noun phrases (NPs) or verb phrases (VPs). Although NP-anaphoric phrases (e.g. pronouns) have received the most study to date in the processing literature, researchers have recently also begun to examine the processing of VP-anaphoric phrases. In view of that, the studies in this dissertation focus on the processing of coordinate sentences involving VP-anaphoric phrases that contain either a null verb phrase (VP) anaphor (i.e. *did too*, with the null anaphor site marked with an underscore) as in (1) or an overt VP anaphor (i.e. *did it too*) as in (2).

(1) John kicked the ball and Bob *did _ too*.
(2) John kicked the ball and Bob *did it too*.
Although there are many constructions where both null and overt VP-anaphoric phrases\(^1\) are acceptable, the use of the null VP-anaphoric phrase is more restricted than that of the overt VP-anaphoric phrase. In the examples below, a star marks an ungrammatical construction. Note that while the null VP-anaphoric phrase can only be used to refer to linguistic antecedents in a parallel syntactic form (cf. (3a) vs. (4a) and (5a)), the overt VP-anaphoric phrase can be used to refer to linguistic antecedents in any syntactic form (as in (3b) and (4b)) or even to non-linguistic antecedents (as in (5b)) (see Murphy, 1985a; Hankamer & Sag, 1976; Sag & Hankamer, 1984, for further discussion).

(3) On Saturday, Bob hit a homerun.
   a. On Sunday, John did too.
   b. On Sunday, John did it too.

(4) On Saturday, a homerun was hit by Bob.
   a. *On Sunday, John did too.
   b. On Sunday, John did it too.

(5) (Event occurs in which Bob hits a homerun and then John hits one)
   a. *John did too!
   b. John did it too!

\(^1\)For ease of exposition, I refer to these phrases as the null VP-anaphoric phrase and the overt VP-anaphoric phrase. There is some debate regarding the linguistic analysis of the two forms, but, critically, both forms take as their referent the event specified in the antecedent VP.
Differentiation vs. Unification Theories

A number of broad-scope linguistic and psycholinguistic theories of anaphoric processing and interpretation offer accounts for the differential usage of the two VP-anaphoric forms. Two types of theories can be distinguished: differentiation theories and unification theories. Differentiation theories propose that the processing and interpretation of one class of anaphors (e.g. the null VP-anaphoric phrase) is qualitatively different from that of another class of anaphors (e.g. the overt VP anaphoric phrase). Specifically, the null VP-anaphoric phrase is claimed to be interpreted via a syntactic copying procedure (Frazier & Clifton, 2001, 2005; Sag & Hankamer, 1984) and the overt VP-anaphoric phrase via discourse reasoning (Sag & Hankamer, 1984). Thus, differentiation theories, like many other theories of cognitive processing, highlight a contrast between rapid, automatic, unconscious processing and slower, controlled, conscious processing.

Although the ideas underlying differentiation theories are intriguing, studies that have directly compared the processing of the two classes have not provided convincing evidence for a qualitative distinction (see Bélanger, 2004 for a critical review). In fact, typically, similar effects have been observed for both classes of

---

2 To date, most theoretical linguistic work on VP-anaphoric constructions has focused on the null VP-anaphoric phrase and has been aimed at building a theory of ellipsis (see Johnson, 2001; Kennedy, 2003, Merchant, 2001 for overviews); unfortunately, to date, no one theory of ellipsis has been completely successful. In addition, theories of ellipsis generally fail to make specific processing predictions regarding the comparison of null and overt VP-anaphoric phrases. Therefore, I do not discuss them in detail here.
anaphors (Martin & McElree, 2008; Murphy, 1985a, 1990). When differential effects have been observed, they have been attributable to factors other than the class of the anaphor. For example, although one study (Tanenhaus & Carlson, 1990) demonstrated differential effects, follow-up studies attributed these effects to the antecedent form or the specific task given to participants (see Mauner, Tanenhaus, & Carlson, 1995, and Murphy, 1990, for discussion). Moreover, in the only other studies that observed differential effects (Hestvik, Nordby, & Karlsen, 2005; Streb, Hennighausen, & Rösler, 2004), the anaphors chosen for the comparison were problematic. In comparing an NP anaphor of one class to a VP anaphor of the other class, these studies left open the possibility that the observed differences were due to the phrasal category rather than the class of the anaphor (see Callahan, 2008, for discussion). Furthermore, other studies that claimed that the interpretation of the null anaphoric phrase relies on distinct processes did not directly compare anaphoric phrases from the two classes. For example, although a lack of complexity effects in null VP-anaphoric constructions has been cited as evidence for a syntactic copying procedure (Frazier & Clifton, 2001), to our knowledge, these effects have never been directly investigated in overt VP-anaphoric constructions. Thus, although the lack of complexity effects could be related to the anaphoric form, it could also be related to some other aspect of these constructions such as that they involve parallelism, a factor that will be investigated in Chapter 5 of this dissertation.

In contrast to differentiation theories, unification theories propose that the same processing mechanism underlies the processing of all anaphoric forms (e.g.
In these theories, differential usage is related to the claim that, under optimal conditions, speakers produce anaphors that help the hearer identify the appropriate antecedent. Specifically, two factors interact to determine which anaphoric form is used: a) characteristics of the anaphor that determine how effective it is in uniquely identifying its antecedent and b) characteristics of the antecedent that determine its accessibility to the hearer. For example, anaphors vary in terms of their phonological size, the amount of information they carry, and whether they are ambiguous in the relevant context whereas antecedents vary in terms of their distance from the anaphor, their prominence in the relevant context, and the number of competitors they have. According to these theories, the fact that the null and overt VP-anaphoric phrases are used to refer to different types of antecedents is just one consequence of a general tendency for speakers to pair relatively less informative anaphors (e.g. the null VP-anaphoric phrase) with relatively more accessible antecedents and relatively more informative anaphors (e.g. the overt VP-anaphoric phrase) with relatively less accessible antecedents.

Although they do not posit a critical role for phrasal category per se, unification theories do note that the accessibility of the antecedent may be related to its phrasal category (e.g. a VP may be less accessible than an NP because it contains more information).
In the examples above, questionable acceptability is marked with a question mark. These examples illustrate that as the distance between the antecedent and the anaphor increases (and hence the relative accessibility of the antecedent decreases), the acceptability of the null VP-anaphoric phrase (cf. (6a) and (7a)), but not the overt VP-anaphoric phrase (cf. (6b) and (7b)) is affected.

The results of previous behavioral and electrophysiological investigations of the processing of different overt anaphoric forms are generally consistent with unification theories. Specifically, for a given anaphor, contexts involving more distant antecedents (e.g. Gordon & Scearce, 1995; Streb, Rösler, & Hennighausen, 1999; Streb, Hennighausen, & Rösler, 2004) or multiple possible antecedents (e.g. Garnham, 2001; van Berkum, Koorneef, Otten, & Nieuwland, 2006) are associated with increased processing. Moreover, in contexts involving prominent antecedents, over-
informative anaphoric forms elicit increased processing (e.g. Gordon, Grosz, & Gilliam, 1993; Swaab, Camblin, & Gordon, 2004). In addition, aside from having independent effects, characteristics of the antecedent often interact with characteristics of the anaphor (i.e. affecting some forms more than others, cf. Streb, Rösler, & Hennighausen, 1999).

Unification theories have also found cross-linguistic support in behavioral investigations of pro-drop languages such as Spanish and Italian where overt subject pronouns (e.g. as in Él comió.- “He ate.”) alternate with null subject pronouns (e.g. as in _ Comió.- “S/he ate.”). Previous studies (e.g. Alonso-Ovalle, Fernández-Solera, Frazier, & Clifton, 2002, Carminati, 2002) have demonstrated that comprehenders tend to interpret less informative forms as referring to highly accessible antecedents and more informative forms as referring to less accessible antecedents. Moreover, participants experience processing difficulty when these interpretative preferences are violated (i.e. when an over-informative form is used to refer to a highly accessible antecedent). Comparisons of null and overt subject pronouns are especially well-suited to evaluating the effects of anaphoric form because constructions can be generated that differ only in terms of the form of the anaphor (i.e. minimal pairs); however, the properties of the English language prevent this comparison (i.e. in English, null and overt pronouns do not freely alternate in subject position). For this reason, the studies in this dissertation instead compare another set of constructions that form minimal pairs (i.e. those involving null and overt VP-anaphoric phrases) to provide evidence that distinguishes between differentiation and unification theories. Specifically, the
studies described in Chapters 2 and 3 investigate the effects of anaphoric form on preference (as measured via a natural rating questionnaire) and processing (as measured via eye-tracking and electrophysiological techniques) and the study described in Chapter 4 investigates whether variation in anaphoric form is associated with variation in comprehension accuracy or interpretation.

All of the studies in Chapters 2, 3, and 4 compare null and overt VP-anaphoric phrases in contexts where both are grammatical and there is a single highly-accessible antecedent. As mentioned previously, such contexts are particularly interesting because the effects of anaphoric form are isolated from the effects of other factors (e.g. grammaticality, antecedent preference). Furthermore, these contexts are identical in terms of the syntactic and semantic information provided to the processor before and after the anaphor site. This feature is crucial for investigations of real-time processing using temporally-sensitive measures (e.g. eye-movements, event-related potentials (ERPs)). Finally, since the majority of previous psycholinguistic research involved NP-anaphoric constructions, comparisons of VP-anaphoric constructions provide an opportunity to evaluate the relevant theoretical claims in a new domain.

Parallelism

Although it is clear that anaphoric form is a key factor influencing differential usage and variation in anaphoric processing, other factors may also play an important role. As discussed earlier, the null, but not the overt VP-anaphoric phrase is restricted to contexts where the antecedent and anaphor are syntactically-parallel. This hints at a
relation between parallelism and anaphoric form. In fact, parallelism has long been proposed to play a key role in anaphoric interpretation (see Chambers & Smyth, 1998; Kehler, 2002; Smyth, 1994) with an anaphor preferentially interpreted as referring to an antecedent in a parallel syntactic position or with a parallel thematic role. Moreover, studies in English (e.g. Gordon, Grosz, & Gilliam, 1993) as well as in pro-drop languages (e.g. Alonso-Ovalle et al., 2002; Carminati, 2002) have suggested that anaphoric form interacts with this preference (i.e. while null subject pronouns are preferentially interpreted as referring to antecedents in subject position, overt subject pronouns are not),

Psycholinguistic researchers have only recently begun to study the effects of parallelism on processing; however, it is clear that, in contexts involving coordinate sentences, parallel conjuncts as in (3) are preferred to non-parallel conjuncts as in (4).

(8) Bob vacuumed the living room and Joe mopped the kitchen.
(9) Bob vacuumed the living room and the kitchen was mopped by Joe.

In addition, parallelism between the conjuncts facilitates processing for the second conjunct (the so-called parallelism effect). For example, one study that investigated the processing of coordinate sentences (Frazier et al., 1984) compared reading times for second conjuncts that were parallel or non-parallel to the first conjunct on one of

---

4 Parallelism effects occur in both comprehension and production (e.g. Bock, 1986); however, since the two types of effects are not identical, it is unclear whether the underlying mechanisms are the same (for discussion, see Frazier et al., 2000; Knoeferle & Crocker, accepted).
several dimensions (e.g. voice, attachment site, thematic roles, animacy). Parallelism on any of these dimensions resulted in faster reading times for the second conjunct.

Further studies have revealed that parallelism effects occur in both syntactically ambiguous and unambiguous contexts (Apel et al., 2007; Frazier et al., 1984; Frazier et al., 2000; Knoeferle & Crocker, accepted); however, they are restricted to conjoined structures (Frazier et al., 2000, Exp. 4; Apel et al., 2007, Exp. 2) where the conjunction implies parallelism (i.e. and, while) (Knoeferle, 2007; Shapiro & Hestvik, 1995). Moreover, parallelism effects appear to be driven by the processor rather than the grammar because these effects occur even for features that are not controlled by the grammar (e.g. the internal structure of conjuncts, Frazier et al., 2000, Exp. 3; semantic similarity, Knoeferle & Crocker, accepted, Exp. 3). In addition, since parallelism effects are observed for features that are not considered in an early structural analysis (e.g. animacy, semantic similarity) and are not observable in the earliest processing measures (cf. Frazier et al., 2000; Knoeferle & Crocker, accepted), the underlying mechanism is likely not the facilitation of specific early syntax-based parsing routines.

The study described in Chapter 5 of this dissertation investigates the underlying processes and timecourse of parallelism effects in coordinate sentences. Given the relationship between parallelism and anaphora, this study evaluates a hypothesis (i.e. the sustained reactivation hypothesis) that proposes that a well-known context-driven memory-based reactivation process operative in anaphoric processing (cf. Nicol & Swinney, 2002) underlies parallelism effects. Essentially, the sustained
reactivation hypothesis proposes that, when a cue that implies parallelism is encountered, previously processed material is reactivated. This material then remains active throughout the second conjunct until it can be integrated with material from the second conjunct, resulting in facilitated processing for linguistic material that is parallel to the reactivated material on any dimension. Specifically, the sustained reactivation hypothesis claims that, in coordinate structures, encountering a conjunction that implies parallelism (e.g. *and*, see Knoeferle, 2007) results in the reactivation of the previous conjunct; nonetheless, it is possible that other linguistic or non-linguistic cues to parallelism might also be sufficient to elicit reactivation.

Since, in natural language, coordinate sentences are the prototypical environment for parallelism as well as a highly typical environment for VP-anaphoric phrases, the study in Chapter 5 evaluates the activation of a parallel element (i.e. a verb) throughout two *and*-conjoined sentences involving a VP-anaphoric phrase. Verbs are a good candidate for testing sentence-level parallelism effects because they play a crucial syntactic and semantic role in the sentence and activate not only their meaning (e.g. de Goede, 2006; Shapiro et al., in prep) but also typically associated arguments, argument structures, and thematic role information (Altmann & Kamide, 1999; Boland et al., 1990; Ferretti et al., 2001; Friederici & Frisch, 2000; McRae et al., 1997; McRae et al., 1998; Shapiro et al., 1987, 1989). Moreover, comprehenders have been demonstrated to use the information provided by verbs to facilitate processing by generating expectations about upcoming context (e.g. Altmann & Kamide, 1999; Shapiro et al., 1993; Trueswell & Kim, 1998; Trueswell et al., 1993).
Contributions of this dissertation

Collectively, the studies contained in this dissertation explore the effects of the interrelated factors of anaphoric form and parallelism on processing and interpretation. By testing fully grammatical constructions that form minimal pairs, this research isolates variation due to anaphoric form from variation in acceptability or interpretative preference, unlike the majority of previous research. In addition, by focusing on coordinate sentences involving VP anaphors, these studies extend previous psycholinguistic and neurolinguistic research into the relatively unexplored domain of verbal anaphora. Finally, the sustained reactivation hypothesis complements theories of anaphora by formulating clear testable predictions about the processing of coordinate sentences. By drawing theoretical connections between two distinct literatures, this work proposes, in the spirit of parsimony, that seemingly disparate phenomena may result from common underlying processes. In so doing, this work achieves broad relevance and contributes to the development of wide-scope theories of language processing.
References


knowledge to resolving discourse roles: Bonding and resolution. *Journal of Memory and Language, 42,* 526-544.


Knoeferle, P. (2007). Eye tracking parallelism effects in coordinate constructions: the effects of “and”, “but” and “while”. Poster presented at the 20th Annual CUNY, La Jolla, USA.


Chapter 2

Effects of variation in anaphoric form on the processing and interpretation of VP-anaphoric constructions: An eye-tracking study
Abstract

This study investigates the effects of variation in anaphoric form by comparing the processing and interpretation of a null VP-anaphoric phrase (i.e. *did too*) and an overt VP-anaphoric phrase (e.g. *did it too*) in contexts where both forms are grammatical. Unification theories predict that in contexts involving a highly accessible antecedent, participants prefer the least informative grammatical anaphoric form and that over-informative forms will be associated with processing difficulty. In Experiment 1, participants rated the null VP-anaphoric phrase as more natural than the overt VP-anaphoric phrase in contexts involving a highly accessible antecedent. In Experiment 2, participants’ eye-movements were monitored while reading the same sentences. Longer first-pass and total reading times were observed for the anaphor region when it contained an overt VP-anaphoric phrase. This effect remained when length differences were factored out. Overall, the results from both experiments support unification theories. Moreover, the eye-movement results suggest that the processing difficulty related to over-informative anaphors is best conceptualized as a violation of the expectation for a less informative anaphor.
The term, *anaphora*, denotes the relationship between linguistic phrases that refer to the same real-world entity or concept (i.e. the antecedent and the anaphor). Anaphors appear in a variety of forms ranging from those that carry detailed semantic information (e.g. a repeated noun phrase) to those that carry only featural information (e.g. an overt pronoun marked for gender and number) to those that are unpronounced and carry no information (e.g. a null pronoun). Anaphors also correspond to a variety of phrasal categories such as noun phrases (NPs), verb phrases (VPs), or even whole inflectional phrases (IPs) (i.e. clauses). Although NP-anaphoric phrases (e.g. pronouns) have received the most study to date in the processing literature, researchers have recently begun to examine the processing of VP-anaphoric and IP-anaphoric phrases in more depth.

The research in the current study focuses on sentences involving VP-anaphoric phrases that contain either a null verb phrase (VP) anaphor (i.e. *did too*) or an overt VP anaphor (i.e. *did it too*)\(^1\) and investigates the effect of this variation in anaphoric form on participants’ conscious naturalness ratings as well as their eye-movements during reading. Although there is significant overlap, the null and overt VP-anaphoric

\(^1\) For ease of exposition, we refer to these phrases as the null VP-anaphoric phrase and the overt VP-anaphoric phrase. Although there is some debate regarding the linguistic analysis, critically, both forms take as their referent the event specified in the antecedent VP.
phrases of interest in the current study are not used completely interchangeably in natural language. In particular, the two phrases differ in the types of antecedents to which they can refer.

(1) On Saturday, Bob hit a homerun.
   a. On Sunday, John did too.
   b. On Sunday, John did it too.

(2) On Saturday, a homerun was hit by Bob.
   a. *On Sunday, John did too.
   b. On Sunday, John did it too.

(3) (Event occurs in which Bob hits a homerun and then John hits one)
   a. *John did too!
   b. John did it too!

While the null VP-anaphoric phrase can only be used to refer to linguistic antecedents in a parallel syntactic form (as in the comparison of (1a) vs. (2a) and (3a) with ungrammatical constructions marked with a star), the overt VP-anaphoric phrase can be used to refer to linguistic antecedents in any syntactic form (as in (1b) and (2b)) or even to non-linguistic antecedents (as in (3b)) (see Murphy, 1985a; Hankamer & Sag, 1976; Sag & Hankamer, 1984, for further discussion).

A number of broad-scope linguistic and psycholinguistic theories of anaphoric processing and interpretation offer accounts for the differential usage of the VP-
anaphoric forms in the current study. These theories basically fall into two classes: differentiation theories and unification theories. Differentiation theories propose that the processing and interpretation of one class of anaphors (e.g. the null VP-anaphoric phrase in the current study) is qualitatively different from that of another class of anaphors (e.g. the overt VP anaphoric phrase in the current study). In particular, the processing of the null VP-anaphoric phrase is claimed to involve a syntactic copying procedure (Frazier & Clifton, 2001, 2005; Sag & Hankamer, 1984). Although this idea is intriguing, the results of direct comparisons of the two classes of anaphors have not strongly supported a qualitative processing distinction (see Bélanger, 2004 for a critical review). Generally, similar effects have been observed for both classes of anaphors (Callahan, Nicol, Love, & Swinney, 2007, submitted; Martin & McElree, 2008; Murphy, 1985a, 1990) or differential effects have been attributable to factors other than the class of the anaphor. Specifically, although one study (Tanenhaus & Carlson, 1990) demonstrated differential effects, follow-up studies attributed these effects to the antecedent form or the specific task given to participants (see Mauner, 2

---

2 To date, most of the theoretical linguistic work on VP-anaphoric constructions has focused exclusively on VP-elliptical constructions (e.g. the null VP-anaphoric phrase in the current study) (see Johnson, 2001; Kennedy, 2003, Merchant, 2001 for overviews); however, to date, no one theory of ellipsis has been completely successful and, for the most part, these theories are silent as to underlying processing mechanisms. Since these theories fail to make specific processing predictions regarding the comparison of null and overt VP-anaphoric phrases, we do not discuss them in detail here.

3 Frazier & Clifton (2001) cited a lack of complexity effects in null VP-anaphoric constructions as evidence for a syntactic copying procedure; however, to our knowledge, complexity effects have never been directly investigated in overt VP-anaphoric constructions. Thus, it is unclear whether the lack of complexity effects is related to the anaphoric form or to some other characteristic of these constructions.
Tanenhaus, & Carlson, 1995, and Murphy, 1990, for discussion). Likewise, the only other studies that observed differential effects (Hestvik, Nordby, & Karlsen, 2005; Streb, Hennighausen, & Rösler, 2004) compared an NP anaphor of one class to a VP anaphor of the other class, leaving open the possibility that the observed differences were due to the phrasal category rather than the class of the anaphor (see Callahan, 2008, and Callahan et al., submitted, for discussion).

In contrast to differentiation theories, unification theories propose that the same processing mechanism underlies the processing of all anaphoric forms (e.g. Almor, 1998; Ariel, 1990, 1991; Givón, 1983; Gordon & Hendrick, 1998; Grosz, Joshi, & Weinstein, 1995; Murphy, 1985b; Walker, Joshi, & Prince, 1998). In order to account for differential usage, these theories make reference to a pragmatic principle that states that, under optimal conditions, speakers produce anaphors that help the hearer identify the appropriate antecedent (see Grice, 1975, for details and discussion of the pragmatic principles underlying conversational speech). Specifically, unification theories claim that two factors interact to determine the anaphoric form used: a) characteristics of the anaphor that determine how effective it is in uniquely identifying its antecedent and b) characteristics of the antecedent that determine its accessibility to the hearer. Anaphors vary in terms of their phonological size, the amount of information they carry, and whether they are ambiguous in the relevant context. Antecedents vary in terms of their distance from the anaphor, their prominence in the relevant context, and the number of competitors they have. Fundamentally, unification theories propose that, for a given context, comprehenders
are sensitive to the relative frequency with which speakers use different anaphoric forms. We will delve more deeply into this notion of relative frequency in the General Discussion, but for now we simply note that speakers tend to pair less informative anaphors (e.g. the null VP-anaphoric phrase in the current study) with more accessible antecedents and more informative anaphors (e.g. the overt VP-anaphoric phrase in the current study) with less accessible antecedents.

(4) John raked the leaves in the backyard.
    Later, Bill did too.

(5) John raked the leaves in the backyard.
    This was much more fun than studying for exams?
    Later, Bill did too.

(adapted from Bélanger, 2004)

Comprehenders’ sensitivity to this tendency is illustrated by the comparison between the contexts in (4) and (5) above where increasing the distance between the antecedent and the anaphor decreases the accessibility of the antecedent and hence the acceptability of the null VP-anaphoric phrase (the question mark denotes questionable acceptability).

Unlike differentiation theories, unification theories and models have found general support in investigations of the processing and interpretation of different overt anaphoric forms in English (e.g. Almor, 1998; Gordon, Grosz, & Gilliam, 1993). In
addition and of even greater relevance to the current study, these models have also
found support in investigations of the processing and interpretation of null and overt
subject pronouns in pro-drop languages (e.g. Spanish and Italian, see Alonso-Ovalle,
Fernández-Solera, Frazier, & Clifton, 2002, and especially Carminati, 2002 for details
and further discussion). Participants tend to interpret less informative forms as
referring to highly accessible antecedents and more informative forms as referring to
less accessible antecedents. Moreover, participants experience processing difficulty
when these interpretative preferences are violated (i.e. when an over-informative form
is used to refer to a highly accessible antecedent).

The current study focuses on investigating the claims of unification theories in
a new domain: namely, the processing of VP-anaphoric constructions. Null VP-
anaphoric phrases (e.g. *did too*, as in (6) below) are compared with overt VP-
anaphoric phrases (e.g. *did it too*, as in (7) below) in contexts where both forms are grammatical.

(6) The typist corrected the error and the editor did too using red pen.

(7) The typist corrected the error and the editor did it too using red pen.

Investigating contexts in which both VP-anaphoric phrases are grammatical allows
effects related to the form of the anaphor to be isolated from effects related to
(un)grammaticality. In addition, by presenting both forms in the same context, the
characteristics of the anaphor can be varied while the characteristics of the antecedent
are kept constant. The antecedent in the contexts tested in the current study, (e.g. *corrected the error*), is highly accessible (i.e. it is close to the anaphor, prominent, and has no competitor). Although unification theories propose that the same processing mechanisms underlie the processing of all anaphoric forms, they predict variation in preference and ease of processing based on an interaction between the characteristics of the antecedent and the anaphor. Thus, these theories predict a preference for the least informative anaphor in these contexts (i.e. the null VP-anaphoric phrase).

Experiment 1 tests this prediction by collecting naturalness ratings for the two anaphoric forms. If unification theories are correct, then the null VP-anaphoric phrase should be rated as more natural than the overt VP-anaphoric phrase. In addition to predicting a preference for the less informative anaphoric form in contexts with highly accessible antecedents, unification theories predict that over-informative anaphoric forms in these contexts (e.g. the overt VP-anaphoric phrase in the current study) will be associated with processing difficulty. Experiment 2 tests this prediction by examining the eye-movements of participants while reading sentences involving the two anaphoric phrases. If the prediction is correct, the overt VP-anaphoric phrase should be associated with increased reading times and other indicators of processing difficulty.⁴

---

⁴ Of course, since the overt VP-anaphoric phrase contains an additional word, it will be necessary to factor out the possible confound of length. We will address this by performing the analyses on residual reading times as well as raw reading times. Please see the results section of Experiment 2 for more details.
Experiment 1- Naturalness Rating Questionnaire

Methods

Participants

The participants in Experiment 1 were recruited from the University of California, San Diego and received course credit for their participation. All participants were right-handed, neurologically-unimpaired, and had normal hearing and normal or corrected-to-normal vision. All participants were also native English speakers with no exposure to a second language before age six. Forty-five participants were initially tested; however, one participant was excluded from further analysis for not following instructions. Therefore, 44 participants were included in further analysis (15 male, 29 female; mean age = 20.31, SD = 1.81).

Materials

For the critical experimental items, 56 items of similar form were constructed. Each experimental item consisted of two sentences joined by the conjunction, \textit{and}. The first conjunct always consisted of a subject and a VP and the second conjunct always consisted of a different subject, an anaphoric phrase, and a final phrase. For the two VP-anaphoric conditions (please see Table 1 for sample sentences and the Appendix for a full list), the anaphoric phrase in the second conjunct was either a null VP-anaphoric phrase (e.g. \textit{did too}) or an overt VP-anaphoric phrase (e.g. \textit{did it too}). In
anticipation of the eye-tracking experiment (Experiment 2), two NP-anaphoric conditions were also included to replicate a standardly observed effect of the frequency of a word on eye-tracking measures and to investigate previous observations of a reverse word frequency effect at a corresponding anaphor (Heine, Tamm, Hofmann, Hutzler, & Jacobs, 2006; van Gompel & Majid, 2004; but c.f. Simner & Smyth, 1999). In these two conditions (please see Table 1 for a sample sentence and the Appendix for a full list), the first conjunct consisted of a subject and a VP containing either a comparatively high-frequency object (as in (8) below; Mean Kucera-Francis frequency = 114, e.g. error) or a comparatively low-frequency object (as in (9) below) Mean Kucera-Francis frequency = 2.91, e.g. digit) and the second conjunct consisted of a different subject and a repeated verb with an NP-anaphoric object (i.e. it).

(8) The typist corrected the error and the editor corrected it too using red pen.

(9) The typist corrected the digit and the editor corrected it too using red pen.

Although the NP-anaphoric conditions were mainly of interest in the eye-tracking experiment, they were also included in the naturalness rating experiment to ensure that the frequency manipulation did not affect participants’ ratings of the naturalness of the entire sentence and so that the same set of items would be tested in both experiments.

---

5 In keeping with the literature on frequency effects, the Kucera-Francis frequency for the higher-frequency object in a particular item was at least ten times that of the lower-frequency object. For ease of exposition, we will refer to these phrases as either high-frequency or low-frequency.
For purposes of analysis, the NP-anaphoric items will be considered separately from the VP-anaphoric items.

To minimize the possibility that participants would develop strategies during testing, participants were also presented with 40 filler items of varying structures that did not involve VP-anaphoric phrases of any kind. The 56 experimental and 40 filler items were assembled into a single pseudo-randomized list with the constraint that no experimental condition ever appeared twice in a row. Four counterbalanced versions of the list were created by systematically varying the condition across the lists for the experimental items. Although each item only appeared once in a given list, across lists each item appeared in each condition. Participants responded to items in all four conditions, but only read any given item once. Thus, over the course of the experiment, each participant was presented with 14 items from each of the two critical VP-anaphoric conditions (Null VP-Anaphoric Phrase (NVP) and Overt VP-Anaphoric Phrase (OVP)), 14 items from each of the two NP-anaphoric conditions (High-frequency NP (HFNP) and Low-Frequency NP (LFNP)), and 40 filler items.

Procedure

Before beginning the experiment, participants gave informed consent and completed a questionnaire about their language use and language experience. Next, the participants received instructions for the naturalness rating task. Participants were instructed to read each sentence and rate how natural the sentence sounded using a
Like Likert scale ranging from 1 (very natural) to 5 (very awkward). The entire experiment lasted less than one hour.

**Experimental Design**

For presentation purposes, the critical conditions were considered as part of a single within-subjects repeated-measures design with four conditions (NVP, OVP, HFNP, LFNP); however, for analysis, the two VP-anaphoric conditions (NVP, OVP) were considered separately from the two NP-anaphoric conditions (HFNP, LFNP).

**Results**

Mean ratings for the four experimental conditions ranged between 3.5 and 3.7 where a rating of 3 represented “OK” and a rating of 4 represented “Somewhat awkward”. An omnibus one-way ANOVA showed a significant effect of condition ($F_1(3, 129) = 4.94, p = .003, F_2(3, 165) = 9.09, p < .001$). The theoretical hypotheses of interest were then evaluated via paired two-tailed t-tests. Ratings for the high-frequency NP-anaphoric condition (Mean = 3.5, SD = .7) and the low-frequency NP-anaphoric condition (Mean = 3.5, SD = .7) were not significantly different ($t_1(43) = .005, p = 1, t_2(55) = .07, p = .95$) indicating that the frequency of the object of the VP in the first conjunct did not affect participants’ ratings of the naturalness of the sentence. Critically, ratings for the Overt VP-anaphoric condition (Mean = 3.7, SD = .7) were significantly higher than for the Null VP-anaphoric condition (Mean = 3.5,
SD = .7), indicating that participants rated the Overt VP-anaphoric condition as less natural ($t_1(43) = 3.44, p = .001, t_2(55) = 3.92, p < .001$).

Discussion

The results of the naturalness rating study support the claim of unification theories that comprehenders are sensitive to the fact that highly accessible antecedents are typically referred to by the least informative anaphor possible. In Experiment 1 involving contexts where both VP-anaphoric phrases are grammatically possible, participants displayed a preference for the least informative anaphoric form (i.e. the null VP-anaphoric phrase). The preference for the null VP-anaphoric phrase observed in this experiment is also consistent with a previous naturalness rating study conducted in our lab (see Callahan et al., submitted, for details) using very similar stimuli (e.g. *The novelist edited the book one last time and the publisher did (it) too.*).

Naturalness ratings, while informative about participants’ preferences, represent a final conscious preference for one type of anaphoric phrase over another; yet, unification theories also claim that violations of these preferences are associated with real-time processing difficulty. Investigating this theoretical claim requires a methodology that is sensitive to real-time language processing and has excellent temporal resolution. To this end, Experiment 2 utilizes eye-tracking techniques to investigate the real-time processing of the same stimuli from Experiment 1.
Experiment 2- Eye-tracking Study

Methods

Participants

The participants in Experiment 2 were recruited from the University of Arizona, Tucson and received course credit for their participation. All participants were right-handed, neurologically-unimpaired, and had normal vision or vision that was corrected-to-normal via soft contact lenses. All participants were also native English speakers with no exposure to a second language before age six. Thirty-six participants were tested (11 male, 25 female; mean age = 21.61, SD = 6.44).

Materials

The materials in Experiment 2 were identical to those in Experiment 1. See the Appendix for a full list. As in Experiment 1, over the course of the experiment, each participant was presented with 14 items from each of the two critical VP-anaphoric conditions (Null VP-Anaphoric Phrase (NVP) and Overt VP=Anaphoric Phrase (OVP)), 14 items from each of the two NP-anaphoric conditions (High Frequency NP (HFNP) and Low-Frequency NP (LFNP)), and 40 filler items. Please see Table 1 for sample sentences. In Experiment 2, participants were also presented with 12 practice items and 32 yes/no comprehension questions randomly distributed throughout the list to ensure participants were understanding the sentences, e.g. ITEM: *The boss printed the invoice and the accountant printed it too using red ink.* QUESTION: *Was the ink
blue?. There were an equal number of yes/no responses to the comprehension questions across the experiment.

Procedure

Before beginning the experiment, participants gave informed consent and completed a questionnaire about their language use and language experience. Next, the participants received instructions for the reading task. Participants were instructed to read each sentence carefully for understanding but otherwise to read normally. Participants were also informed that some of the sentences would be followed by a yes/no question and that they should indicate their response using the two buttons on the response box in front of them. Once participants indicated that they understood the instructions, the experiment began.

Participants’ eye movements were recorded with a Dr. Bouis oculometer by reflecting an 8.5mA beam of infrared light off the right eye (although viewing was binocular). The eye-tracker monitored gaze location with millisecond resolution to determine fixation onset and duration. Sentences were presented as single lines of text on a 21-inch CRT monitor. The distance from the eye to the monitor was approximately 60 cm allowing for single character resolution. A bite plate and headrest were used to minimize head movements. The eye-tracker was calibrated at the beginning of the experiment and then recalibrated after every four trials. The entire experiment lasted less than one hour.
Experimental Design

The experimental design was the same as in Experiment 1.

Results

Three critical regions were identified for analysis: 1) the VP1 region (e.g. corrected the error), 2) the anaphor region (e.g. did too), and 3) the post-anaphor region (e.g. using red pen.). Please see Table 1 for sample sentences with the critical regions for analysis marked. For each critical region, two standard eye-tracking measures were examined: 1) first-pass time and 2) total reading time. First-pass time refers to the sum of all fixations in a region from the first fixation in that region until the first fixation in any other region. Total reading time refers to the sum of all fixations in a region. The first-pass measure reflects early processing while the total time measure reflects not only early but also later processing including re-processing (i.e. re-reading a region after leaving the region).

Omnibus one-way ANOVAs in the critical regions showed a significant effect of condition on first-pass times in the VP1 region ($F_1(3, 105) = 8.80, p < .001, F_2(3, 165) = 7.54, p < .001$), in the anaphor region ($F_1(3, 105) = 61.69, p < .001, F_2(3, 165) = 65.57, p < .001$), and in the post-anaphor region for the subject analysis ($F_1(3, 105) = 3.15, p = .03$), but not the item analysis ($F_2(3, 165) = 1.20, p = .31$). In addition, 6

---

6 Go-past times (sometimes called regression path time) were also examined. The pattern was nearly identical to the pattern for total times and did not affect the interpretation of the findings, so for the sake of brevity, we will not present these results.
significant effects of condition on total time were observed in the VP1 region ($F_1(3, 105) = 12.00, p < .001, F_2(3, 165) = 6.46, p < .001$), the anaphor region ($F_1(3, 105) = 38.31, p < .001, F_2(3, 165) = 47.39, p < .001$), and the post-anaphor region ($F_1(3, 105) = 7.82, p < .001, F_2(3, 165) = 6.95, p < .001$). Since the theoretical hypotheses of interest pertained only to the comparison between the two NP-anaphoric conditions and the comparison between the two VP-anaphoric conditions, the results of the omnibus ANOVAs will not be discussed further.

NP-Anaphoric Conditions

First, we examine the reading times for the NP-anaphoric items. These items were included in the design to replicate a standardly observed effect of longer reading times for low-frequency as compared to high-frequency words (e.g. Inhoff & Rayner, 1986; Rayner & Duffy, 1986) and to further investigate a previous report of faster first fixation and first-pass times on the word(s) following an anaphor that refers to a low-frequency as compared to a high-frequency antecedent (van Gompel & Majid, 2004). Based on this previous study, the theoretical hypotheses of interest were evaluated via paired one-tailed t-tests. Please see Table 2 for means and standard errors.
As shown in Figure 1, in the NP-anaphoric conditions, the VP1 region was associated with longer first-pass times when it contained a low-frequency object (e.g. digit) than when it contained a high-frequency object (e.g. error) \( t_1(35) = 4.49, p < .001, \ t_2(55) = 3.66, p = .001 \). The same pattern was observed in total times \( t_1(35) = 5.20, p = .000, \ t_2(55) = 3.09, p = .002 \). In the anaphor region, the comparison of the two NP-anaphoric conditions did not reach significance for either measure (all \( p > .05 \)). In contrast, the post-anaphor region was associated with longer first-pass times when it followed an anaphor that referred to a high-frequency as compared to a low-frequency antecedent (LFNP: 599 ms, HFNP: 645 ms). This difference was significant in the subject analysis \( t_1(35) = 2.54, p = .01 \) and marginal in the item analysis \( t_2(55) = 1.50, p = .07 \). Although the numerical difference between the HFNP and LFNP conditions was in the same direction for total times (LFNP: 706 ms, HFNP: 722 ms), this difference failed to reach significance in total times (all \( p > .05 \)). The marginal significance of the effect in the items analysis of first-pass times is probably due to the
fact that the post-anaphor region contained more words and was associated with more variability in reading times in the current study than in the previous study (van Gompel & Majid, 2004). In fact, when one item with an overall mean that exceeded the average item mean by more than two standard deviations was excluded, the effect in the post-anaphor region reached standard levels of significance in both the subjects and items analysis for first-pass times ($t_1(35) = 2.76, p = .005, t_2(54) = 1.70, p = .05$). The effect still failed to reach significance in total times (all $p > .05$).

**VP-Anaphoric Conditions**

The replication of previous research provided by the NP-anaphoric items suggests that participants are paying attention and performing appropriately. Next, we turn to the critical theoretical hypotheses under investigation in this study and examine the reading times for the VP-anaphoric items. Please see Table 3 for means and standard errors. In the two VP-anaphoric conditions, the VP1 region was identical (i.e. there was no manipulation of the frequency of the object) and thus no significant difference was predicted in this region. Indeed, as Figure 2 shows, the two VP-anaphoric conditions did not differ in the VP1 region in first-pass or total times (all $p > .05$). In contrast, the anaphor region was associated with longer first-pass times when it contained an overt VP-anaphoric phrase (i.e. *did it too*) as compared to a null VP-anaphoric phrase (i.e. *did too*) ($t_1(35) = 10.19, p < .001, t_2(55) = 9.38, p < .001$). The same pattern was observed in total times ($t_1(35) = 7.71, p < .001, t_2(55) = 9.60, p < .001$). In addition, the final region was associated with marginally longer total times...
when it followed a null VP-anaphoric phrase than when it followed an overt VP-
anaphoric phrase ($t_1(35) = 1.99, p = .06, t_2(55) = 2.04, p = .05$). This effect was not
observed in first-pass times (all $p > .05$).

In this study, the anaphor region contained more characters in the Overt VP-
anaphoric condition than in the Null VP-anaphoric condition. This length confound is
unavoidable since the comparison of interest is between a phrase containing a null
anaphor and one containing an overt anaphor. Thus, it is important to rule out the
possibility that length rather than anaphor form is responsible for the longer first-pass
and total reading times observed in the Overt VP-anaphoric condition. One way to
address this question would be to divide the time for the region by the number of
characters yielding an estimate of the average reading time per character; however,
previous research (Ferreira & Clifton, 1986; Trueswell, Tanenhaus & Garnsey, 1994)
demonstrated that this method is inaccurate under certain circumstances and suggested
an alternate calculation that yields an estimate of residual reading time. Following this
research, regression analyses were performed on the eye-movement data for each
subject to generate a linear equation that best approximated the relationship between
number of characters and reading time for each individual subject. These equations
were then used to generate predicted times for each region based on the number of
characters in the region. Residual reading times for each region were calculated by
subtracting the predicted times from the observed reading times. Thus, positive
residual reading times indicate that participants were slower to read that region than
would be predicted by the length of the region and their average reading speed.
Negative residual reading times indicate that participants were faster than would be
predicted. The calculated mean residual first-pass and total times for the anaphor
region were then entered into paired two-tailed t-tests.

Residual reading times for the anaphor region showed the same pattern of
effects as raw reading times. Specifically, the anaphor region was associated with
longer residual first-pass times when it contained an overt VP-anaphoric phrase (i.e.
did it too) than when it contained a null VP-anaphoric phrase (i.e. did too) (OVP: 44,
NVP: -34; \( t_1(35) = 5.15, p < .001, t_2(55) = 6.17, p < .001 \)). The same pattern was
observed in total times (OVP: 60, NVP: -37; \( t_1(35) = 3.98, p < .001, t_2(55) = 5.69, p <
.001 \)). Thus, the analysis of residual reading times is consistent with the analysis of
raw reading times and suggests that the observed differences between the VP-
anaphoric conditions are not due to differences in length.
General Discussion

NP-Anaphoric Conditions

Consistent with previous investigations of the effects of word frequency on reading times (e.g. Inhoff & Rayner, 1986; Rayner & Duffy, 1986; Schilling, Rayner, & Chumbley, 1998), longer first-pass and total reading times were observed in the VP1 region when it contained a low-frequency as compared to a high-frequency object. A previous investigation of the effect of antecedent frequency on anaphor processing (van Gompel & Majid, 2004) also revealed longer first fixation and first pass (but not total) times for the word(s) following an anaphor with a high-frequency as compared to a low-frequency antecedent. Van Gompel and Majid interpreted this effect as evidence for a saliency-based account (e.g. Pynte & Colonna, 2000) by which infrequent antecedents are easier to access than frequent antecedents because they are more salient, attract more attention, and require more processing effort. Consistent with this previous study, the current study revealed longer first-pass (but not total) times in the post-anaphor region. Thus, the current study fits with previous research demonstrating that although high-frequency words are typically processed more easily than low-frequency words, this effect is reversed at a corresponding anaphor (i.e. anaphors that refer to low-frequency antecedents are processed more easily) (Heine et al., 2006; van Gompel & Majid, 2004; but c.f. Simner & Smyth, 1999). More generally, these findings also fit with observations that increased processing during encoding results in easier retrieval (e.g. O’Brien & Myers, 1985; Tyler, Hertel, McCallum, & Ellis, 1979).
VP-Anaphoric Conditions

In the naturalness rating questionnaire in Experiment 1, sentences containing an overt VP-anaphoric phrase were rated as less natural than sentences containing a null VP-anaphoric phrase. This finding is consistent with the general claim of unification theories that in contexts where multiple anaphoric forms are grammatical, over-informative anaphors are dispreferred. This claim is based in the idea that speakers are ideally as informative as necessary, but not over-informative (i.e. the Gricean maxim of quantity (Grice, 1975)). Based on this principle, comprehenders expect speakers to produce the least informative anaphoric form possible in these contexts (i.e. the null VP-anaphoric phrase) and thus find the over-informative form less natural.

On-line processing investigations of unification theories have revealed that in addition to displaying a conscious preference for less informative anaphors, comprehenders experience processing difficulty when encountering over-informative anaphors (Gordon et al., 1993; Carminati, 2002). In the eye-tracking experiment in the current study (Experiment 2), this processing difficulty was manifested as longer first-pass and total reading times for the anaphor region when it contained an overt VP-anaphoric phrase as compared to a null VP-anaphoric phrase. This difference was observed even when length differences between the conditions in this region were factored out by the calculation of residual reading times. When considered in the context of the results from the naturalness rating questionnaire, the eye-tracking
results suggest that encountering an anaphoric phrase in a form that contradicts an expectation elicits additional processing.

Since both anaphoric forms were syntactically and semantically legal in the constructions tested in this study, the increased reading time cannot be related to the processing of a syntactic or semantic violation. Furthermore, since the sentences are not ambiguous and there is only one possible antecedent for the VP-anaphoric phrase, the effect cannot be related to the need to revise some previous incorrect syntactic analysis (as in garden-path sentences, for example) or incorrect antecedent selection (as has been suggested for comparisons of null and overt subject pronouns in pro-drop languages (e.g. Carminati, 2002)). Instead, we suggest that the effect does not reflect a reanalysis process at all, but rather a violation of the expectation for the less informative (and more frequent) form. The fact that the effect in the current study is observable in the earliest measure, first-pass time, and in the earliest region, the anaphor region, also suggests an early expectation-driven effect.

Consistent with an expectation-based explanation, the only ERP study to date that compared null and overt VP-anaphoric constructions (Callahan et al., submitted) observed an early centro-posterior negativity that peaked approximately 100ms after the onset of the overt anaphor (i.e. *it*). The authors interpreted that effect as reflecting the deployment of processing resources to an unexpected continuation. In addition, that study also observed a frontal positivity between 300 and 500ms after the onset of the last word of the VP-anaphoric phrase (i.e. *too*) that was greater following a null than an overt VP-anaphoric phrase. Interestingly, the eye-tracking experiment in the
current study also revealed a marginal effect of longer reading times in the post-anaphor region following a null as compared to an overt VP-anaphoric phrase. We interpret this reversal pattern as reflecting a trade-off in resource allocation (i.e. pay now vs. pay later). When comprehenders encounter an overt VP-anaphoric phrase that violates their expectations, they devote additional processing resources to that phrase and begin the process of identifying, retrieving, and integrating the antecedent information. This resource allocation is observable in the increased reading time in the anaphor region in the current eye-tracking study and in the early negativity in the ERP study. In contrast, when comprehenders encounter a null VP-anaphoric phrase, this process is not begun until the word after the anaphor is encountered and the null anaphor site is detected. This resource allocation is observable in the marginally longer reading times for the post-anaphor region in the eye-tracking study and the frontal positivity in the ERP study. Although the results of these two studies are suggestive, further research in this domain is necessary to support and refine this interpretation of the later effect, especially given that the effect in the post-anaphor region just missed significance in the current eye-tracking study.

Overall, the results of the naturalness rating, eye-tracking, and ERP experiments described above suggest that comprehenders are sensitive to VP-anaphoric form both in terms of conscious preference and unconscious processing. In contexts where multiple anaphoric forms are grammatical, comprehenders expect speakers to produce the least informative anaphoric form possible and violations of this expectation elicit additional processing. This sensitivity to anaphoric form is
consistent with the predictions of unification theories (e.g. Almor, 1998; Ariel, 1990, 1991; Givón, 1983; Gordon & Hendrick, 1998; Murphy, 1985b) because, essentially, these theories propose that, for a given context, comprehenders’ expectations reflect the relative frequency with which speakers use different anaphoric forms (i.e. the relative context-specific frequency).

It is also well-known that comprehenders are sensitive to the relative frequency of linguistic terms across all contexts (i.e. the relative global frequency). Given that, in general, little is known about the frequencies of the null and overt VP-anaphoric phrase, one important question for future research is whether the expectation-driven effects observed in the current study reflect global or context-specific frequency. The naturalness rating data (as well as our intuition) suggests that the null form is used more frequently than the overt form in the contexts tested in the current study; however, corpus analyses will be crucial to assessing not only the relative frequency of the two forms in contexts where both are grammatical but also the relative global frequency of the two forms. On the basis of such corpus analyses, contexts can be identified where both forms are grammatical, but one or the other is more frequent. Direct comparisons of the two phrases in these particular contexts could then disentangle the effects of global and context-specific frequency on processing variation related to anaphoric form. We plan to pursue this line of research in the future. Moreover, we suggest that, in general, further research on the processing and interpretation of different anaphoric forms in contexts where multiple anaphoric forms
are grammatical will be essential to evaluating the effects of form on anaphoric
processing and interpretation.

This chapter, in full, has been submitted for publication as it may appear in
Journal of Experimental Psychology: Learning, Memory, & Cognition, Callahan,
S.M., Nicol, J., & Love, T. The dissertation author was the primary investigator and
author of this paper.
References


Appendix

NP-Anaphoric Conditions (HFNP/LFNP)

1. The policeman found the (theater/cabaret) and the fireman found it too using a map.
2. The kid ate the (cereal/caviar) and the teenager ate it too because he was hungry.
3. The boss printed the (article/invoice) and the accountant printed it too using red ink.
4. The musician purchased the (car/urn) and the actress purchased it too paying cash.
5. The actor danced the (twist/waltz) and the choreographer danced it too smiling widely.
6. The exterminator swatted the (air/bee) and the workman swatted it too several times.
7. The spy learned the (language/alphabet) and the diplomat learned it too while in school.
8. The repairman reported the (damage/mildew) and the janitor reported it too around noon.
9. The merchant swept the (floor/patio) and the barber swept it too using a wide broom.
10. The customer read the (study/label) and the pharmacist read it too noting the warning.
11. The architect drew the (building/bungalow) and the illustrator drew it too using pencil.
12. The developer checked the (inventory/blueprint) and the contractor checked it too every day.
13. The priest organized the (dinner/raffle) and the widow organized it too selling punch.
14. The host explained the (game/skit) and the announcer explained it too during taping.
15. The parent grabbed the (bag/toy) and the kidnapper grabbed it too yelling loudly.
16. The zookeeper cleaned the (window/aviary) and the janitor cleaned it too twice a week.
17. The novelist edited the (book/memo) and the publisher edited it too finding errors.
18. The grandma licked the (knife/spoon) and the child licked it too while baking cookies.
19. The cellist recorded the (piece/polka) and the guitarist recorded it too exactly once.
20. The doctor swallowed the (formula/vitamin) and the child swallowed it too without water.
21. The preacher played the (piano/banjo) and the organist played it too every Sunday.
22. The invader looted the (village/pyramid) and the barbarian looted it too during the war.
23. The toddler tossed the (hat/toy) and the mother tossed it too having lots of fun.
24. The rebel pursued the (boat/raft) and the colonel pursued it too following for days.
25. The consultant rated the (product/mascara) and the consumer rated it too several times.
26. The gardener sprayed the (water/shrub) and the girl sprayed it too using the hose.
27. The butcher weighed the (beef/chop) and the lady weighed it too checking twice.
28. The wife froze the (meat/loaf) and the daughter froze it too sealing the container.
29. The gangster waved the (saw/bat) and the thug waved it too while making threats.
30. The model buttoned the (jacket/blazer) and the designer buttoned it too helping her.
31. The producer accepted the (award/bonus) and the actor accepted it too thanking God.
32. The nurse tested the (machine/adapter) and the doctor tested it too before the surgery.
33. The mom decorated the (food/cake) and the sister decorated it too using sprinkles.
34. The park ranger climbed the (tree/silo) and the worker climbed it too without a rope.
35. The salesman folded the (shirt/tunic) and the manager folded it too checking the tag.
36. The dictator revised the (treaty/memoir) and the leader revised it too before agreeing.
37. The beggar shared the (blanket/chowder) and the vagrant shared it too every cold night.
38. The Spaniard fled the (country/dungeon) and the Russian fled it too bringing nothing.
39. The designer picked the (sign/vase) and the florist picked it too despite the color.
40. The doorman dropped the (box/mug) and the maid dropped it too after being startled.
41. The handyman locked the (door/cart) and the butler locked it too fearing thieves.
42. The cowboy inspected the (horse/steed) and the jockey inspected it too before bidding.
43. The carpenter mailed the (bill/deed) and the plumber mailed it too later that week.
44. The bartender served the (drink/vodka) and the waitress served it too every night.
45. The pianist memorized the (music/polka) and the dancer memorized it too using tricks.
46. The salesman studied the (report/coupon) and the analyst studied it too during lunch.
47. The painter bought the (brush/putty) and the sculptor bought it too later that day.
48. The agent sent the (memo/note) and the receptionist sent it too although much later.
49. The gardener admired the (grass/hedge) and the golfer admired it too before putting.
50. The businessman poured the (oil/tea) and the geisha poured it too during the meal.
51. The typist corrected the (error/digit) and the editor corrected it too using red pen.
52. The cop described the (event/brawl) and the victim described it too during the trial.
53. The farmer washed the (china/basil) and the cook washed it too removing all the dirt.
54. The advisor changed the (writing/wording) and the writer changed it too several times.
55. The columnist reviewed the (play/skit) and the critic reviewed it too creating buzz.
56. The cameraman rode the (horse/float) and the veteran rode it too enjoying the parade.

VP-Anaphoric Conditions (NVP/OVP)

57. The policeman found the theater and the fireman (did too/did it too) using a map.
58. The kid ate the cereal and the teenager (did too/did it too) because he was hungry.
59. The boss printed the invoice and the accountant (did too/did it too) using red ink.
60. The musician purchased the car and the actress (did too/did it too) paying cash.
61. The actor danced the waltz and the choreographer (did too/did it too) smiling widely.
62. The exterminator swatted the bee and the workman (did too/did it too) several times.
63. The spy learned the language and the diplomat (did too/did it too) while in school.
64. The repairman reported the damage and the janitor (did too/did it too) around noon.
65. The merchant swept the floor and the barber (did too/did it too) using a wide broom.
66. The customer read the label and the pharmacist (did too/did it too) noting the warning.
67. The architect drew the building and the illustrator (did too/did it too) using pencil.
68. The developer checked the inventory and the contractor (did too/did it too) every day.
69. The priest organized the raffle and the widow (did too/did it too) selling punch.
70. The host explained the game and the announcer (did too/did it too) during taping.
71. The parent grabbed the bag and the kidnapper (did too/did it too) yelling loudly.
72. The zookeeper cleaned the window and the janitor (did too/did it too) twice a week.
73. The novelist edited the book and the publisher (did too/did it too) finding errors.
74. The grandma licked the spoon and the child (did too/did it too) while baking cookies.
75. The cellist recorded the piece and the guitarist (did too/did it too) exactly once.
76. The doctor swallowed the vitamin and the child (did too/did it too) without water.
77. The preacher played the piano and the organist (did too/did it too) every Sunday.
78. The invader looted the village and the barbarian (did too/did it too) during the war.
79. The toddler tossed the toy and the mother (did too/did it too) having lots of fun.
80. The rebel pursued the boat and the colonel (did too/did it too) following for days.
81. The consultant rated the product and the consumer (did too/did it too) several times.
82. The gardener sprayed the water and the girl (did too/did it too) using the hose.
83. The butcher weighed the beef and the lady (did too/did it too) checking twice.
84. The wife froze the meat and the daughter (did too/did it too) sealing the container.
85. The gangster waved the bat and the thug (did too/did it too) while making threats.
86. The model buttoned the jacket and the designer (did too/did it too) helping her.
87. The producer accepted the award and the actor (did too/did it too) thanking God.
88. The nurse tested the machine and the doctor (did too/did it too) before the surgery.
89. The mom decorated the cake and the sister (did too/did it too) using sprinkles.
90. The park ranger climbed the tree and the worker (did too/did it too) without a rope.
91. The salesman folded the shirt and the manager (did too/did it too) checking the tag.
92. The dictator revised the treaty and the leader (did too/did it too) before agreeing.
93. The beggar shared the blanket and the vagrant (did too/did it too) every cold night.
94. The Spaniard fled the country and the Russian (did too/did it too) bringing nothing.
95. The designer picked the vase and the florist (did too/did it too) despite the color.
96. The doorman dropped the box and the maid (did too/did it too) after being startled.
97. The handyman locked the door and the butler (did too/did it too) fearing thieves.
98. The cowboy inspected the horse and the jockey (did too/did it too) before bidding.
99. The carpenter mailed the bill and the plumber (did too/did it too) later that week.
100. The bartender served the drink and the waitress (did too/did it too) every night.
101. The pianist memorized the music and the dancer (did too/did it too) using tricks.
102. The salesman studied the report and the analyst (did too/did it too) during lunch.
103. The painter bought the brush and the sculptor (did too/did it too) later that day.
104. The agent sent the memo and the receptionist (did too/did it too) although much later.
105. The gardener admired the grass and the golfer (did too/did it too) before putting.
106. The businessman poured the tea and the geisha (did too/did it too) during the meal.
107. The typist corrected the error and the editor (did too/did it too) using red pen.
108. The cop described the event and the victim (did too/did it too) during the trial.
109. The farmer washed the basil and the cook (did too/did it too) removing all the dirt.
110. The advisor changed the wording and the writer (did too/did it too) several times.
111. The columnist reviewed the play and the critic (did too/did it too) creating buzz.
112. The cameraman rode the horse and the veteran (did too/did it too) enjoying the parade.
Acknowledgements

During the completion of this study, the first author was supported by an NSF Graduate Research Fellowship as well as a traineeship awarded by NIH to the Center for Research in Language at the University of California, San Diego. This research was also funded by NIH-DC000494. The authors wish to thank Jeffrey Witzel for his assistance with data collection and data processing in the eye-tracking study.
Table 1. Sample Sentences

| NP-Anaphoric Conditions |  |
|-------------------------|  |
| HFNP        | The typist [corrected the error] and the editor [corrected it too] [using red pen.] |
| LFNPs       | The typist [corrected the digit] and the editor [corrected it too] [using red pen.] |

| VP-Anaphoric Conditions |  |
|-------------------------|  |
| NVP         | The typist [corrected the error] and the editor [did too] [using red pen.] |
| OVP         | The typist [corrected the error] and the editor [did it too] [using red pen.] |

Note. Brackets indicate critical regions for analysis. HFNP = High-Frequency NP; LFNPs = Low-Frequency NP; NVP = Null VP-Anaphoric Phrase; OVP = Overt VP-Anaphoric Phrase.
<table>
<thead>
<tr>
<th>Region</th>
<th>VP1</th>
<th>Anaphor</th>
<th>Post-Anaphor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First-pass time</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HFNP</td>
<td>589 (25)</td>
<td>518 (21)</td>
<td>645 (34)</td>
</tr>
<tr>
<td>LFNP</td>
<td>669 (34)</td>
<td>498 (21)</td>
<td>599 (31)</td>
</tr>
<tr>
<td>Difference</td>
<td>-80*</td>
<td>20</td>
<td>46*</td>
</tr>
<tr>
<td><strong>Total time</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HFNP</td>
<td>757 (41)</td>
<td>653 (28)</td>
<td>722 (38)</td>
</tr>
<tr>
<td>LFNP</td>
<td>850 (51)</td>
<td>644 (31)</td>
<td>706 (39)</td>
</tr>
<tr>
<td>Difference</td>
<td>-93*</td>
<td>9</td>
<td>16</td>
</tr>
</tbody>
</table>

Note. Standard errors appear in parentheses. HFNP = High-Frequency NP; LFNP = Low-Frequency NP. Differences (i.e. HFNP minus LFNP) marked with a star (*) were significant at $p < .05$ while those marked with a pound sign (#) were significant at $p < .07$. 
Table 3. Means (ms) and Standard Errors for the VP-Anaphoric Conditions

<table>
<thead>
<tr>
<th>Region</th>
<th>VP1</th>
<th>Anaphor</th>
<th>Post-Anaphor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First-pass times</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OVP</td>
<td>607 (28)</td>
<td>471 (24)</td>
<td>662 (39)</td>
</tr>
<tr>
<td>NVP</td>
<td>595 (27)</td>
<td>321 (18)</td>
<td>654 (31)</td>
</tr>
<tr>
<td>Difference</td>
<td>12</td>
<td>150*</td>
<td>8</td>
</tr>
<tr>
<td>Total times</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OVP</td>
<td>752 (38)</td>
<td>627 (29)</td>
<td>751 (42)</td>
</tr>
<tr>
<td>NVP</td>
<td>769 (41)</td>
<td>451 (26)</td>
<td>797 (38)</td>
</tr>
<tr>
<td>Difference</td>
<td>-17</td>
<td>176*</td>
<td>-46#</td>
</tr>
</tbody>
</table>

Note. Standard errors appear in parentheses. OVP = Overt VP-Anaphoric Phrase; NVP = Null VP-Anaphoric Phrase. Differences (i.e. OVP minus NVP) marked with a star (*) were significant by subjects at \( p < .05 \) while those marked with a pound sign (#) were significant at \( p < .06 \).
Figure 1. First-pass (top) and total (bottom) reading times for the NP-anaphoric conditions. Bars indicate standard errors. HFNP = High-Frequency NP; LFNP = Low-Frequency NP. Comparisons marked with a star (*) were significant at \( p < .05 \) while those marked with a pound sign (#) were significant at \( p < .07 \).
Figure 2. First-pass (top) and total (bottom) reading times for the VP-anaphoric conditions. Bars indicate standard errors. OVP = Overt VP-Anaphoric Phrase; NVP = Null VP-Anaphoric Phrase. Comparisons marked with a star (*) were significant at $p < .05$ while those marked with a pound sign (#) were significant at $p < .06$. 
Chapter 3

An ERP investigation of the effects of anaphoric form: Processing null and overt verb phrase anaphors
Abstract

The current study investigated the effects of anaphoric form on processing by comparing constructions involving null and overt verb phrase (VP) anaphors (i.e. *did too* vs. *did it too*). Differentiation theories suggest that qualitatively different mechanisms underlie the processing of these two constructions while unification theories suggest a single set of mechanisms that is modulated by variation in anaphoric form. In this study, relative to the null VP-anaphoric condition, the overt VP-anaphoric condition elicited an early negativity that peaked approximately 100ms post-onset of the pronoun, *it*. Since a naturalness rating study suggested that the overt VP anaphor is less natural, this early negativity is interpreted as reflecting the additional processing required for an unexpected continuation. Relative to the overt VP-anaphoric condition, the null VP-anaphoric condition elicited an anterior positivity between 300-500ms post-onset of the word following the null anaphor site (i.e. *too*). This effect is interpreted as reflecting a delay in the onset of the contingent negative variation (CNV) that is related to the delayed detection of the anaphoric relationship.

Overall, the results were more consistent with unification theories than differentiation theories since they suggested that processing variation due to anaphor form is reflected in the modulation of certain electrophysiological patterns rather than qualitatively different patterns. In addition, we also observed effects of a task-related plausibility manipulation. Implausible subjects elicited an N400/P600 complex that fits with the growing literature that suggests that a P600 can be elicited by semantic violations.
An ERP investigation of the effects of anaphoric form: Processing null and overt verb phrase anaphors

Language researchers have always paid special attention to the study of anaphoric processing: how, during real-time sentence and discourse processing, the language processing system establishes relationships between two linguistic phrases that refer to the same real-world entity or concept (i.e. the antecedent and the anaphor). Behavioral research has suggested that anaphoric interpretation involves several processes including: a) detecting the anaphor, b) selecting and retrieving the antecedent, and c) integrating the antecedent (see Nicol & Swinney, 2002; Trueswell & Tanenhaus, 2005 for reviews). Now, with the advent of modern neurophysiological techniques, the systematic investigation of the neurophysiological aspects of these processes has begun. Electrophysiological techniques are particularly useful because they record electrical activity in the brain (via electrodes placed on the scalp) with millisecond resolution. Averaging the tiny voltage fluctuations related to the presentation of a stimulus across many trials and participants yields an event-related potential (ERP) that reflects the brain’s response to the stimulus.

**ERP Studies of Anaphora**

Thus far, the majority of electrophysiological studies dealing with anaphoric constructions have involved either relationships involving an overt antecedent and an overt anaphor (e.g. relationships between a noun phrase and a pronoun) or
relationships involving a filler (i.e. an overt “moved” wh-phrase) and its gap (i.e. a phonetically null empty “placeholder” for the filler) (see Callahan, 2008, for a recent review). From a processing perspective, two interrelated factors distinguish these two types of relationships: a) whether the antecedent or anaphor provides cues to guide the linking process and b) whether the anaphor is overt or null. When the anaphor is phonetically overt, it carries featural and/or semantic cues to the antecedent and the processor uses these cues to initiate a search in previous context for an appropriate antecedent. In contrast, when the anaphor is phonetically null, its position must be hypothesized and the processor uses cues provided by the antecedent (i.e. features that mark displacement from the canonical position) to initiate a strategic search in upcoming context for the first possible anaphor site (i.e. the active filler strategy, cf. Frazier & Flores D’Arcais, 1989).

ERP studies have revealed certain differences in the processing of relationships involving overt and null anaphors (see Callahan, 2008, for details), but have generally suggested that the retrieval and integration of antecedent information at the anaphor site is similar. Retrieval is indexed electrophysiologically by a left anterior negativity between 300 and 500ms post-onset of the critical word (i.e. a phasic LAN) (e.g. Anderson & Holcomb, 2005; King & Kutas, 1995; Streb, Rösler, & Hennighausen, 1999) and integration is indexed by modulations of components related to semantic and syntactic processing difficulty, a centro-posterior negativity (i.e. an N400) (e.g. Mecklinger, Schriefers, Steinhauer, & Friederici, 1995; Streb, Hennighausen, & Rösler, 2004; Swaab, Camblin, & Gordon, 2004) and a late positivity (i.e. a P600)
Thus, processing at the anaphor site does not appear to be directly affected by whether the anaphor is null or overt. This is surprising because, as compared to an overt anaphor, a null anaphor is more difficult to detect and provides fewer cues to its antecedent; however, in the null anaphoric relationships studied to date (e.g. filler-gap relationships), the antecedent is marked in a salient way that signals the upcoming anaphor site. The question remains, then, whether anaphoric form (i.e. null vs. overt) affects processing when the anaphoric relationship cannot be detected until the anaphor site.

To answer this question, it is necessary to investigate null anaphoric relationships where the antecedent is not specially marked. To date, these relationships have not received much study using electrophysiological techniques. The only previous studies have involved control constructions and gapping constructions. In the so-called control construction in (1) below, the subject of the matrix clause, the man, is co-referential with the subject of the infinitival clause, a null anaphor\(^1\) (indicated by PRO). In the gapping construction in (2) below, the verb in the first clause, bought, is co-indexed with a null anaphor in the second clause (indicated by e, as is the convention).

\(^1\) Transformational theories posit a null anaphor in this position while non-transformational theories do not; however, there is little to no processing evidence clearly distinguishing the two positions (cf. Fodor, 1989).
From a processing perspective, these null anaphoric relationships are both similar to and different from the relationships discussed previously. As in other null anaphoric relationships, the language processor must infer the presence of the anaphor; however, as in overt anaphoric relationships, the antecedent is not marked in any way. Thus, the language processor only initiates a search in previous context for the antecedent upon encountering a phrase that violates the phrase structure requirements, implying that a previous structural position must be filled by a null anaphor. For example, in (2) above, the verb, *bought*, is in no way marked as an antecedent and the phrase, *bread*, is the first indication that the verb in the second clause is represented by a null anaphor. Moreover, as in other null anaphoric relationships (but in contrast to overt anaphoric relationships) the anaphor carries no featural or semantic information to guide the search for the antecedent.

In control constructions, a central or posterior positivity between 300 and 600ms post-onset of the word following the null anaphor has been reported (Featherston, Gross, Münte, & Clahsen, 2000; Walenski, 2002). Moreover, this effect is related to how frequently a specific verb appears in a null anaphoric construction (Walenski, 2002). In contrast, encountering the word following the anaphor in gapping constructions is associated with an early fronto-central or centro-posterior negativity between approximately 100 and 300ms post-onset of the word following the anaphor.
site (i.e. an ELAN) (Kaan, Wijnen, & Swaab, 2004; Streb et al., 2004). Since the ELAN is typically associated with phrase structure violations (cf. Friederici, 2002), this effect has been interpreted as reflecting the fact that the phrase following the anaphor site is unexpected and is initially taken as a phrase structure violation. Thus, although the particular electrophysiological index of inferring the anaphor differs between gapping constructions and control constructions, both effects appear to be related to the expectation (or lack thereof) for the null anaphoric construction.

Once the null anaphor is detected, the antecedent information must be retrieved and integrated into the sentence at the anaphor site. As in the other anaphoric relationships studied to date, this retrieval process is sometimes indexed by a left anterior negativity between 200 and 400ms post-onset of the first word after the null anaphor position in control constructions (i.e. a phasic LAN) (Featherston et al., 2000; but cf. Walenski, 2002); in contrast, in the only study of gapping constructions, this process was indexed by an anterior positivity between 300 and 500ms post-onset (Kaan et al., 2004). Finally, similar to other anaphoric relationships, semantic and syntactic integration difficulty at the anaphor site are reflected by modulations of a centro-parietal negativity between 300 and 500ms (i.e. an N400) and a left posterior positivity between 600 and 900ms post-onset (i.e. a P600) post-onset of the next word (Kaan et al., 2004; but cf. Demestre, Meltzer, García-Albea, & Vigil, 1999).

The results of previous studies of null anaphoric constructions in which the antecedent is not specially marked suggest certain tentative conclusions. First, indices that reflect detection of the anaphor (i.e. the late positivity in control constructions and
the early negativity in gapping constructions) appear to be related to expectation for the null anaphor. Second, retrieval of the antecedent appears to be reflected by different electrophysiological indices in control and gapping constructions (i.e. a left anterior negativity and an anterior positivity respectively). Finally, as in other types of anaphoric relationships, integration of the antecedent information at the anaphor site is reflected by modulations of a centro-parietal negativity between 300 and 500ms (i.e. an N400) and a posterior positivity between 600 and 900ms (i.e. a P600) post-onset of the word following the anaphor site, reflecting semantic and syntactic processing difficulty.

The fact that detection and retrieval are reflected by different electrophysiological indices in control and gapping constructions suggests there may be an underlying distinction between the processing of these two types of null anaphoric relationships. In particular, in control constructions, the lexical entry of the verb contains the information that it can be associated with an infinitival complement involving the null anaphor, PRO. Because the matrix verb precedes the null anaphor position (at least in the languages tested so far: English German, and Spanish), this verb may signal the likelihood of an upcoming null anaphor, resulting in processing correlates that resemble those for relationships where the antecedent is marked (i.e. a phasic LAN and a P600). In contrast, in gapping constructions, it is nearly impossible to predict the null anaphor, resulting in a different set of processing correlates (i.e. an early negativity and an anterior positivity).
Another possibility is that the observed processing variation results from variation in the phrasal category of the antecedent and anaphor: control constructions involve nominal antecedents and anaphors while gapping constructions involve verbal antecedents and anaphors. Although the evidence is mixed, neurophysiological research has suggested that verbs and nouns may be stored or processed differently (Damasio & Tranel, 1993; Perani et al., 1999; but cf. Tyler, Russell, Fadili, & Moss, 2001) and previous electrophysiological studies have linked the processing of verbs with anterior positivity (e.g. Federmeier, Segal, Lombrozo, & Kutas, 2000). Further research on different types of constructions involving null anaphors is needed to distinguish between these hypotheses and solidify and refine the tentative conclusions outlined above.

Constructions That Allow Both Null and Overt Anaphors

The current study contributes to the growing literature on the electrophysiological correlates of processing anaphoric relationships by directly investigating the effects of the anaphoric form (i.e. null vs. overt) on processing. As described previously, it is difficult to draw strong conclusions about the effects of anaphoric form from a comparison of the null and overt anaphoric relationships studied to date. Moreover, previous studies have suggested that different null anaphoric constructions may be associated with different electrophysiological patterns. To avoid these possible confounds, it is necessary to compare constructions that differ only in terms of anaphoric form. Fortuitously, there do exist several types of
constructions where both null and overt anaphors are grammatical. In the current study, we compare VP-anaphoric constructions involving a null VP anaphor with those involving an overt VP anaphor. In the null VP-anaphoric construction in (3), a null anaphor (indicated by e) refers to the antecedent VP, *edited the book*, while in the overt VP-anaphoric construction in (4), an overt anaphor, *it*, refers to the same antecedent.  

(3) The novelist *edited the book* one last time and the publisher *did e too.*  
(4) The novelist *edited the book* one last time and the publisher *did it too.*

Although the null and overt forms of the VP anaphor occupy the same syntactic position and are both grammatical in the constructions tested herein, the two types are not used completely interchangeably in natural language. In particular, it has been claimed that the null anaphor requires a syntactically-parallel antecedent in linguistic context while the overt anaphor may take antecedents of various forms and types. Nonetheless, critically, the two constructions are identical in terms of the syntactic and semantic information provided to the processor before the anaphor site. Thus, in both, the anaphoric relationship can only be detected once the anaphor site is encountered. Moreover, the word following the anaphor site, *too*, is identical in both constructions,

\[^{2}\text{The null VP anaphor must have a linguistic VP antecedent, but the overt VP anaphor technically can refer to any event. For the current study, the critical point is that both forms take the event specified in the VP as the antecedent. Therefore, for ease of exposition, we class both as VP anaphors.}\]
allowing a direct electrophysiological comparison of the processing elicited by the two anaphoric forms.

**Theoretical Approaches**

Several broad-scope linguistic and psycholinguistic theories of anaphoric processing and interpretation offer accounts for the differential usage of the two VP-anaphoric forms. These theories basically fall into two classes: differentiation theories and unification theories. Differentiation theories propose that the processing and interpretation of one class of anaphors (e.g. the null VP anaphor) is qualitatively different from that of another class of anaphors (e.g. the overt VP anaphor). In particular, the null VP anaphor is claimed to be linked to the surface syntactic form of the sentence and interpreted via a syntactic copying procedure while the overt VP anaphor is linked to the discourse model and interpreted via active reasoning (Frazier & Clifton, 2001, 2005; Sag & Hankamer, 1984). This is an appealing idea, but when the two classes of anaphors have been directly compared, the results have not strongly supported such a distinction (see Bélanger, 2004 for a critical review).

Generally,

---

3 Most theoretical linguistic work on VP-anaphoric constructions has focused exclusively on VP-elliptical constructions (e.g. the null VP anaphor) (see Johnson, 2001; Kennedy, 2003, Merchant, 2001 for overviews); however, no one theory of ellipsis has been completely successful. Since these theories fail to make specific processing predictions regarding the comparison of null and overt VP anaphors, we do not discuss them in detail here.

4 Frazier & Clifton (2001) also cited a lack of complexity effects in null VP-anaphoric constructions as evidence for a syntactic copying procedure; however, complexity effects have never been investigated in overt VP-anaphoric constructions. Thus, it is unclear whether the lack of complexity effects is related to the class of anaphor or to
similar effects have been observed for both classes of anaphors (Martin & McElree, 2008; Murphy, 1985a, 1990). One of the few observations of differential effects (as in Tanenhaus & Carlson, 1990) was later attributed to factors other than the class of the anaphor (i.e., the antecedent form and the specific task given participants, cf. Mauner, Tanenhaus, & Carlson, 1995; Murphy, 1990). Moreover, the only other studies that observed differential effects (Hestvik, Nordby, & Karlsen, 2005; Streb, Hennighausen, & Rösler, 2004) compared an NP anaphor of one class to a VP anaphor of the other class, leaving open the possibility that the observed differences were due to the phrasal category rather than the class of the anaphor (see Callahan, 2008, for discussion). In particular, although the only previous ERP study addressing this question (Streb et al., 2004) claimed to find qualitatively different processing, this study compared nominal anaphors (e.g. pronouns and proper names) with verbal anaphors (e.g. elided verbs in gapping constructions), so the results may have been related to the phrasal category rather than the class of the anaphor.

In contrast to differentiation theories, unification theories propose that the same processing mechanism underlies the processing of all anaphoric forms (e.g. Ariel, 1990, 1991; Givón, 1983; Murphy, 1985b). To account for differential usage, these theories introduce a pragmatic principle that states that, under optimal conditions, speakers produce anaphors that help the hearer identify the appropriate antecedent (see Grice, 1975, for details). Specifically, unification theories claim that two factors interact to determine the anaphoric form used: a) characteristics of the some other characteristic of these constructions, perhaps parallelism (cf. Callahan, Shapiro, & Love, submitted).
anaphor that determine how effective it is in uniquely identifying its antecedent and b) characteristics of the antecedent that determine its accessibility to the hearer.

Anaphors vary in terms of their phonological size, the amount of information they carry, and whether they are ambiguous in the relevant context. Antecedents vary in terms of their distance from the anaphor, their prominence in the relevant context, and the number of competitors they have. Fundamentally, speakers tend to pair less informative anaphors (e.g. the null VP-anaphor) with more accessible antecedents, and more informative anaphors (e.g. the overt VP-anaphor) with less accessible antecedents. Related unification processing models suggest this correspondence is also relevant to the processing and interpretation of anaphors (e.g. Almor, 1998; Gordon & Hendrick, 1998; Grosz, Joshi, & Weinstein, 1995; Murphy, 1985b; Walker, Joshi, & Prince, 1998).

Unification theories have found general support in investigations of the processing and interpretation of different overt anaphoric forms in English (e.g. Almor, 1998; Gordon, Grosz, & Gilliam, 1993) as well as the processing and interpretation of null (as compared to overt) subject pronouns in pro-drop languages (e.g. Spanish and Italian, cf. Alonso-Ovalle, Fernández-Solera, Frazier, & Clifton, 2002, and especially Carminati, 2002). In these studies, less informative forms tend to be interpreted as referring to highly accessible antecedents and more informative forms as referring to less accessible antecedents. Moreover, processing difficulty is observed when an over-informative form refers to a highly accessible antecedent (Callahan, Nicol, & Love, submitted; Carminati, 2002).
The Current Study

The current study investigates the hypothesis that variation in anaphoric form (i.e. null vs. overt) results in variation in processing, and that this variation will be reflected in the character of the ERPs. Differentiation theories suggest that the processing mechanisms underlying the two constructions are qualitatively different and hence might be reflected by different electrophysiological waveforms (cf. Streb et al., 2004). In contrast, unification theories suggest that the processing mechanisms underlying the two constructions are the same and hence rather than eliciting different electrophysiological waveforms, processing variation related to the type of the anaphor should be reflected in a modulation of the electrophysiological pattern.

In terms of the specific pattern of electrophysiological correlates elicited by the manipulation of anaphoric form, we predicted an early centro-posterior negativity and a subsequent anterior positivity appearing at the onset of the word following the null anaphor site. This prediction is based on previous ERP studies that compared null and overt verbal anaphoric constructions (Kaan et al., 2004; Streb et al., 2004). In the current study, the plausibility of the subject of the second conjunct was also manipulated so that participants could perform a sensicality judgment task. Implausible versions of each experimental item were created by replacing an animate NP with an inanimate NP matched in terms of length and frequency (see the Methods section for details). This manipulation also served as a replication of previous research on semantic incongruity; therefore, since the implausible NP fits poorly with the prior
semantic context, we predicted that a comparison of the plausible and implausible NPs
would reveal an N400 (Kutas & Hillyard, 1980).

Methods

Participants

Seventy-seven undergraduate students from the University of California, San
Diego participated for course credit. All participants were right-handed as assessed by
the Edinburgh inventory (Oldfield, 1971), neurologically-unimpaired with normal
hearing and normal or corrected-to-normal vision, and native English speakers with no
exposure to a second language before age six. Participants’ data was excluded from
further analysis if they did not meet a pre-set criterion of 85% accuracy on the
behavioral task (n = 3), did not respond to more than 15% of items (n = 2), or if the
EEG was contaminated by artifacts for more than 33% of trials for any condition (in
any time window of interest, since with auditory presentation critical time windows
often overlap) (n = 11). Thus, 61 participants’ data were included for further analysis
(25 male, 36 female; Age: Mean = 19.95; SD = 1.51; Edinburgh handedness score:
Mean = 80.5; SD = 16.5).

Materials

This paper presents a study comparing sentences involving null and overt VP
anaphors; however, during the experiment, participants also heard sentences for
another study that will not be discussed here. Thus, over the course of the experiment,
participants were presented with 120 experimental items for the current study, 60 experimental items for another study, and 180 filler items. The experimental and filler items were assembled into a single pseudo-randomized list with the constraint that no condition appeared more than twice in a row. Six counterbalanced versions of the list were created by systematically varying anaphor form and plausibility across the lists for the experimental items. Although each item only appeared once in a given list, across lists each item appeared in each condition. Participants responded to items in all conditions, but only heard any given item once.

-------------------------------------

Insert Table 1 about here

-------------------------------------

The study discussed in this paper involved four conditions: P-NVP (Plausible-Null VP), P-OVP (Plausible-Overt VP), I-NVP (Implausible-Null VP), I-OVP (Implausible-Overt VP) (see Table 1 for sample sentences). Each experimental item consisted of two sentences conjoined by the conjunction, and. The first conjunct was always the same while the second conjunct varied in terms of the form of the anaphor (Null VP or Overt VP) (see (5) below).

(5) The novelist edited the book one last time and the publisher (did/did it) too.

(6) The novelist edited the book one last time and the pineapple (did/did it) too.
For the sensicality judgment task, an implausible version of each item was created by replacing the animate subject of the second conjunct (e.g. *publisher*) with an inanimate noun that was matched on length and frequency, but was implausible given the context (e.g. *pineapple*) (see (6) above). To minimize the possibility that participants would develop strategies during testing, filler items were also constructed. So that participants would not become aware of the critical conditions, half of the filler items consisted of conjoined sentences involving VP anaphoric phrases other than those involved in the critical experimental items (e.g. *has too*). The other half consisted of sentences of varying structures that did not involve VP anaphoric phrases. For purposes of the sensicality judgment task, half of the filler items were rendered implausible by replacing a noun at various points throughout the sentence with a noun that was implausible given the context. In the filler items, the placement of the implausible noun varied so that participants would not develop expectations about the position of the critical word for the task. For the sensicality judgment task, there were an equal number of yes/no responses across the entire experiment.

*Naturalness Rating Pretest*

Previous studies of null anaphoric constructions have indicated that expectancy plays a role in processing (Kaan et al., 2004; Streb et al., 2004; Walenski, 2002); thus, we wanted to assess whether participants expect a particular anaphoric form in the sentential contexts under investigation. Given the difficulty of obtaining accurate
estimates of how often each anaphoric form is used in these contexts, we instead collected naturalness ratings for the plausible versions of the 180 experimental items. Thirty participants rated the naturalness of each item on a Likert scale ranging from 1 (very natural) to 5 (very awkward). The results indicated mean ratings for the two VP anaphoric conditions that fell between 3 and 4 where a rating of 3 represented “OK” and a rating of 4 represented “Somewhat awkward”. Although participants’ ratings were similar for both VP-anaphoric forms (Null VP: Mean = 3.2, SD = .76; Overt VP: Mean = 3.7, SD = .62), Overt VP anaphors were rated as significantly less natural, $t_1(29) = 4.04, p = .0004, t_2(179) = 18.36, p < .0001$. This suggests that, overall, participants have an expectation for the null VP-anaphor in these contexts.

**Procedure**

This study was conducted in keeping with proper ethical guidelines and with the approval of the Institutional Review Board at the University of California, San Diego. Before beginning the experiment, each participant gave informed consent and completed a screening questionnaire about their language use and language experience as well as the Edinburgh handedness inventory (Oldfield, 1971).

Participants were seated in a comfortable chair in a sound-attenuated cubicle in front of a two-button response box. Following electrode application, participants were given instructions for the experiment. They were told that they would be hearing sentences and that their task was to decide whether or not each sentence made sense. They were also instructed to refrain from blinking or moving during each sentence and
to keep their fingers resting lightly on the buttons throughout the experiment. At the start of each trial, an aural cue (i.e. “Ready?”) alerted participants to fixate a red dot on the screen. After 250ms of silence, the aural presentation of the sentence began. The end of the sentence was followed by a variable period of silence (750, 1000, or 1250ms) to prevent participants from developing expectations about when the critical response would be required. Following this period of silence, a square-wave onset tone cued participants to make their response via the response box. Participants had a total of 3s to respond. Then, after a 500ms pause, the next trial began. At the beginning of the experiment, participants completed a ten-sentence practice session to familiarize them with the environment and procedure and received feedback on their performance. Then, they were presented with one of the six pseudorandomized lists.

The entire experiment lasted approximately two hours, with electrode application requiring approximately one hour and the experimental task requiring approximately one hour. Participants were given short breaks at various points throughout the task portion.

Data Acquisition and Analysis

The electroencephalogram (EEG) was recorded using a QuikCap (Compumedics, Inc.) embedded with 32 sintered Ag/AgCl electrodes arranged in a modified international 10-20 system including midline (Fz, FCz, Cz, CPz, Pz, Oz) and lateral sites (FP1/2, F3/4, F7/8, C3/4, T7/8, P3/4, P7/8, O1/2) as well as FC3/4, FT7/8, CP3/4, TP7/8. Bipolar loose electrodes were also applied to the outer canthus of both
eyes and above and below the left eye to monitor eye-movements. Electrode impedances were kept below 8kΩ. The EEG signal was amplified 100x with a SynAmps amplifier (3dB cut-off), band-pass filtered between 0.15 and 30Hz, and digitized at 500Hz. All electrodes were referenced on-line to the linked mastoids.

Time windows of 1000ms were examined that began at the onset of the word of interest with the 200ms prior to the onset of the word of interest used as a baseline. Time windows containing artifacts from amplifier blocking, signal drift, excessive eye-movements, muscle activity, or eye-blinks were rejected off-line, resulting in the rejection of 12% of all possible time windows (due to auditory presentation, critical time windows often overlapped). When necessary, the data for electrodes contaminated for the majority of the experiment was interpolated from surrounding electrodes. In the creation of the average ERPs, all artifact-free trials were included. Data from the extreme frontal (FP1, FP2) and extreme occipital electrodes (O1, O2, Oz) was excluded from statistical analysis since these electrodes were not involved in any hypotheses of interest and were often contaminated by eye-movements and muscle activity. Repeated-measures ANOVAs including experimental and electrode factors were thus conducted for the subset of electrodes of interest (Fz, F3/4, F7/8, FCz, FC3/4, Cz, C3/4, CPz, CP3/4, Pz, P3/4, P7/8, FT7/8, T7/8, TP7/8). P-values for comparisons with more than one degree of freedom in the numerator were adjusted using the correction proposed by Huynh and Feldt (1970). In these cases, original degrees of freedom and the corrected p-value are reported.
Experimental Design

For purposes of analysis, the critical conditions in the current study were considered as part of a 2 X 2 within-subjects repeated-measures design involving the factors Anaphor Form (Null VP or Overt VP) and Plausibility (Plausible or Implausible).

Results

Accuracy

Participants were highly accurate on the sensicality judgment task in the overall study (Mean = 95.13%, SD = 2.84%), indicating that they were paying attention and performing the task to the best of their ability. Participants were also highly accurate when only the study discussed in this paper (i.e. the conditions: P-NVP, P-OVP, I-NVP, I-OVP) were considered (Mean = 94.39%; SD = 8.29%) (see Table 2 for condition means and standard deviations). A repeated-measures ANOVA involving the factors, Plausibility (Plausible or Implausible) and Anaphor Form (Null VP or Overt VP) revealed that the only significant effect was a main effect of Plausibility ($F_1(1, 60) = 36.23, p = .000, F_2(1, 179) = 105.88, p = .000$), reflecting significantly higher accuracy for Implausible (Mean = 98.25%) than Plausible items (Mean = 90.52%). This result is not surprising since the Implausible items contain a salient violation. Crucially, accuracy did not differ between the anaphoric forms ($F_1(1, 60) = 5.51, p = .55, F_2(1, 179) = 34.71, p = .42$).
ERP Analyses of the Effects of Plausibility

Figure 1 shows grand average ERPs for plausible and implausible second subjects. Visual inspection of the waveforms revealed a centro-medial negative peak for the Implausible condition approximately 400-700ms post-onset of the second subject (e.g. publisher/pineapple) and a posterior medial positivity for the Implausible condition approximately 700-1000ms post-onset of the second subject. Repeated-measures ANOVAs including experimental and electrode factors were thus conducted for these two time windows of interest following the onset of the second subject (400-700ms and 700-1000ms). Separate analyses were performed on data from midline electrodes including the factors, Plausibility (Plausible or Implausible) and Anteriority (Frontal, Fronto-Central, Central, Centro-Posterior, or Posterior). Analyses performed on data from lateral electrodes included the factors, Plausibility and Anteriority as well as the additional factors, Laterality (Medial or Lateral) and Hemisphere (Right or Left).5

5 Analyses were initially performed including the factor, Anaphor Form (Null VP or Overt VP). Since all effects related to Anaphor Form failed to reach significance (all p > .05), the data reported here are collapsed across Anaphor Form.
400-700ms. Statistical comparisons performed in a 400-700ms window following the onset of the second subject revealed a significant main effect of Plausibility at midline ($F(1, 60) = 39.37, p = .000$) and lateral electrodes ($F(1, 60) = 36.2, p = .000$) and an interaction between Plausibility and Laterality at lateral electrodes ($F(1, 60) = 12.92, p = .001$). There was also a significant main effect of Anteriority at both midline ($F(4, 240) = 62.13, p = .000$) and lateral electrodes ($F(4, 240) = 87.96, p = .000$). No other effects reached significance. In keeping with the descriptive observations above, the statistical results in the 400-700ms time window reflect the fact that the grand average ERP was significantly more negative for Implausible items and that the observed difference had a centro-medial maximum. The timing and distribution suggest that this effect is an instance of the N400 (Kutas & Hillyard, 1980).

700-1000ms. Statistical comparisons performed in the 700-1000ms window following the onset of the second subject revealed a significant interaction between Plausibility and Anteriority at midline ($F(4, 240) = 46.71, p = .000$) and lateral electrodes ($F(4, 240) = 40.88, p = .000$). For lateral electrodes, there was also a significant interaction between Plausibility and Laterality ($F(1, 60) = 5.22, p = .03$), a
significant interaction between Anteriority and Laterality ($F(4, 240) = 4.01, p = .004$), and significant three-way interactions between Plausibility, Laterality, and Anteriority ($F(4, 240) = 8.73, p = .000$) and Plausibility, Laterality, and Hemisphere ($F(1, 60) = 4.38, p = .04$). There was also a significant main effect of Anteriority at both midline ($F(4, 240) = 105.78, p = .000$) and lateral ($F(4, 240) = 121.42, p = .000$) electrodes and Laterality at lateral electrodes ($F(1, 60) = 86.06, p = .000$). No other effects reached significance. In keeping with the descriptive observations above, the statistical results reflect the fact that the grand average ERP in this time window was significantly more positive for Implausible items and that the effect had a posterior medial maximum. The timing and distribution suggest that this effect is an instance of the P600 (Hagoort, Brown, & Groothusen, 1993; Osterhout & Holcomb, 1992).

**ERP Analyses of Effects of Anaphor Form**

Figure 2 shows grand average ERPs for items involving null and overt VP anaphors from the onset of “too”. Visual inspection of the waveforms reveals a posterior medial negative peak for the Overt VP condition approximately 50-150ms post-onset of “too” and an anterior lateral positivity for the Null VP condition approximately 300-500ms post-onset. Repeated-measures ANOVAs including experimental and electrode factors were thus conducted for time windows of interest following the onset of “too” (50-150ms and 300-500ms). All ERP analyses of the effects of Anaphor Form were performed on Plausible items only. Separate analyses were performed on data from midline electrodes including the factors, Anaphor Form
(Null VP or Overt VP) and Anteriority (Frontal, Fronto-Central, Central, Centro-Posterior, or Posterior). Analyses performed on data from lateral electrodes included the factors, Anaphor Form and Anteriority as well as the additional factors, Laterality (Medial or Lateral) and Hemisphere (Right or Left).

-------------------------------------
Insert Figure 2 about here
-------------------------------------

50-150ms. Statistical comparisons in the 50-150ms window following the onset of “too” revealed a significant main effect of Anaphor Form at midline ($F(1, 60) = 5.15, p = .03$) and lateral ($F(1, 60) = 4.42, p = .04$) electrodes. There was also a significant interaction between Anaphor Form and Anteriority at midline ($F(4, 240) = 7.26, p = .002$) and lateral ($F(4, 240) = 4.95, p = .01$) electrodes and a significant interaction between Anaphor Form and Laterality at lateral electrodes ($F(1, 60) = 6.44, p = .01$). At lateral electrodes, there was also a significant main effect of Hemisphere ($F(1, 60) = 15.61, p = .000$) and Laterality ($F(1, 60) = 7.30, p = .009$). No other effects reached significance. In keeping with the descriptive observations above, the statistical results reflect the fact that the grand average ERP was more negative in this time window for the Overt VP condition and that the effect had a posterior medial maximum. The interpretation of this effect will be discussed in detail in the discussion section.
300-500ms. Statistical comparisons in a 300-500ms window following the onset of “too” revealed a significant interaction between Anaphor Form and Anteriority at midline ($F(4, 240) = 2.90, p = .05$) and lateral ($F(4, 240) = 4.08, p = .02$) electrodes. At lateral electrodes, there was also a significant main effect of Laterality ($F(1, 60) = 13.19, p = .001$) as well as a significant interaction between Anteriority and Hemisphere ($F(4, 240) = 4.08, p = .007$) and Anteriority and Laterality ($F(4, 240) = 3.76, p = .01$). In keeping with the descriptive observations above, the statistical results reflect the fact that the grand average ERP was more positive for the Null VP condition in this time window and that the effect has an anterior lateral distribution. The interpretation of this effect will be discussed in more detail in the discussion section.

Discussion

The current study used electrophysiological techniques to investigate the effect of anaphoric form on the processing of anaphoric constructions. To this end, VP-anaphoric constructions were compared that consisted of two conjoined phrases and differed only in terms of the form of the anaphor in the second conjunct (i.e. a null VP anaphor, did too, vs. an overt VP anaphor, did it too). Participants performed a sensicality judgment task and as a task manipulation, the plausibility of the subject of the second conjunct was also varied.
Effects of Plausibility

The grand average ERPs were examined for effects of plausibility from the onset of the subject of the second conjunct. Implausible subjects elicited a negativity between 400 and 700ms post-onset with a centro-medial maximum as well as a positivity between 700 and 1000ms post-onset with a posterior medial maximum. The timing and distribution of these effects suggests that they represent an N400/P600 complex. The observation of an N400 to the implausible subject is not surprising since the N400 is typically associated with semantic incongruity (Brown, Hagoort, & Kutas, 2000; Kutas & Hillyard, 1980) and the implausible subject fits poorly with the previous semantic context. Since all the implausible subjects were also inanimate, the N400 may also reflect a violation of the expectation for animate subjects (cf. Weckerly & Kutas, 1999).

At first glance, the observation of a P600 in the current study is more surprising since this effect was originally described as being elicited by syntactic violations and the constructions tested here are completely syntactically legal; recently, however, several studies have reported P600 effects in response to semantic violations (e.g. Hoeks, Stowe, & Doedens, 2004; Kim & Osterhout, 2005; Kuperberg, Sitnikova, Caplan, & Holcomb, 2003; van Herten, Kolk, & Chwilla, 2005). In a recent review (Kuperberg, 2007), Kuperberg suggested that the P600 reflects conflict between a semantic memory-based processing stream sensitive to lexical associative and categorical relationships and a combinatorial processing stream sensitive to morphosyntactic and thematic-semantic constraints. Furthermore, Kuperberg...
delineated a number of linguistic and experimental conditions that increase the likelihood of observing a P600 in response to a semantic violation: a) violations of a verb’s animacy constraints, b) semantic associations between a verb and its arguments, c) an explicit acceptability judgment task, and d) a coherent surrounding sentence or discourse context. In the current study, three of these conditions are met. First, all of the critical verbs required animate subjects; therefore, the implausible second subject violated the animacy constraints of the verb. Second, participants performed a sensicality judgment task. Finally, the first conjunct (which was grammatical and sensical in all experimental items) provided a coherent surrounding context. The only condition identified by Kuperberg that did not occur in the current study was the presence of semantic associations between the verb and its arguments. Thus, the current study supports previous studies (Friederici & Frisch, 2000; Hoeks et al., 2004; Kuperberg et al., 2007) in suggesting that a P600 may be observed to a semantic violation even in the absence of semantic associations between the verb and its arguments. The current study also observed an N400/P600 complex even though the verb information in the second conjunct was only represented by a VP-anaphoric phrase. This finding fits with a previous study (Kaan et al., 2004) that reported an N400/P600 following an implausible object in gapping sentences (where the verb is represented by a null anaphor) and a previous study (Friederici & Frisch, 2000) that reported an N400/P600 complex in an implausible context where the object of the verb was an anaphor. This suggests that semantic and thematic information is available
immediately at an anaphor site and that retrieved antecedent information is sufficient to elicit a P600.

In addition, the observation of the N400/P600 complex at the implausible second subject rather than at the anaphoric verb phrase itself suggests that verb information is already available at the point following the conjunction. One possible explanation is that the parser initially pursues a conjoined noun phrase analysis. Then, the observed effects reflect the semantic and thematic violations that result if the second subject is taken as the object of the antecedent verb. A more intriguing explanation is that the expectation for parallelism elicited by the conjunction, and, results in the reactivation of information from the first conjunct. Several studies have suggested that, in parallel clauses, information from the first conjunct facilitates processing of the second conjunct (Frazier, Munn, & Clifton, 2000; Frazier, Taft, Roeper, Clifton, & Ehrlich, 1984; Knoeferle & Crocker, accepted) and a study of similar VP-anaphoric constructions (Callahan, Shapiro, & Love, submitted), found that information from the verb in the first conjunct was reactivated immediately following the conjunction.

The nature of the mechanism that underlies such parallelism effects is currently unclear. Frazier and Clifton (2001, 2005) suggested a copying mechanism (i.e. copy α) that recovers the elided material in elliptical constructions with unambiguous scope; however, evidence that verb information from the first conjunct is available following the conjunction in non-elliptical constructions (i.e. Kaan et al., 2004; Knoeferle & Crocker, accepted) suggests that the processing mechanism that underlies
these effects may have a broader scope. Kaan and colleagues proposed syntactic persistence as a possible mechanism since the first and second conjuncts of the gapping constructions tested in their study had identical syntactic structures (cf. Kaan et al., 2004); however, overt VP-anaphoric constructions such as those in the current study do not require syntactic parallelism between the antecedent and anaphor so semantic parallelism may be as useful as syntactic parallelism in the processing and interpretation of parallel constructions (cf. Kehler, 2002). Further research on the processing of parallelism is needed to specify the nature of the mechanism underlying these effects.

**Effects of Anaphor Form**

The grand average ERPs were examined for effects of anaphor form at the word following the anaphor site (i.e. *too*). Overall, the waveforms for the null and overt conditions were highly overlapping, suggesting it is unlikely that the two anaphoric forms invoke different underlying processing mechanisms, at least in contexts where both are grammatical; nonetheless, two effects related to anaphor form were observed. First, an early centro-posterior negativity between 50 and 150ms post-onset was observed in the overt VP-anaphoric condition. The early centro-posterior negativity peaked approximately 100ms after the onset of the pronoun, *it*. Since this effect only occurs in the overt condition, it is likely related to the processing of the overt pronoun and may represent an N100. Although the auditory N100 has been interpreted as reflecting the initial sensory processing of the auditory stimulus, it has
also been demonstrated to be sensitive to cognitive factors such as attention (Davis, 1964; Picton & Hillyard, 1974) and expectancy (Starr, Aguinaldo, Roe, & Michalewski, 1997).

Interestingly, this effect resembles the early centro-posterior negativity observed following the null anaphor in a previous ERP study of gapping constructions (Kaan et al., 2004). Kaan and colleagues interpreted this effect as an instance of the ELAN even though it did not have the typical left-anterior distribution and suggested that it reflected the fact that the word following the null anaphor site is initially taken as a phrase structure violation; however, if the effect in the current study has the same source (the timing and distribution are intriguingly similar, although the effect in the current study is somewhat more peaked) then it can be also be elicited when there is no chance of mistakenly detecting a phrase structure violation. Taking these studies together and placing them in the context of other studies of null anaphoric constructions, we suggest that these early negativities index the detection and initial processing of the anaphor and are related to expectation (or lack thereof) for the anaphor. In gapping constructions, the null anaphor is the unexpected continuation; however, in the current study the overt anaphor was the unexpected continuation (i.e. sentences containing overt VP anaphors were rated as less natural). Thus, rather than marking a phrase structure violation per se, the early negativity observed in the current study may more generally reflect the deployment of additional processing resources to an unexpected continuation.
In the current study, we also observed an anterior positivity between 300 and 500ms post-onset for the null anaphor condition (i.e. *did too*). This effect also resembles an effect observed in a previous ERP study of gapping (Kaan et al., 2004). Along with Kaan and colleagues, we suggest that this positivity might be conceptualized as a delay in the onset of the contingent negative variation (CNV)-like negativity that is elicited by closed-class words (i.e., the N280; cf. Brown, Hagoort, & ter Keurs, 1999; Neville, Mills, & Lawson, 1992) or the CNV that precedes participants making a response (Rockstroh, Elbert, Birbaumer, & Lutzenberger, 1982; Walter, Cooper, Aldridge, & McCallum, 1964). The fact that this effect is observed (or is larger) for the null condition suggests that processing at word following the anaphor site is increased relative to the overt condition, reflecting the delayed detection of the anaphor site. Essentially, then, the pattern in the current study reveals differences in when processing resources are deployed during the processing of constructions involving null and overt VP anaphors. In the overt condition, the overt anaphor, *it*, elicits additional processing as reflected by the early negativity. In contrast, in the null condition, the word following the anaphor site, *too*, elicits additional processing as reflected by the anterior positivity.

**Conclusion**

This study compared the processing of null and overt VP-anaphoric constructions. Relative to the null VP anaphor, the overt VP anaphor, *it*, elicited an early negativity that peaked approximately 100ms post-onset. Based on a study that
suggested that the overt VP anaphor is less natural, we suggest that this early negativity reflects the additional processing required for an unexpected continuation. In the context of previous studies of null anaphoric constructions, we view these early negativities as reflecting a lack of expectation for the anaphor. An anterior positivity between 300-500ms post-onset of the word following the null anaphor site was also observed. We interpret this effect as reflecting a delay in the onset of the CNV and suggest it reflects increased processing related to the delayed detection of the anaphoric relationship.

The observed pattern of effects does not support the predictions of differentiation theories because, overall, we found no evidence of qualitatively different processes for null and overt VP anaphors. Rather, we suggest, that consistent with unification theories, the two anaphoric forms elicit similar processes and the difficulty or costliness of these processes is modulated by factors related to the form of the anaphor. Further electrophysiological research comparing the processing of null and overt anaphoric constructions will help to solidify and refine the interpretations of the effects observed in this study. In particular, we suggest that further research in constructions or languages where null anaphors are very frequent and alternate with overt anaphors (e.g. Spanish, Italian) will prove especially fruitful.

This chapter, in full, has been submitted for publication as it may appear in *Brain and Language*, Callahan, S.M., Nicol, J., Love, T., & Swinney, D. The dissertation author was the primary investigator and author of this paper.
Acknowledgments

During the completion of this study, the first author was supported by an NSF Graduate Research Fellowship as well as a traineeship awarded by NIH to the Center for Research in Language at the University of California, San Diego. This research was also funded by NIH (R01-DC000494). Although David Swinney passed away before the preparation of this manuscript, he was integral to the design and data collection phases of the study and thus is included as a co-author.
References


Carminati, M.N. (2002). The processing of Italian subject pronouns. (Doctoral
dissertation, University of Massachusetts at Amherst, 2002). Dissertation Abstracts International, A: The Humanities and Social Sciences, 63, 1, 163-A.


Table 1. Sample Sentences

| Condition | Sample Sentence |
|-----------|----------------|---|
| P-NVP     | The novelist edited the book one last time and the publisher did too. |
| P-OVP     | The novelist edited the book one last time and the publisher did it too. |
| I-NVP     | The novelist edited the book one last time and the pineapple did too. |
| I-OVP     | The novelist edited the book one last time and the pineapple did it too. |

Table 2. Means and Standard Deviations in terms of Percent Correct on the Sensicality Judgment Task

<table>
<thead>
<tr>
<th>Condition</th>
<th>Percent Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-NVP</td>
<td>90.49 (10.2)</td>
</tr>
<tr>
<td>P-OVP</td>
<td>90.55 (9.91)</td>
</tr>
<tr>
<td>I-NVP</td>
<td>98.58 (2.15)</td>
</tr>
<tr>
<td>I-OVP</td>
<td>97.92 (3.29)</td>
</tr>
</tbody>
</table>

Figure 1. Grand average ERPs from the onset of the second subject for Plausible (dotted lines) and Implausible items (solid lines). Negative is plotted upwards. PLAUS = Plausible; IMPLAUS = Implausible.
Figure 2. Grand average ERPs from the onset of “too” for items involving Null VP anaphors (dotted lines) and items involving Overt VP anaphors (solid lines). Negative is plotted upwards. P-NVP = Plausible Null Verb Phrase; P-OVP = Plausible Overt Verb Phrase.
Chapter 4

The interpretation of null and overt VP-anaphoric constructions
Abstract

This study directly compares the interpretation of constructions involving a verb phrase (VP)-anaphoric phrase containing a null anaphor (i.e. *did too*) with those involving a VP-anaphoric phrase containing an overt anaphor (i.e. *did it too*). Although these VP-anaphoric forms are not used completely interchangeably in natural language, the current study investigates constructions where both are grammatical. Previous studies have shown that in these contexts, the null VP-anaphoric phrase is preferred and the overt VP-anaphoric phrase elicits increased processing. The current study expands on this work by directly comparing participants’ conscious interpretation of the null and overt VP-anaphoric phrase in constructions involving different types of antecedent VPs. The results demonstrate extremely high comprehension accuracy for unambiguous constructions involving either form. In addition, this study is the first to document experimentally that participants interpret the null and overt VP-anaphoric phrase similarly in both ambiguous contexts (i.e. both strict and sloppy interpretation available) and biased contexts (i.e. only sloppy interpretation available). Overall, the results suggest that previous observations of variation related to anaphoric form in conscious preference and unconscious processing are likely not due to differences in comprehension or interpretation.
The interpretation of null and overt VP-anaphoric constructions

The term, *anaphora*, denotes the relationship between linguistic phrases that refer to the same real-world entity or concept (i.e. the antecedent and the anaphor). Anaphors appear in a variety of forms ranging from those that carry detailed semantic information (e.g. a repeated noun phrase) to those that carry only featural information (e.g. an overt pronoun marked for gender and number) to those that are unpronounced and carry no information (e.g. a null pronoun). Anaphors also correspond to a variety of phrasal categories such as noun phrases (NPs), verb phrases (VPs), or even whole inflectional phrase (IPs) (i.e. clauses). Although NP-anaphoric phrases (e.g. pronouns) have received the most study to date in the psycholinguistic literature, researchers have recently begun to examine the processing of VP-anaphoric and IP-anaphoric phrases in more depth.

The research in the current study focuses on sentences involving VP-anaphoric phrases that contain either a null verb phrase (VP) anaphor (i.e. *did too*) or an overt VP anaphor (i.e. *did it too*)\(^1\) and investigates the effect of this variation in anaphoric form on interpretation. For ease of exposition, we will refer to these phrases as the null VP-anaphoric phrase and the overt VP-anaphoric phrase. Despite significant overlap, the two forms are not used completely interchangeably in natural language. In particular, the two forms differ in the types of antecedents to which they can refer. While the null VP-anaphoric phrase can only be used to refer to linguistic antecedents

\(^1\) Although there is some debate regarding the linguistic analysis, critically, both forms take as their referent the event specified in the antecedent VP.
in a parallel syntactic form, the overt VP-anaphoric phrase can be used to refer to linguistic antecedents in any syntactic form or even to non-linguistic antecedents (see Murphy, 1985a; Hankamer & Sag, 1976; Sag & Hankamer, 1984, for further discussion).

Direct comparisons of the on-line processing of constructions involving the null and overt VP-anaphoric phrases have revealed effects related to variation in the form of the anaphoric phrase. The observed effects suggest that, in contexts where both forms are grammatical, the null VP-anaphoric phrase is preferred (Callahan, Nicol, & Love, submitted; Callahan, Nicol, Love, & Swinney, submitted) and the overt VP-anaphoric phrase is associated with increased processing as indexed by increased reading time in an eye-tracking study (Callahan, Nicol, Love, Witzel, & Swinney, 2008; Callahan, Nicol, & Love, submitted) and a peaked centro-posterior negativity following the overt anaphor, it, in an ERP study (Callahan, Nicol, Love, & Swinney, 2007, submitted). Taken together, the results of these studies suggest that comprehenders are sensitive to VP-anaphoric form both in terms of conscious preference and unconscious processing. In contexts where multiple anaphoric forms are grammatical, comprehenders expect speakers to produce the least informative anaphoric form possible and violations of this expectation elicit additional processing. This sensitivity to anaphoric form is consistent with the predictions of a number of theories and models of processing that link the usage, interpretation, and processing of different anaphoric forms to the interaction between characteristics of the anaphor that determine how effective it is in uniquely identifying its antecedent and characteristics
of the antecedent that determine its accessibility to the hearer. (e.g. Almor, 1998; Ariel, 1990, 1991; Givón, 1983; Gordon & Hendrick, 1998; Murphy, 1985b).

A number of previous studies have investigated the interpretation of the null VP-anaphoric phrase (e.g. Frazier & Clifton, 2000); however, to our knowledge, there have been no experimental investigations of the interpretation of the overt VP-anaphoric phrase and no direct comparison of the conscious interpretation of the two forms in constructions where both are grammatical. In order to confidently link the previous findings of variation related to anaphoric form to an expectation for a particular form, it is important to investigate whether participants’ are equally accurate in comprehending the two forms and whether they interpret both similarly.

To this end, the current study takes advantage of an interesting property of the constructions of interest in the current study, the so-called “strict/sloppy” ambiguity. In particular, in VP-anaphoric constructions, when the antecedent VP involves an anaphor such as the possessive pronoun, his, in (1) below, the interpretation of the VP-anaphoric phrase is ambiguous.

(1) The lawyer who was worried about a possible audit saved his receipts and the accountant who knew all the tax regulations (did too/did it too), according to the IRS agent.

The ambiguity arises because the possessive pronoun contained within the VP-anaphoric phrase can be interpreted as referring to either the subject of the first
conjunct (i.e. the lawyer) or the subject of the second conjunct (i.e. the accountant).

The interpretation that the accountant saved the lawyer’s receipts is referred to as the “strict” reading since the possessive pronoun has the same reference in the antecedent and anaphor clause (i.e. it refers to the subject of the first conjunct). In contrast, the interpretation that the accountant saved the accountant’s receipts is referred to as the “sloppy” reading since the possessive pronoun has a different reference in the antecedent and anaphor clause.

Investigations of the interpretation of the strict/sloppy ambiguity in constructions involving the null VP-anaphoric phrase have revealed that, in ambiguous contexts similar to (1) above, adults\(^2\) choose the strict and sloppy interpretations equally often (Frazier & Clifton, 2000) or prefer the sloppy interpretation (Foley, Nuñez del Prado, Barbier, & Lust, 2003, fn.3; Shapiro, Hestvik, Lesan, & Garcia, 2003). Although participants typically report that both interpretations are available in ambiguous contexts, the strict interpretation can be blocked by lexical characteristics of the verb in the antecedent VP or other pragmatic factors. For example, in (2) below, the VP, *batted her eyelashes*, involves a verb of inalienable possession.

\(^2\) Although child language is beyond the scope of this paper, we note that a study by Foley and colleagues that investigated the strict/sloppy ambiguity in children showed effects generally similar to those documented in adults; nonetheless, younger children and older children did show somewhat different interpretative patterns (see Foley et al., 2003, for details).
(2) The dentist who was attractive and flirtatious *batted her eyelashes* and the artist who was a little bit tipsy (*did too/did it too*), according to the hostess at the party.

For verbs of inalienable possession, the action can only be performed on an object belonging to the subject of the verb (i.e. you can only bat your own eyelashes, not someone else’s). In these contexts in tasks probing conscious final interpretation, participants consistently choose the sloppy interpretation as the only possible interpretation (Shapiro et al., 2003).³

In both the linguistic and psycholinguistic literature, the strict/sloppy ambiguity has been discussed mostly in the context of constructions involving the null VP-anaphoric phrase (i.e. VP-elliptical constructions) (Foley et al., 2003; Shapiro & Hestvik, 1995; Shapiro et al., 2003); however, the ambiguity also arises in various other constructions including, critically, those involving the overt VP-anaphoric phrase (also deaccenting and so-called “paycheck” pronouns, see Gardent, 1997 for examples and further discussion). The current study expands on previous work by directly comparing participants’ conscious interpretation of the null and overt VP-anaphoric phrase in constructions involving ambiguous, biased, or unambiguous antecedent VPs (see Table 1 for sample sentences and the Methods section for more details). Based on intuition, we predict that for constructions involving ambiguous and

---

³ Interestingly, despite the conscious availability of only the sloppy interpretation, both the strict and sloppy readings appear to be computed at the anaphor site in unconscious on-line processing (see Shapiro & Hestvik, 1995, and Shapiro et al., 2003, for details).
biased VPs, interpretation of the two VP-anaphoric forms will be similar. We also predict that for constructions involving unambiguous VPs, comprehension accuracy will be comparable for the two VP-anaphoric forms. If the results support these predictions, it will suggest that previous observations of variation in preference and unconscious processing are unlikely to be related to differences in interpretation or comprehension accuracy.

Methods

Participants

All participants were recruited from the University of California, San Diego and received course credit for their participation. All participants were right-handed, neurologically-unimpaired, and had normal or corrected-to-normal vision. All participants were also native English speakers with no exposure to a second language before age six. Thirty-six participants were initially tested; however, one participant was excluded for accuracy below 90% on the Unambiguous conditions (see the Materials section for more details). Therefore, 35 participants were included in further analysis (6 male, 29 female; Mean age = 19.97, SD = 1.92).
Materials

For the critical experimental items, 120 items of similar form were constructed. Please see Table 1 for sample sentences and the Appendix for a full list. Each experimental item consisted of two sentences joined by the conjunction, *and*. The first conjunct consisted of a subject and a VP and the second conjunct consisted of a different subject and a VP-anaphoric phrase. For length purposes and in anticipation of a follow-up priming study, each subject was modified by a relative clause or a prepositional phrase and a final phrase followed the second conjunct. The antecedent VP in the first conjunct was varied among three types: 1) an ambiguous VP involving a transitive verb with an object modified by a possessive pronoun (e.g. *saved his receipts*), 2) a pragmatically-biased VP involving a verb of inalienable possession with an object modified by a possessive pronoun (e.g. *batted her eyelashes*), or 3) an unambiguous VP involving a transitive verb with an object modified by a lexical possessor (e.g. *pulled the criminal's file*). Within each of these verb types, the VP-anaphoric phrase in the second conjunct was also varied between two forms: 1) a null VP-anaphoric phrase (e.g. *did too*) or 2) an overt VP-anaphoric phrase (e.g. *did it too*). Thus, the crossing of the factors antecedent VP type and VP-anaphoric form resulted in six conditions: 1) Ambiguous-Null (A-N), 2) Ambiguous-Overt (A-O), 3) Biased-Null (B-N), 4) Biased-Overt (B-O), 5) Unambiguous-Null (U-N), and 6) Unambiguous-Overt (U-O). Half of the items in the Biased conditions (i.e. 20 items) were taken with minor modification from the stimuli involving verbs of inalienable possession in Shapiro, Hestvik, Lesan, & Garcia (2003).
The 120 experimental items were assembled into a single pseudo-randomized list with the constraint that no experimental condition ever appeared more than twice in a row. Two versions of the list were created by systematically varying the form of the VP-anaphoric phrase across the lists. Although each item only appeared once in a given list, across lists each item appeared with both VP-anaphoric forms. Participants responded to items in all six conditions, but only read any given item once. Thus, over the course of the experiment, each participant was presented with 20 items from each of the six critical conditions (A-N, A-O, B-N, B-O, U-N, U-O).

**Procedure**

Before beginning the experiment, participants gave informed consent and completed a questionnaire about their language use and language experience. Next, the participants received instructions for the rating task. Participants were instructed to read each sentence (e.g. *The lawyer who was worried about a possible audit saved his receipts and the accountant who knew all the tax regulations did too, according to the IRS agent.*) and answer a question regarding their interpretation of the VP-anaphoric phrase (e.g. *Whose receipts did the accountant save?*). So that participants could indicate whether they were aware of more than one possible interpretation as well as their relative preference for each interpretation, participants answered this question using a Likert scale ranging from 1 to 7 where 1 indicated that Subject 1 was the only possible interpretation (e.g. *The lawyer’s*) and 7 indicated that Subject 2 was the only possible interpretation (e.g. *The accountant’s*). For the Ambiguous and Biased
conditions, Subject 1 corresponded to the strict interpretation while Subject 2 corresponded to the sloppy interpretation. For the Unambiguous conditions, the correct answer corresponded to 1 and 7 an equal number of times. Please see Figure 1 for a sample item from the questionnaire and an illustration of the Likert scale. Once participants indicated that they understood the instructions, the experiment began. The entire experiment lasted less than one hour.

Experimental Design

For presentation purposes, the six critical conditions were considered as part of a single 3 X 2 mixed factorial design with the factors, anaphoric form (Null vs. Overt), and antecedent VP type (Ambiguous vs. Biased vs. Unambiguous). Over the course of the experiment, each subject contributed data to each condition and each item appeared in all conditions, but a given subject contributed data for a given item in only one condition.

Results

Although the six experimental conditions were considered as part of a single mixed factorial design where each subject contributed data to each condition,
antecedent VP type was a between-items factor. Thus, for analysis, the effect of the factor anaphoric form will be evaluated separately for each antecedent VP type via paired two-tailed t-tests. The relevant statistical comparisons are described in more detail below.

**Unambiguous Conditions**

In the Unambiguous conditions, the antecedent VP contained a transitive verb and an object modified by a lexical possessor. These conditions were included in the study as a comparison of comprehension accuracy for the two VP-anaphoric forms (see Table 1 for a sample sentence and question) and also served as control conditions to ensure that participants were paying attention and completing the task to the best of their ability. Overall, response accuracy for the Unambiguous items was extremely high (Mean = 97.93%, SD = 2.94%). Moreover, as shown in Table 2, a comparison of the Unambiguous-Null and the Unambiguous-Overt conditions revealed that mean accuracy was similar for the two VP-anaphoric phrase forms (U-N: Mean = 98.43%, SD = 2.65%; U-O: Mean = 97.43%, SD = 3.91%). This pattern was confirmed by the results of a paired two-tailed t-test in which the difference was only marginally significant by subjects and not significant by items ($t_1(34) = 1.87, p = .07, t_2(39) = 1.41, p = .17$).

The majority of responses to Unambiguous items took the form of ratings of 1 or 7; however, in approximately 10% of responses to Unambiguous items, participants indicated a rating other than 1 or 7. If this rating indicated a preference for the correct interpretation, it was counted as correct. In other words, if the correct interpretation was “the criminal’s” and this corresponded to an interpretative preference rating of 1, ratings of 1-3 were accepted as correct.
Ambiguous Conditions

In the Ambiguous conditions, the VP contained a transitive verb and an object modified by a possessive pronoun (see Table 1 for a sample sentence and question). In these cases, the interpretation of the VP-anaphoric phrase is ambiguous. Since both the strict and sloppy interpretations are available, participants should not produce ratings of 1 or 7, but rather should produce ratings that fall somewhere between 2 and 6. In fact, as shown in Table 3, a comparison of the Ambiguous-Null and the Ambiguous-Overt conditions revealed that participants produced ratings that corresponded to the middle of the rating scale (A-N: Mean = 4.34, SD = .61; A-O: Mean = 4.32, SD = .68, indicating that both interpretations were equally available. Moreover, when the two conditions were directly compared via a paired two-tailed t-test, no significant difference related to the anaphoric form was observed ($t_1(34) = .21, p = .84$, $t_2(39) = .22, p = .83$).
Biased Conditions

The Biased conditions involved antecedent VPs containing verbs of inalienable possession and an object modified by a possessive pronoun (see Table 1 for a sample sentence and question). For verbs of inalienable possession (e.g. bat the eyelashes), the action can only be performed on an object belonging to the subject of the verb (i.e. you can only bat your own eyelashes, not someone else’s). Therefore, regardless of VP-anaphoric form, ratings for these items should correspond to the sloppy interpretation (e.g. a rating of 7). As predicted and as shown in Table 3, a comparison of the Biased-Null and the Biased-Overt conditions revealed that participants produced ratings that corresponded to the sloppy interpretation in both conditions (B-N: Mean = 6.82, SD = .35; B-O: Mean = 6.82, SD = .35). Moreover, when the two conditions were directly compared via a paired two-tailed t-test, no significant difference related to anaphoric form was observed ($t_1(34) = .14, p = .89, t_2(39) = .10, p = .92$).

Discussion

Previous studies that investigated constructions where both the null and overt VP-anaphoric phrases are grammatical observed variation related to the form of the anaphoric phrase. Specifically, participants displayed a preference for the null VP-anaphoric phrase and the overt VP-anaphoric phrase was associated with increased processing (Callahan et al., 2008; Callahan, Nicol, & Love, submitted; Callahan,
Nicol, Love, & Swinney, 2007, submitted). These effects were interpreted as reflecting an expectation for the least informative grammatical anaphoric form in contexts with highly accessible antecedents; however, it is important to rule out the possibility that differences in comprehension or interpretation are responsible for the observed effects. To this end, the current study directly compared the comprehension and interpretation of the two VP-anaphoric forms in contexts where both are grammatical.

The first goal of the current study was to ascertain whether off-line comprehension accuracy is similar for constructions involving null and overt VP-anaphoric phrases. For unambiguous sentences involving the two VP-anaphoric forms, comprehension accuracy was extremely high overall. In addition, accuracy was similar for the two VP-anaphoric forms (U-N: 98.43%; U-O: 97.43%) with statistical analysis revealing a difference that was only marginally significant by subjects and not significant by items. Comprehension is clearly highly accurate for constructions involving either VP-anaphoric form. Even if the marginal effect reflects a real difference, the magnitude is extremely small.

The second goal of the current study was to investigate whether there are differences in the interpretation of the null and overt VP-anaphoric phrases. Participants were asked to indicate their interpretation of the two forms in ambiguous contexts (i.e. where the construction involved the strict/sloppy ambiguity). Previous psycholinguistic research on the strict/sloppy ambiguity has focused solely on the null VP-anaphoric phrase (Frazier & Clifton, 2000; Shapiro & Hestvik, 1995; Shapiro et
al., 2003); however, the strict/sloppy ambiguity can also be observed in constructions involving the overt VP-anaphoric phrase (as well as other constructions, see Gardent, 1997, for more examples). The results in the current study replicate one previous experimental investigation of the interpretation of the strict/sloppy ambiguity in constructions involving the null VP-anaphoric phrase (Frazier & Clifton, 2000) in demonstrating that participants find the strict and sloppy interpretation equally available and extend this finding to the overt VP-anaphoric phrase. Although other studies have suggested a preference for the sloppy interpretation (Foley et al., 2003, fn. 3; Shapiro et al., 2003), these previous investigations provide few details about the stimuli and methods, so it is difficult to know whether stimuli- or task-related differences across studies might be responsible for the increased preference for the sloppy interpretation. For example, the fact that participants reported their interpretation using a Likert scale in the current study may have encouraged them to generate and evaluate both possible interpretations. Future research will be needed to clarify this issue.

In addition to ambiguous contexts, the current study also investigated contexts where the strict interpretation is blocked by the pragmatic bias of the antecedent VP (i.e. constructions involving verbs of inalienable possession). Previous research on conscious interpretation in these constructions (Shapiro et al., 2003) has demonstrated that only the sloppy interpretation is available for sentences involving the null VP-anaphoric phrase. In the current study, we replicated this finding for the sentences involving the null VP-anaphoric phrase. Participants overwhelmingly indicated that
the sloppy interpretation was the only possible interpretation. We also found the exact same results for sentences involving the overt VP-anaphoric phrase. Again, participants indicated that the sloppy interpretation was the only possible interpretation.

Overall, then, the results of the current study revealed no differences in the conscious interpretation of constructions involving null and overt VP-anaphoric phrases. We tested two different antecedent VP types and found different patterns of interpretative preference based on the type of antecedent VP. Both the strict and sloppy interpretation was available when the antecedent VP was ambiguous and involved a transitive verb and a possessive pronoun; yet, when the antecedent VP was pragmatically-biased and involved a verb of inalienable possession, only the sloppy interpretation was available. To our knowledge, this study is the first to document experimentally identical patterns of interpretative preference for the null and overt VP-anaphoric phrases.

Taken together with the results from the unambiguous contexts, the overall pattern suggests that participants are equally good at retrieving and integrating antecedent information at a VP-anaphoric phrase regardless of the form, at least in constructions where both forms are grammatical and when given the unlimited time of an off-line task; moreover, these results provide complementary evidence that previous observations of processing variation related to anaphoric form are likely not due to differences in overall comprehension accuracy or interpretation. Thus, the results of the current study highlight the fact that variation in processing may be
observed despite a lack of variation in conscious comprehension and interpretation. In order to develop broad theories and models of anaphoric processing that can account for a wide range of effects, it is important to examine both differences and similarities in the processing and interpretation of forms. We suggest that research that directly compares different anaphoric forms is particularly relevant and can reveal important generalizations about anaphoric processing and interpretation.
References


Appendix

Unambiguous Items

1. The cop who was investigating the robbery pulled the criminal's file and the attorney who was preparing for the defense (did too/did it too), according to the clerk at the courthouse.
2. The banker who was a stickler for the rules processed the applicant's paperwork and the real estate agent who would make a large commission off the sale (did too/did it too), according to the seller of the house.
3. The mom who had planned the whole party frosted the kid's birthday cake and the sister who had come up with the mermaid theme (did too/did it too), according to the neighbor who blew up the balloons.
4. The pianist who loved performing in front of an audience memorized the composer's music and the dancer who had been practicing the steps for weeks (did too/did it too), according to the director of the production.
5. The surgeon with the Ivy League education read the patient's chart and the neurologist with the warm and caring demeanor (did too/did it too), according to the nurse assigned to the ward.
6. The engineer with the expensive black suit printed the consultant's invoice and the administrator with the trendy designer eyeglasses (did too/did it too), according to the secretary at the firm.
7. The deliveryman who was in a big hurry dropped the tenant's package and the doorman who hadn't received a Christmas tip (did too/did it too), according to the security camera in the lobby.
8. The tailor with twenty years experience constructing wedding dresses examined the designer's fabric and the bride with the beautiful long blonde hair (did too/did it too), according to the maid of honor.
9. The commentator whose contract had just been renewed questioned the judge's ruling and the prosecutor who had been involved in the case (did too/did it too), according to the recap on the nightly news.
10. The juggler who had four balls in the air observed the child's face and the magician who was pulling a rabbit out of a hat (did too/did it too), according to the mother standing in the crowd.
11. The politician who had recently been re-elected presented the lobbyist's position and the legislator who had received large campaign contributions (did too/did it too), according to the special prosecutor.
12. The boy in the back of the room answered the teacher's question and the girl at the desk by the door (did too/did it too), according to the principal observing the class.
13. The aunt who was out of town for the party signed the relative's birthday card and the uncle who had a previous engagement (did too/did it too), according to the niece who delivered the envelope.
14. The rebel who needed food to survive pursued the captain's boat and the colonel who was investigating arms smuggling on the river (did too/did it too), according to the natives on the shore.
15. The patrolman who was just ending a long shift identified the suspect's car and the victim who had been robbed at gunpoint (did too/did it too), according to the district attorney.
16. The social worker in charge of the case investigated the baby's safety and the judge hearing arguments in family court (did too/did it too), according to the grandmother.
17. The sniper who was trained for hostage situations circled the gunman's house and the policeman who was worried about the children inside (did too/did it too), according to the chief of police on the scene.
18. The columnist from the New York Times reviewed the actor's play and the critic from the New York Post (did too/did it too), according to the message left by the agent.
19. The niece who had grown a lot lately returned the grandfather's gift and the cousin who hated the color pink (did too/did it too), according to the saleswoman at the store.
20. The inmate who made the confession requested the prosecutor's presence and the officer who did everything by the book (did too/did it too), according to the prison guard.
21. The dermatologist who was invited to a party located the host's house and the pilot who brought a bottle of wine (did too/did it too), according to the other attendees.
22. The father who had a bad temper kicked the policeman's car and the adolescent who always got in trouble (did too/did it too), according to the police report.
23. The king on the jewel-encrusted throne kissed the queen's cheek and the diplomat from the Italian consulate (did too/did it too), according to the video shown on the evening news.
24. The husband who was always very chivalrous carried the wife's bags and the bellboy who wanted a big tip (did too/did it too), according to the desk clerk.
25. The clerk in the bustling train station punched the baker's ticket and the conductor on the platform at the transfer point (did too/did it too), according to company records.
26. The boy who was known as a troublemaker yanked the sister's hair and the baby who laughed and giggled (did too/did it too), according to the mom who saw the whole thing.
27. The chef at the annual Pillsbury bakeoff evaluated the contestant's entry and the host from the television network (did too/did it too), according to the viewers watching at home.
28. The celebrity who was making a big comeback perused the reporter's article and the publicist who had arranged the interview (did too/did it too), according to the maid dusting the piano.
29. The cook in the fancy four-star restaurant followed the chef's recipe and the wife who had bought the best-selling cookbook (did too/did it too), according to the delighted diners.

30. The teen who lived down the street babysat the mayor's children and the nanny who was originally from Sweden (did too/did it too), according to the mothers at the park.

31. The chemist who was skeptical of the results replicated the researcher's experiment and the physicist who was interested in broader applications (did too/did it too), according to the papers published in the journal.

32. The author who was a native New Yorker autographed the fan's book and the illustrator who was originally from Spain (did too/did it too), according to the security guard monitoring the crowd.

33. The friend who had gone to the same high school defended the suspect's innocence and the relative who had lived next door for years (did too/did it too), according to the detective on the case.

34. The drycleaner who had owned the store for twenty years folded the man's shirts and the cashier who wasn't busy at the moment (did too/did it too), according to the woman picking up a dress.

35. The butler who was trained in England polished the grandmother's silver and the housekeeper who was getting ready for the dinner party (did too/did it too), according to the caterer delivering the food.

36. The principal of the local high school criticized the adolescent's behavior and the mother with three small children (did too/did it too), according to the secretary listening at the door.

37. The cop who had witnessed the incident complimented the woman's courage and the anchorwoman who heard the story (did too/did it too), according to the relative watching the report.

38. The merchant with the short black curly hair displayed the celebrity's merchandise and the hairstylist with the long painted nails (did too/did it too), according to the company representative.

39. The saleswoman who was excited to make such a big sale calculated the socialite's bill and the manager who was worried about errors (did too/did it too), according to the personal assistant.

40. The tenant who was disgruntled about being evicted destroyed the landlord's apartment and the vandal who found the door unlocked (did too/did it too), according to the testimony in court.

*Ambiguous Items*

41. The artist who rented a studio downtown washed his car and the writer who liked being outside on a sunny afternoon (did too/did it too), according to the kids playing on the corner.
42. The fisherman who was always especially careful wore his life jacket and the lifeguard who had extensive safety training (did too/did it too), according to the man conducting the surprise inspection.
43. The cowboy who was getting ready to compete inspected his boots and the clown who would be working in the ring that night (did too/did it too), according to the announcer at the rodeo.
44. The manufacturer who was late for a flight to Hong Kong initialed his contract and the diplomat who really wanted to close the deal quickly (did too/did it too), according to the attorney who prepared the papers.
45. The miner who might have been exposed to the toxin noted his temperature and the paramedic who was feeling nauseous and lightheaded (did too/did it too), according to the center for disease control.
46. The barber who was obsessed with cleanliness swept his floor and the vendor who had tracked in some dirt from outside (did too/did it too), according to a mailman who saw everything.
47. The roofer who got stuck with the cramped middle seat on the plane adjusted his pillow and the plumber who realized that the movie screen was blocked (did too/did it too), according to the stewardess serving drinks.
48. The psychiatrist who worked late almost every day walked her dog and the housewife who lived down the street (did too/did it too), according to joggers in the neighborhood.
49. The builder who was working on a large construction project checked his inventory and the custodian who was head of the local union (did too/did it too), according to the annual report issued at the meeting.
50. The hairstylist who was sitting chatting on the front porch patted her cat and the designer who had been working hard in the garden (did too/did it too), according to the nosy neighbor across the street.
51. The explorer who had just gotten back from a long trip edited his manuscript and the producer who was considering new projects (did too/did it too), according to the gossip columnist.
52. The agent who spent all day talking on the phone watered her plants and the illustrator who was known for having quite a green thumb (did too/did it too), according to the woman at the adjacent desk.
53. The politician who was moving to a new office dusted her bookshelves and the secretary who was procrastinating on writing a letter (did too/did it too), according to the janitor working in the hall.
54. The teenager who was usually extremely lazy rinsed his plate and the dishwasher who always left the kitchen spotless (did too/did it too), according to the cook stirring the soup.
55. The sculptor who wanted better light in the afternoon moved her table and the architect who needed more space for storing supplies (did too/did it too), according to the landlord who rented the loft.
56. The engineer who had just finished the marathon took his pulse and the doctor who was also an avid runner (did too/did it too), according to the organizer of the race.

57. The grandmother who was making ice cream sundaes for dessert licked her spoon and the child who couldn't wait a second longer (did too/did it too), according to the nephew waiting at the table.

58. The singer who had unrealistic expectations for the meeting brought his contract and the entrepreneur who was organizing the concert (did too/did it too), according to the impartial mediator.

59. The mathematician who had recently bought a weedwacker admired his grass and the exterminator who mowed the lawn every Saturday (did too/did it too), according to the handyman washing the windows.

60. The dean who was showing an important slide used his pointer and the chemist who was a visiting scholar from Harvard (did too/did it too), according to the faculty who were at the talk.

61. The designer who was focused on resolving the argument ignored her phone and the supplier who was determined to get a fair price (did too/did it too), according to the assistant taking notes.

62. The dermatologist who usually avoided eating chocolate tasted her dessert and the model who only indulged once in a while (did too/did it too), according to the astonished waiter.

63. The physicist who was applying for a patent tested his machine and the mechanic who wanted to revolutionize the industry (did too/did it too), according to the supervising lab technician.

64. The astronomer who was submitting a new grant application revised his document and the meteorologist who had been awarded many grants (did too/did it too), according to the drafts found in the trash.

65. The scholar who had won the nobel prize last year praised his book and the intellectual who taught at Columbia University (did too/did it too), according to the librarian in charge of purchasing.

66. The executive who was the head of the marketing department forwarded his proposal and the programmer who supervised the production team (did too/did it too), according to the conclusions of the internal investigation.

67. The teller who was extremely conscientious counted her money and the manager who had already been given a warning (did too/did it too), according to the tape from the hidden security camera.

68. The bridesmaid who hated the purple dress with puffy sleeves sniffed her corsage and the neighbor who thought the ceremony was boring (did too/did it too), according to the photographer in the front pew.

69. The cheerleader who was getting ready for the big competition curled her hair and the mother who had driven all the way from Georgia (did too/did it too), according to the person delivering room service.
70. The driver who was responsible for the accident assessed his injuries and the passenger who hadn't been wearing a seatbelt (did too/did it too), according to witnesses at the scene.
71. The nurse who had been a vegetarian since childhood analyzed her diet and the psychologist who only ate organic produce (did too/did it too), according to the personal trainer.
72. The electrician who had just moved into the neighborhood painted his apartment and the handyman who enjoyed a good weekend project (did too/did it too), according to the landlord of the building.
73. The inmate who had information about the murder reviewed his options and the prosecutor who had recently been assigned to the case (did too/did it too), according to the cop who made the arrest.
74. The speaker who always sold out the large venue quoted his book and the anchorman who was a local legend in the state (did too/did it too), according to a blogger on the internet.
75. The graduate who was excited about the ceremony ironed his suit and the father who accepted nothing less than perfection (did too/did it too), according to the housekeeper at the hotel.
76. The lawyer who was worried about a possible audit saved his receipts and the accountant who knew all the tax regulations (did too/did it too), according to the IRS agent.
77. The salesman who usually worked on Mondays visited his mother and the waiter who only ever worked the weekend shifts (did too/did it too), according to the nursing home sign in sheet.
78. The translator who had taken dictation for a letter corrected his mistake and the attorney who had worked very late the night before (did too/did it too), according to the records in the file.
79. The diplomat who wanted to cover all the bases recorded his meetings and the spy who understood the power of blackmail (did too/did it too), according to the testimony at the trial.
80. The housewife who was accused of running a scam admitted her involvement and the waitress who was under great pressure from the detective (did too/did it too), according to official transcripts.

**Biased Items** (Items 81-99 adapted from Shapiro et al., 2003)

81. The gambler who had just won ten hands in a row winked his eyes and the butler who was in on the elaborate scheme (did too/did it too), according to the private eye at the house.
82. The bride who was heading for the bathroom shuffled her feet and the pianist who had a blister on her big toe (did too/did it too), according to the guests waiting for drinks.
83. The jeweler who was just suspicious by nature raised his eyebrows and the astronaut who had actually been cheating for years (did too/did it too), according to the others at the table.
84. The skier with the badly broken leg twiddled his thumbs and the boxer with a couple of rib fractures (did too/did it too), according to the world-famous surgeon.
85. The kid who desperately wanted some french fries smacked his lips and the baker who was thirty-five pounds overweight (did too/did it too), according to the other customers watching the demonstration.
86. The fisherman who had always been allergic to pollen twitched his nose and the rancher who had a very bad head cold (did too/did it too), according to the passengers on the bus trip.
87. The caddie who was always extremely impatient drummed his fingers and the plumber who hated waiting for anything (did too/did it too), according to the other people in the group.
88. The conductor who was a complete perfectionist tossed his hair and the student who was very vain about his looks (did too/did it too), according to the person sitting in the third row.
89. The dentist who was attractive and flirtatious batted her eyelashes and the artist who was a little bit tipsy (did too/did it too), according to the hostess at the party.
90. The acrobat who was extremely conceited and vain flexed his muscles and the bodyguard who had an inflated self-image (did too/did it too), according to someone else in the weight room.
91. The chef who was frustrated by the blackout clenched his fists and the dishwasher who was secretly afraid of the dark (did too/did it too), according to the delivery truck driver.
92. The criminal who would not admit any guilt blinked his eyes and the reporter who was determined to get the story (did too/did it too), according to the cameraman filming the interview.
93. The paramedic who was in charge of the chaotic scene wrinkled her forehead and the neighbor who was afraid it might be too late (did too/did it too), according to the concerned on-lookers.
94. The priest who saw the display of Halloween candy clicked his tongue and the barber who was a health food enthusiast (did too/did it too), according to the person behind them in the checkout line.
95. The zoo-keeper who hated working overtime pouted her lips and the messenger who could not make the scheduled drop (did too/did it too), according to the first draft of the novel.
96. The king who was always impatient in meetings cleared his throat and the mayor who grew up in New York City (did too/did it too), according to the new unauthorized biography.
97. The plaintiff who feared testifying in front of an audience fluttered his eyelashes and the translator who was holding back a flood of tears (did too/did it too), according to the court reporter.
98. The burglar who was hiding in the shadows hunched his shoulders and the
doorman who wore only a thin summer coat (did too/did it too), according to the
hot dog vendor.
99. The Martian who was undergoing interrogation rolled his eyes and the detective
who had been the first on the scene (did too/did it too), according to the cover
story of the Enquirer.
100. The cowboy who suggested making the wager shrugged his shoulders and the
carpenter who was not worried about losing (did too/did it too), according to eye-
 witnesses questioned later.
101. The accountant who believed strongly in the power of prayer bowed his head
and the inventor who had been an Atheist for many years (did too/did it too),
according to the other people at the funeral.
102. The hiker who was afraid of surprising a wild animal snapped his fingers and
the biologist who was concerned about safety (did too/did it too), according to the
leader of the group.
103. The veterinarian who was taking swimming classes held his breath and the
olympian who had lots of underwater experience (did too/did it too), according to
the swimming instructor.
104. The professor who was too far away from the blackboard squinted his eyes and
the physician who was extremely nearsighted (did too/did it too), according to
another man sitting in the back row.
105. The runner who knew that a strong core improved performance tightened his
abs and the teenager who saw a pretty girl on the sidelines (did too/did it too),
according to spectators on the sidelines.
106. The millionaire who was listening to the orchestra tapped his foot and the
senator who had been given tickets by a friend (did too/did it too), according to the
newspaper article.
107. The singer who was nervous about the performance ground his teeth and the
vendor who hated the loud squealing of feedback (did too/did it too), according to
the manager of the venue.
108. The monk who meditated for an hour every day focused his attention and the
banker who came to the retreat for some solitude (did too/did it too), according to
the journals kept by participants.
109. The coach who was checking out the sold-out crowd craned his neck and the
custodian who was preparing for a long night (did too/did it too), according to fans
in the stands.
110. The juggler who was carefully walking the tightrope pointed his toes and the
clown who was doing a funny dance (did too/did it too), according to the laughing
crowd.
111. The guitarist who just found out the performance was cancelled cocked his
head and the entrepreneur who was confused about the reason (did too/did it too),
according to the assistant who informed them of the change.
112. The swimmer who was waiting for the starting whistle tensed his body and the photographer who was preparing to capture the action (did too/did it too), according to the referee standing nearby.

113. The lawyer who had been working all night long rubbed his eyes and the writer who could barely see the words on the page (did too/did it too), according to reports of the meeting.

114. The bullfighter who was courageous and excited stamped his feet and the sailor who was looking on from the stands (did too/did it too), according to the reporter who was there.

115. The landlord who thought the check had not been sent furrowed his brow and the mailman who was responsible for the delivery (did too/did it too), according to the investigator of the incident.

116. The soldier who was fleeing the battlefield at night strained his hamstring and the surgeon who was carrying another wounded man (did too/did it too), according to the official report.

117. The jockey who was warming up for the race swiveled his hips and the stuntman who was practicing jumps involving twists (did too/did it too), according to the trainers at the gym.

118. The principal who was very disappointed in the situation shook his head and the relative who had been called away from work (did too/did it too), according to the child who was in trouble.

119. The comedian who agreed with the terms of the contract nodded his head and the publisher who had made the initial offer (did too/did it too), according to witnesses present.

120. The musician who had been playing for hours cracked his knuckles and the waiter who had just finished a long shift (did too/did it too), according to the observant bartender.
Table 1. Sample Sentences

**Unambiguous Conditions**

The cop who was investigating the robbery pulled the criminal’s file and the attorney who was preparing the defense (did too/did it too), according to the clerk at the courthouse. QUESTION: Whose file did the attorney pull?

**Ambiguous Conditions**

The lawyer who was worried about a possible audit saved his receipts and the accountant who knew all the tax regulations (did too/did it too), according to the IRS agent. QUESTION: Whose receipts did the accountant save?

**Biased Conditions**

The dentist who was attractive and flirtatious batted her eyelashes and the artist who was a little bit tipsy (did too/did it too), according to the hostess at the party. QUESTION: Whose eyelashes did the artist bat?
Table 2. Mean Accuracy (in %) and Standard Deviations for the Unambiguous Conditions

<table>
<thead>
<tr>
<th>Condition</th>
<th>Accuracy (in %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unambiguous-Null</td>
<td>98.43 (2.65)</td>
</tr>
<tr>
<td>Unambiguous-Overt</td>
<td>97.43 (3.91)</td>
</tr>
</tbody>
</table>

Note. Standard deviations appear in parentheses.
Table 3. Mean Naturalness Ratings and Standard Deviations

<table>
<thead>
<tr>
<th>Naturalness Rating</th>
<th>Ambiguous Conditions</th>
<th>Biased Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1 = Strict, 7 = Sloppy)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ambiguous-Null 4.34 (.68)</td>
<td>Biased-Null 6.82 (.35)</td>
</tr>
<tr>
<td></td>
<td>Ambiguous-Overt 4.32 (.68)</td>
<td>Biased-Overt 6.82 (.35)</td>
</tr>
</tbody>
</table>

Note. The strict interpretation corresponds to a rating of 1 and the sloppy interpretation corresponds to a rating of 7. Standard deviations appear in parentheses.
SAMPLE QUESTION

The lawyer who was worried about a possible audit saved his receipts and the accountant who knew all the tax regulations did too, according to the IRS agent.

Whose receipts did the accountant save?

the lawyer’s  1  2  3  4  5  6  7  the accountant’s

1 = "The lawyer’s" is the only possible interpretation.  
2 = I really prefer "the lawyer’s" but "the accountant’s" is also a possible interpretation.  
3 = I somewhat prefer "the lawyer’s" but "the accountant’s" is also a possible interpretation.  
4 = I like "the lawyer’s" and "the accountant’s" equally as possible interpretations.  
5 = I somewhat prefer "the accountant’s" but "the lawyer’s" is also a possible interpretation.  
6 = I really prefer "the accountant’s" but "the lawyer’s" is also a possible interpretation.  
7 = "The accountant’s" is the only possible interpretation.

Figure 1. Sample question from the questionnaire illustrating the Likert scale.
Chapter 5

Parallelism effects and verb activation: The sustained reactivation hypothesis
Abstract

The current study investigated the processes underlying parallelism by evaluating the activation of a parallel element (i.e. a verb) throughout and-coordinated sentences. Four points were tested: 1) approximately 1600ms after the verb in the first conjunct (PP1), 2) immediately following the conjunction (PP2), c) approximately 1100ms after the conjunction (PP3), d) at the end of the second conjunct (PP4). The results revealed no activation for the verb at PP1, suggesting decay occurs relatively quickly following initial activation; however, activation was observed at PP2, PP3, and PP4, suggesting the conjunction elicits reactivation that is sustained until the end of the second conjunct. These findings support a specific hypothesis about parallelism, the sustained reactivation hypothesis. This hypothesis claims that, in conjoined structures, a cue that implies parallelism elicits the reactivation of material from the first conjunct and that this activation is sustained until integration with the second conjunct can be completed.
Parallelism effects and verb activation: The sustained reactivation hypothesis

The term, parallelism, refers to a likeness, correspondence, or similarity. In natural language, coordinate structures are the prototypical environment for the expression of parallelism. Along these lines, participants who are asked to comprehend sentences strongly prefer parallel conjuncts as in (1) to nonparallel conjuncts as in (2) and processing for the second conjunct is facilitated by parallelism between the conjuncts (the so-called parallelism effect, originally reported by Frazier et al., 1984).¹

(1) Bob vacuumed the living room and Joe mopped the kitchen.
(2) Bob vacuumed the living room and the kitchen was mopped by Joe.

Since this report, researchers have investigated parallelism effects with the goal of building a successful theory of parallelism. In the following section, we summarize previous research in terms of its import for the fundamental aspects of such a theory: namely, the scope of parallelism, the types of representations it operates upon, and the types of processes that are involved.

¹ A similar facilitation is observed in production (originally reported by Bock, 1986) although it is unclear whether effects in the two domains reflect the same underlying mechanisms (for discussion, see Frazier et al., 2000; Knoeferle & Crocker, accepted).
**Previous Research**

In the first report of parallelism effects on the processing of coordinate sentences (Frazier et al., 1984), reading times were compared for second conjuncts that were parallel or non-parallel to the first conjunct on one of several dimensions (e.g. voice, attachment site, thematic roles, animacy). Strikingly, faster reading times were observed for the second conjunct when the conjuncts were parallel on any of the dimensions tested. A later study (Frazier et al., 2000) found the same parallelism effect even when the acceptability of the parallel and non-parallel forms was controlled and also demonstrated parallelism effects related to parallelism in the internal structure of the conjunct.

Knoeferle and colleagues conducted a series of eye-tracking experiments to investigate parallelism effects in German (Apel et al., 2007; Knoeferle and Crocker, accepted). These experiments replicated previous findings of parallelism effects in English in both structurally ambiguous and unambiguous contexts. Furthermore, these investigations found similar effects regardless of whether the constituent order was marked or unmarked (i.e. canonical). Finally, a critical experiment (Knoeferle & Crocker, accepted, Exp. 3) manipulated both syntactic parallelism (i.e. constituent order) and semantic parallelism (i.e. semantic similarity of constituents). Although the effect of syntactic parallelism was larger, both syntactic and semantic parallelism resulted in parallelism effects; moreover, the strongest parallelism effects occurred when both types of information were parallel, suggesting an additive nature to parallelism.
The pattern of results in these previous studies suggests a number of conclusions about the scope, processes, and underlying representations of parallelism effects. First, since these effects are observed for sentences that do not involve syntactic ambiguity as well as those that do (Apel et al., 2007; Frazier et al., 1984; Frazier et al., 2000; Knoeferle & Crocker, accepted), the scope of parallelism is not restricted to contexts involving syntactic ambiguity. Thus, parallelism does not hinge on an ambiguity resolution mechanism. Second, although parallelism effects have been observed in a wide range of contexts, two previous experiments that tested non-conjoined contexts (Frazier et al., 2000, Exp. 4; Apel et al., 2007, Exp. 2) failed to find parallelism effects. Thus, it appears that the scope of parallelism is restricted to coordinate structures. Along these lines, a very interesting study (Knoeferle, 2007) investigated whether parallelism effects in coordinate structures are tied to the use of specific conjunctions. They found parallelism effects in and- and while-conjoined contexts where the conjunction implies parallelism of the conjuncts, but not in but-conjoined contexts where it does not. This finding suggests a critical role for semantic or discourse information and a link between parallelism effects and the discourse relationship between the conjuncts as represented by a particular conjunction (c.f. Kehler, 2002; Knoeferle & Crocker, accepted; Shapiro & Hestvik, 1995). Third, the observation of parallelism effects related to parallel features that are not controlled by the grammar (e.g. the internal structure of conjuncts, Frazier et al., 2000, Exp. 3; semantic similarity, Knoeferle & Crocker, accepted, Exp. 3) suggests that the mechanisms underlying parallelism effects are related to the processor rather than to
the grammar. This processing explanation is also supported by the fact that parallelism effects occur even when the parallel and non-parallel contexts are equally acceptable. Finally, since parallelism effects are observed for features that are not considered in an early structural analysis (e.g. animacy, discourse linking, semantic similarity) and are not observable in the earliest processing measures (c.f. Frazier et al., 2000; Knoeferle & Crocker, accepted), the underlying mechanism is likely not the facilitation of specific early syntax-based parsing routines.

*The Reactivation Hypothesis*

Since parallelism effects are restricted to some extent in terms of scope (i.e. to conjoined contexts where the conjunctions imply parallelism), parallelism effects cannot simply reflect the increased activation of recently processed structural material (i.e. pure syntactic priming mechanism); nonetheless, it is still possible that parallelism effects are related to the activation of the previously processed parallel phrase (i.e. a more general priming mechanism). If the activated representation of the previously processed material contains syntactic and semantic information, then such a mechanism would be consistent with findings that semantic information can elicit parallelism effects.

We suggest one intriguing theory about the processes involved in such a priming mechanism: the sustained reactivation hypothesis. This hypothesis makes use of a well-known context-driven memory-based reactivation process. This process has been demonstrated to be operative in sentence processing in the context of relations
between phrases that refer to the same real-world entity or concept (i.e. an antecedent and an anaphor). Specifically, as suggested by numerous studies (see Nicol & Swinney, 2002, for a review), the presentation of an anaphor elicits the immediate reactivation of the previously processed antecedent information from memory for integration with the surrounding context (a process reflected in priming for a probe related to the antecedent). Since the sustained reactivation hypothesis relies on a process already documented to be operative in sentence processing, it is attractively parsimonious. Furthermore, there is good reason to believe that similar processes might be involved in processing parallelism and anaphoric relationships. Antecedents and anaphors commonly occur as parallel elements, although they are not restricted to parallel environments, and parallelism has been proposed as an important mechanism in anaphoric interpretation (c.f. Chambers & Smyth, 1998; Kehler, 2002; Smyth, 1994).

Essentially, in the context of parallelism, the sustained reactivation hypothesis proposes that when a cue that implies parallelism is encountered, previously processed material is reactivated. This material then remains active throughout the second conjunct until it can be integrated with material from the second conjunct. One consequence of this activation is facilitated processing for linguistic material that is parallel to the reactivated material on any of a variety of relevant dimensions. Specifically, the sustained reactivation hypothesis claims that, in coordinate structures, a conjunction that implies parallelism (e.g. *and*, see Knoeferle, 2007 for details) elicits the reactivation of the previous conjunct so that the two conjuncts can be integrated;
however, the sustained reactivation hypothesis is theoretically compatible with the idea that other linguistic (or potentially even non-linguistic) cues to parallelism could also elicit reactivation.

The current study investigates the sustained reactivation hypothesis by examining the activation of a parallel element in *and*-coordinated sentences such as (3) below.

(3) The doctor read the chart of the child with the broken arm during his morning rounds, and the insurance agent in the tacky suit did as well in order to become more familiar with the case.

In keeping with the majority of previous research on parallelism, this study tests contexts that involve parallelism between two verb phrases (VPs). Using cross-modal lexical priming techniques that are sensitive to the activation of words within a sentence, we evaluate the activation of the verb from the first conjunct (e.g. *read*) at various points throughout the sentence. The sustained reactivation hypothesis predicts that the activation of the verb will decay after it is initially processed in the first conjunct and then be reactivated in the second conjunct following the conjunction. Furthermore, this hypothesis predicts that the activation of the verb will be sustained throughout the second conjunct.

---

2 The VP in the second conjunct is represented by the VP-anaphoric phrase, *did as well.* Using a VP-anaphoric phrase rather than a repeated VP is more natural in these contexts.
Verbs are a good candidate for testing sentence-level parallelism effects because they play a crucial syntactic and semantic role in the sentence. In addition to denoting the event, the verb carries information about the participants in the event and the roles they play. Furthermore, during sentence processing, verbs have been demonstrated to immediately activate not only their meaning (e.g. de Goede, 2006; Callahan et al., in prep) but also typically associated arguments, argument structures, and thematic role information (Altmann & Kamide, 1999; Boland et al., 1990; Ferretti et al., 2001; Friederici & Frisch, 2000; McRae et al., 1997; McRae et al., 1998; Shapiro et al., 1987, 1989). Of even greater relevance to the current study, comprehenders have been demonstrated to use the information provided by verbs to facilitate processing by generating expectations about upcoming context (e.g. Altmann & Kamide, 1999; Shapiro et al., 1993; Trueswell & Kim, 1998; Trueswell et al., 1993).

Experiment 1- Verb activation before and after the conjunction

Methods

Participants

The participants in this and the following experiment were recruited from the University of California, San Diego and received course credit for their participation. All participants in both experiments were right-handed, neurologically healthy, and had normal hearing and normal or corrected-to-normal vision. All participants were
also native English speakers with no exposure to a second language before age six. Thirty-nine participants were initially tested in Experiment 1; however, five participants were excluded from further analysis for error rates on the on-line naming task that exceeded 10% on experimental items or 15% on all items. In addition, data from one participant were lost due to computer error. Therefore, 33 participants were included in further analysis for Experiment 1 (4 male, 25 female, 4 unknown; mean age = 20.4, range = 18 - 34).

Materials

For the critical experimental items, 40 items of similar form were constructed. (Please see the Appendix for a full list). Each experimental item consisted of two sentences joined by the conjunction, *and*. The first conjunct began with a subject followed by a VP. The second conjunct began with a different subject followed by an anaphoric phrase (e.g. *did too, did the same, did as well, etc.*). Anaphoric phrases were used in this study because they occur very frequently in parallel constructions and typically sound more natural than a repetition of the VP from the first conjunct (a variety was used to lessen expectation effects). In addition, using anaphoric phrases made it possible to test downstream points because the verb was not directly presented in the second conjunct. A sample sentence appears in (4) below with the critical verb in italics.

---

3 Unfortunately, due to an error, gender and age information as well as comprehension scores for four participants in this experiment were lost.
(4) The doctor *read* the chart of the child with the broken arm during his morning rounds, and the insurance agent in the tacky suit did as well in order to become more familiar with the case. (Probes: reviewed (R), reserved (C))

For each experimental item, related (R) and control (C) probe words were selected (see (4) above for examples). Since the activation of a verb was being tested, all probe words were verbs. Related probe words (e.g. *reviewed*) were semantically related to the verb (e.g. *read*) in the first conjunct and were presented in the same tense. Control probe words (e.g. *reserved*) were also presented in the same tense and were matched to related probe words on length, frequency, and mean naming reaction time in an unprimed context, but critically were not semantically related to the antecedent verb or any other word preceding the test points. A pre-test was conducted to confirm that related probe verbs were strongly associated with the verb and control probe verbs were not. A separate group of 48 participants who did not participate in Experiment 1 or 2 were presented with 80 pairs of verbs. Each verb (e.g. *read*) was presented twice, once paired with the related probe verb (e.g. *reviewed*) and once with the control probe verb (e.g. *reserved*). Participants rated how associated each pair of verbs was using a Likert scale ranging from 1 (not associated at all) to 5 (very associated). Related probe verbs were rated as being highly associated to the verbs (Mean = 4.10; SD = .36) while control probe verbs were not (Mean = 1.48; SD = .25).
Paired two-tailed t-tests revealed that this result was highly significant by participants and items (\(t_1(47) = 53.44, p = .000\); \(t_2(39) = 25.16, p = .000\))^4.

During the uninterrupted aural presentation of the experimental items, probe words were presented visually at one of two points (signified by superscripts\(^1\) and \(^2\) above), the positions of which varied based on the experiment. In Experiment 1, the two probe points occurred in the first conjunct: a) approximately halfway between the offset of the verb in the first conjunct and the onset of the conjunction (PP1) and b) immediately following the conjunction (PP2). To minimize the possibility that participants would develop strategies during testing, 40 filler items were also constructed that consisted of two conjoined sentences. For filler items, probe type and probe position did not vary; the same probe word was always presented at the same probe point. This probe point occurred randomly throughout the filler sentences so that participants did not develop expectations about when the probe would be presented.

Using digital recording and editing software (Cool Edit Pro), all sentences were recorded by a female native English speaker at a normal volume and speaking rate. Probe positions were created by playing a 1kHz square-wave-onset tone on a parallel channel to that containing the speech (a channel not audible to participants). This tone was used by the computer program (RTLAB v.1.1) to initiate presentation of a visual word probe (for the naming task) and to begin response timing (which was terminated by the participant’s naming response). Timing for the placement of the

^4 The control probes for three items in this pre-test were later changed to better match the apriori naming time of the related probe. Removing these items did not change these effects.
probe points was chosen by using Cool Edit Pro software to identify the proper point in the sentence.

The 40 experimental and 40 filler items were assembled into a single pseudo-randomized list with the constraint that no experimental condition ever appeared twice in a row. Four counterbalanced versions of the list were created by systematically varying probe type and probe position across the lists for the experimental items. Although each item only appeared once in a given list, across lists each item was paired with a probe from each condition. Participants responded to probes in all four conditions, but only heard any given item once. Thus, over the course of the experiment, each participant was presented with 40 items from the four critical conditions as well as 40 filler items. In addition, five multiple-choice comprehension questions were randomly distributed throughout the list to make certain that participants were attending to the sentences, e.g. Was the sentence you just heard about...? a) a corporal b) an admiral c) a sergeant. An experimenter read the questions and answer choices at the proper points and participants responded by circling the appropriate letter on an answer sheet. Questions were always followed by a filler item in case responding to the question resulted in a disruption of processing.

Procedure

Before beginning the experiment, participants gave informed consent and completed a questionnaire about their language use and language experience. Next, the participants received instructions for the cross-modal naming task. Participants were
informed that they would be listening to sentences and that at some point during the sentence a word would appear on the screen. Participants were instructed to then say the word out loud as quickly as possible. Participants were also told that after some of the sentences the experimenter would ask them a comprehension question and that they should circle the correct answer on their answer sheet. Once participants indicated that they understood the instructions, the experiment began.

Participants sat facing a PC-compatible computer running RTLAB v.1.1 that controlled the presentation and data collection with millisecond accuracy. Each participant was exposed to one version of the four counterbalanced lists containing 40 experimental items and 40 filler items. The first ten sentences of each session were always filler sentences and served as practice trials to familiarize the subjects with the procedure. Each item was presented aurally through headphones and the appropriate visual probe was displayed on the computer screen at the appropriate point while participants’ voiced responses were recorded via microphone by the computer and a SONY digital tape recorder. Timing was initiated along with the onset of the visual probe and stopped by the onset of the response. The entire experiment lasted less than one hour.

**Experimental Design**

In this 2 X 2 multi-factorial design, the critical conditions were considered as part of a repeated-measures design with two within-subjects factors, Probe Point (PP1 or PP2) and Probe Type (Related or Control).
Results

Participants were highly accurate in their responses to comprehension questions (Mean: 94.48%; SD: 9.1%). The probe naming response times were initially screened to remove probes for which no response was recorded or for which the experimenter determined the response to be inappropriate. This initial screen resulted in the removal of 28 data points (or less than 2% of the data), roughly equally distributed across conditions. Subsequent to this screening, a repeated-measures ANOVA was conducted involving the factors Probe Point (PP1 or PP2) and Probe Type (Related or Control). Please see Table 1 for means and standard deviations.

A comparison of the response times for Related and Control probes suggested priming at PP2 (29ms advantage for related probes), but not at PP1 (1ms advantage for control probes). The statistical analysis generally confirmed this pattern by revealing a main effect of Probe Point ($F_1(1, 32) = 15.41, p = .000, F_2(1, 39) = 6.78, p = .01$), but no main effect of Probe Type ($F_1(1, 32) = 2.82, p = .10, F_2(1, 39) = 3.30, p = .08$). In addition, the interaction between Probe Point and Probe Type was numerically supported and reached significance in the subject, but not the item analysis ($F_1(1, 32)$...
= 5.02, p = .03, $F_2(1, 39) = 1.10, p = .30$). Based on a-priori hypotheses, paired one-tailed t-tests were also conducted comparing Related and Control probes at each probe point. These analyses further confirmed a significant difference between the two probe types at PP2 ($t_1(32) = 4.44, p = .000, t_2(39) = 1.92, p = .03$), but not at PP1 ($t_1(32) = .10, p = .45, t_2(39) = .47, p = .32$).

Discussion
In Experiment 1, significant priming for the verb was observed at PP2, but not at PP1. The failure to find priming at PP1 suggests that the activation of the verb related to its initial presentation decayed by, on average, approximately 1600ms later. This observation fits with previous priming experiments that suggest that the initial activation of a word decays relatively quickly after presentation (e.g. Love & Swinney, 1996). Of greater relevance, the pattern of results supports the sustained reactivation hypothesis since reactivation was temporally linked to the presentation of the conjunction (i.e. and). The findings of Experiment 1 will be discussed further along with the findings of Experiment 2 in the General Discussion.

Experiment 2- Verb activation throughout the second conjunct

Having confirmed that reactivation occurs, we now evaluate its time-course or duration. The sustained reactivation hypothesis claims that reactivated information is sustained until it can be completely integrated (i.e. at least until the end of the second
conjunct). Experiment 2 investigates this hypothesis by further inspecting the pattern of priming in the second conjunct. Specifically, two later points in the second conjunct were tested: a) 700ms before the onset of the parallel element in the second conjunct (i.e. the VP-anaphoric phrase) (PP3) and b) immediately following this phrase (i.e. at the end of the second conjunct) (PP4). Thus, PP3 and PP4 occurred on average approximately 1100ms and 2500ms after PP2. Failing to find priming at either PP3 or PP4 would constitute evidence against the sustained reactivation hypothesis, while finding priming at both PP3 and PP4 would constitute evidence for this hypothesis.

Methods

Participants

Thirty-nine participants were initially tested in Experiment 2; however, five participants were excluded from further analysis for error rates on the on-line naming task that exceeded 10% on experimental items or 15% on all items. In addition, data from three participants was lost due to computer error. Therefore, 31 participants were included in further analysis for Experiment 2 (14 male, 17 female; mean age = 19.6, range = 18 – 25, with one participant failing to provide age information).

Materials and Procedure

The materials, task, and procedure were identical to those used in Experiment 1 and have been described above\(^5\). The only modification in Experiment 2 was the

\(^5\) For clarity, the wording in two items was changed slightly.
placement of the probe points. In Experiment 2, the two probe points were chosen to investigate whether the priming of the verb is maintained throughout the second conjunct and occurred a) 700ms before the onset of the parallel element in the second conjunct (i.e. the VP-anaphoric phrase) (PP3) and b) immediately following this phrase (i.e. at the end of the second conjunct (PP4). Thus, PP3 and PP4 occurred on average approximately 1100ms and 2500ms after PP2. These points are marked with superscripts 3 and 4 in the sample sentence in (5) below.

(5) The doctor read the chart of the child with the broken arm\(^1\) during his morning rounds, and\(^2\) the insurance agent in\(^3\) the tacky suit did as well\(^4\) in order to become more familiar with the case. (Probes: reviewed (R), reserved (C))

Experimental Design

In this 2 X 2 multi-factorial design, the critical conditions were considered as part of a repeated-measures design with two within-subjects factors, Probe Point (PP3 or PP4) and Probe Type (Related or Control).

Results

Participants were highly accurate in their responses to comprehension questions (Mean: 93.55%; SD: 10.82%). The probe naming response times were initially screened to remove probes for which no response was recorded or for which
the experimenter determined the response to be inappropriate. This initial screen resulted in the removal of 17 responses (or approximately 1.4% of the data), roughly equally distributed across conditions. Subsequent to this screening, a repeated-measures ANOVA was conducted involving the factors Probe Point (PP3 or PP4) and Probe Type (Related or Control). Please see Table 1 for means and standard deviations.

A comparison of the response times for Related and Control probes suggested priming at PP3 (32ms advantage for related probes) and PP4 (31ms advantage for related probes). The statistical analysis generally confirmed this pattern by revealing a main effect of Probe Point \( (F_1(1, 30) = 13.28, p = .001, F_2(1, 39) = 6.03, p = .02) \) and a main effect of Probe Type \( (F_1(1, 30) = 11.47, p = .002, F_2(1, 39) = 9.32, p = .004) \). The interaction between Probe Point and Probe Type did not reach significance \( (F_1(1, 30) = .02, p = .89, F_2(1, 39) = .19, p = .67) \). Based on a-priori hypotheses, paired one-tailed t-tests were also conducted comparing Related and Control probes at each probe point. These analyses further confirmed a significant difference between the two probe types at PP3 \( (t_1(30) = 2.34, p = .01, t_2(39) = 2.94, p = .003) \) and PP4 \( (t_1(30) = 3.82, p = .001, t_2(39) = 2.02, p = .03) \).

**Discussion**

In Experiment 2, consistent with the sustained reactivation hypothesis, priming effects were observed at both PP3 and PP4. Considering the results from both experiments, priming for the verb from the first conjunct was observed at all points.
tested in the second conjunct (i.e. PP2, PP3, and PP4), indicating that the activation of the verb was sustained at least approximately 2500ms after its initial reactivation following the conjunction. In addition, the numerical magnitude of the priming effect was virtually identical across the three probe points where it was detected (PP2, PP3, and PP4). The implication of these findings will be discussed in the General Discussion.

General Discussion

Previous studies have demonstrated that the processing of the second conjunct of a conjoined phrase is facilitated if it is parallel to the first conjunct (Apel et al., 2007; Frazier et al., 1984, 2000; Knoeferle, 2007; Knoeferle & Crocker, accepted). The current study investigated a specific hypothesis about the processes underlying these parallelism effects, the sustained reactivation hypothesis, by examining the activation of a parallel element (i.e. a verb) throughout two “and”-coordinated sentences. We examined evidence of activation (i.e. priming as indicated by faster response times for semantically related vs. unrelated words) at four points throughout the sentences: 1) approximately 1600ms after the presentation of the verb in the first conjunct (PP1), 2) immediately following the conjunction (PP2), 3) approximately 1100ms after the conjunction (PP3), 4) at the end of the second conjunct (i.e. approximately 2500ms after the conjunction) (PP4).

In the context of previous research, the results at PP1 reveal the time-course of the activation of the verb in the first conjunct. It is well-known that processing a word
immediately activates its lexical representation. For a verb, this representation includes meaning information (de Goede, 2006; Callahan et al., in prep) as well as associated semantic and syntactic information (Altmann & Kamide, 1999; Boland et al., 1990; Ferretti et al., 2001; Friederici & Frisch, 2000; McRae et al., 1997; McRae et al., 1998; Shapiro et al., 1987, 1989). Based on previous research suggesting that activation related to the presentation of a word decays relatively quickly (e.g. Love & Swinney, 1996; Nicol & Swinney, 1989), we predicted that verb activation would not be detected at PP1. Confirming this prediction, no priming was observed at PP1, suggesting that the activation of the verb related to the initial presentation decayed (at least by approximately 1600ms later). Despite being consistent with the previous research on noun processing in English sentences, this finding is inconsistent with a study that investigated verb processing in Dutch sentences and found sustained activation throughout the entire clause related to the verb (de Goede et al., in press); however, the Dutch study differed from previous studies and the current study in a number of ways. In particular, we suspect that the fact that in the Dutch experimental sentences the verb was displaced from the canonical position via head movement might be relevant to the observed pattern of activation. Since head movement is a relatively unexplored psycholinguistic phenomenon, further psycholinguistic research on both head movement and, more generally, the activation of verbs in sentences will be useful in clarifying these issues.

In the context of previous research, the results at PP2, PP3, and PP4 further clarify the activation of the verb in the second conjunct. The sustained reactivation
hypothesis claims that encountering a cue that implies parallelism (e.g. the conjunction, *and*) elicits the reactivation of previously processed relevant material. Based on this hypothesis, we predicted a priming effect at PP2 immediately following the conjunction. Experiment 1 confirmed this prediction, revealing a significant 29ms advantage for related probes at PP2. Experiment 2 then investigated the activation of the verb at later points in the second conjunct. The sustained reactivation hypothesis predicted that the reactivation of the verb would be maintained throughout the second conjunct (i.e. priming effects at PP3 and PP4). Indeed, the results of Experiment 2 revealed a significant 32ms and 31ms advantage for related probes at PP3 and PP4 respectively.

It is worth noting that the numerical magnitude of the priming effect in the current study was similar across the three points tested in the second conjunct. Residual activation due to incomplete decay would be associated with a reduction in the magnitude of the priming effect (Shapiro & Hestvik, 1995). Thus, the nearly identical priming effects at PP2, PP3, and PP4 suggest sustained activation due to active maintenance rather than residual activation due to incomplete decay. Active maintenance processes that are linked to short-term memory as reflected electrophysiologically by a sustained left anterior negativity (c.f. King & Kutas, 1995) have been observed previously in sentence processing when antecedent information cannot be immediately integrated (e.g. as in relations involving displaced arguments). Since the cue to parallelism preceded the processing of the second conjunct in the
current study, information from the first conjunct could not be integrated immediately following reactivation.

The results of Experiments 1 and 2 suggest the following conclusions about the activation pattern of a verb in “and”-coordinated sentences: 1) decay occurs relatively quickly following initial activation, 2) the coordinating conjunction elicits reactivation, and 3) this reactivation is sustained until at least the end of the second conjunct. As discussed previously, a successful psycholinguistic theory of parallelism must describe not only the nature and time-course of the underlying processes but also the relevant mechanisms and underlying representations. The sustained reactivation hypothesis suggests that two processes independently demonstrated to be involved in sentence processing (i.e. reactivation and maintenance) play a key role in generating parallelism effects. Specifically, this hypothesis claims that, in coordinate structures, a cue that implies parallelism (e.g. and, see Knoeferle, 2007) elicits the reactivation of previously processed material from the first conjunct so that this information can be integrated with information from the second conjunct. Following reactivation, the activation of the material is maintained at least until integration can be completed (i.e. until the end of the second conjunct). Although the sustained reactivation hypothesis points to the conjunction, and, as the relevant cue to parallelism in the current study (and in many previous studies), the notion of cue is intended to be broad enough to encompass other linguistic (or potentially even non-linguistic) cues to parallelism.

Having found support for a hypothesis regarding the nature and time-course of the processes underlying parallelism effects, we now delineate some important
remaining questions about a theory of parallelism and suggest some tentative extensions of the current hypothesis. One remaining question concerns the mechanism that underlies the reactivation process. A series of studies by McElree and colleagues (McElree, 2000; McElree et al., 2003; Martin & McElree, 2008) have suggested that reactivation in several types of linguistic dependencies (e.g. long-distance anaphoric relationships, subject-verb relationships) is accomplished via a cue-driven, direct access mechanism: we suggest that such a mechanism may also underlie the reactivation process in parallelism with the conjunction serving as the relevant cue.6 Another question relates to the types of information that are reactivated following the conjunction (e.g. phonological, syntactic, semantic, discourse, etc.). Since both syntactic and semantic parallelism effects have been observed, the reactivated representation must contain at the very least both of these types of information; however, since parallelism can occur at a number of levels ranging from phonology to discourse, we go a step further and suggest that the reactivated representation may contain information at each of these levels.

One final question is whether all parts of a complex first conjunct are reactivated. Based on the results of the current study, we can confidently state that the verb is activated in coordinate sentences; however, as stated previously, we chose to test the activation of the verb because it plays a particularly important role in the sentence. The question remains, then, whether less critical parts of the conjunct are

---

6 As Martin and McElree (2008) note, this mechanism is generally consistent with copy α, the structure sharing mechanism proposed by Frazier and Clifton (2001); however the cue-driven direct access mechanism has the advantage of not being restricted to certain types of constructions.
reactivated. Interestingly, there have been several priming experiments involving similar parallel constructions (Shapiro & Hestvik, 1995; Shapiro et al., 2003) that have investigated the activation of the subject or object from the first conjunct. The authors interpreted the results as suggesting that the subject and object of the VP in the first conjunct were not activated at a point downstream from the conjunction in the second conjunct but were reactivated following a VP-anaphoric phrase at the end of the second conjunct. Despite their interpretation, we note that the effect for the object at the earlier point did reach significance at $p = .05$ ($p = .01$ was the required level when corrected for multiple comparisons), so it is possible that the object was indeed activated at this point. Regardless, although these data do pose a problem for a strong version of the sustained reactivation hypothesis that claims that all potentially parallel material is reactivated following the conjunction and maintained until it can be fully integrated, they are compatible with a more incremental version of this hypothesis whereby individual constituents decay following integration with their parallel counterparts or a version where only certain parts of the first conjunct are reactivated.

Although much progress has been made towards a theory of parallelism, future research will be critical in refining and solidifying the tentative conclusions described above. In particular, investigations of parallelism that evaluate the role of specific conjunctions as well as other linguistic cues will allow the scope of parallelism to be defined more clearly. Similarly, investigations of the activation patterns of different components of complex coordinate structures (e.g. the subject, the object, adjunct phrases, etc.) and different types of information (e.g. phonological, syntactic,
semantic, discourse, etc.) will allow more confident statements regarding the characteristics of the reactivated representation.

This chapter, in full, has been submitted for publication as it may appear in *Journal of Psycholinguistic Research*, Callahan, S.M., Shapiro, L., & Love, T. The dissertation author was the primary investigator and author of this paper.
References


Knoeferle, P. (2007). Eye tracking parallelism effects in coordinate constructions: The effects of “and”, “but” and “while”. Poster presented at the 20th Annual CUNY, La Jolla, USA.


Appendix

1. The admiral lectured on submarines last Saturday in a speech to his men, and the President of the United States did the same a week later in Congress.
   R: spoke   C: added

2. Parisians drink wine which is native to the country, and farmers throughout the Provinces do too because of the high quality and cheap prices available.
   R: swallow   C: skipped

3. The artist improved the design until it was ready to sell, and the businessman in the nearby town did the same with the new line of clothes.
   R: perfected   C: suspected

4. The secretary chatted on the telephone at the desk in the front office all day long, and the Vice President in the adjoining office did too because his assistant was on vacation.
   R: talked   C: topped

5. Athletes train muscle pattern responses necessary for maximum performance, and musicians in symphonic ensembles do the same while rehearsing the more difficult passages of each piece.
   R: practiced   C: described

6. The journalist wrote on Sunday despite the fact that it was here day off, and the researcher in the Neurosciences Department did as well for an article being presented at a conference.
   R: typed   C: named

7. The equestrian trained the quarter horse that was a strong contestant in the race, and the cowboy from a Texas ranch did as well in order to keep the horse in good shape.
   R: taught   C: joined

8. The girl looked for a gift for the birthday party, and the receptionist at the large firm did the same because the boss was always so busy.
   R: searched   C: reserved

9. The mime performed for the tourists on the pier, and the punk rocker with the rainbow hair did too every Sunday afternoon.
   R: acted   C: raced

10. The hunter shot a rattlesnake in the lot around the corner, and the animal control officer in the neighborhood did too providing a safe environment for after school play.
    R: killed   C: walked
11. The doctor read the chart of the child with the broken arm during his morning rounds, and the insurance agent in the tacky suit did as well in order to become more familiar with the case.
   R: reviewed  C: reserved

12. The old man observed the children in the park on sunny afternoons, and the office workers from the corporate annex did the same during their lunch breaks.
   R: watched  C: figured

13. The priest skipped breakfast Tuesday morning because of the time change, and the electrician from up the street did as well because he had to run errands in the morning.
   R: missed  C: topped

14. The farmer fed the animals before going to bed, and the veterinarian who runs the emergency clinic did too before leaving work for the evening.
   R: starved  C: groaned

15. The jeweler examined the rock to see if it was a precious stone, and the geologist who was on a trip to Africa did the same while trying to identify its origin.
   R: studied  C: decided

16. The cyclist crossed the street in the middle of rush hour traffic, and the cabby from New York City did the same on his way to the airport.
   R: traversed  C: elongated

17. The shaman explored the desert terrain in search of particular roots, and the hikers from the East Coast did as well on their vacation to Joshua Tree.
   R: pioneered  C: compacted

18. The chemist discovered the ingredients that were necessary for his experiment, and the witch doctor from a Haitian village did the same because he was intent upon casting his spell.
   R: found  C: faced

19. The Queen of England toured the countryside in order to promote good will among the farmers, and the professor from the archaeology department did as well in order to learn about the area.
   R: traveled  C: suspended
20. The engineer entered the field as a troubleshooter for his company, and the social worker from Child Welfare Services did as well to assess the progress made in her cases.

   R: exited                        C: dialed

21. The dictator put people in jail by the thousands in order to suppress insurrection, and the judge on the State judiciary board did too after 50 years on the bench.

   R: confined                      C: deserted

22. The postman cruised the neighborhood to deliver the mail to each house, and the TV repairman from the cable company did too while looking for the correct addresses.

   R: drove                         C: meant

23. The chef baked a cake for the annual competition, and the housewife in her brand new kitchen did the same for the birthday party.

   R: prepared                      C: accepted

24. The critic hated the film shown last Saturday because of its flagrant inconsistencies, and the steelworker from Detroit did too because of the excessive violence involved.

   R: loved                         C: heard

25. The glassblower created beautiful objects for the Ancient Artisans symposium, and the sculptor from Manhattan did as well, experimenting with modern and innovative techniques.

   R: crafted                       C: gripped

26. The forest ranger looked at the compass to orient himself without sunlight, and the navigator of the Princess cruiseliner did the same to ensure he was not off course.

   R: viewed                        C: shared

27. The miser ate the burger slowly to savor it as long as possible, and the school children at McDonald’s did too while competing to see who chewed each bite the longest.

   R: chewed                        C: matted

28. The paramedics looked at the scene of the accident to see if the victims were still alive, and the bystanders from across the street did the same because they were hoping for a good view.

   R: scanned                      C: clipped
29. The librarian heard the announcement over the intercom system, and the jockey from the local stables did too when it was time for him to join the line-up for the upcoming race.
   R: listened  C: intended

30. The plumber purchased the Big Gulp at 7-11 on his way home from work, and the comedian in the VW bug did too because she was thirsty for a cold drink.
   R: bought  C: jumped

31. The Catholic nun called the number to City Hall to schedule a room for the bible study, and the alcoholic from Kansas did the same so he could verify an appointment scheduled for next week.
   R: phoned  C: coated

32. The botanist placed the plant in the sunlight after finding just the right spot, and the decorator from the Prospect Avenue firm did too because the location by the window was much nicer.
   R: put  C: let

33. The fireman took off the heavy boots that were covered in mud and ash, and the mountain climber from the French Alps did as well when he returned to camp.
   R: removed  C: charged

34. The ophthalmologist gave the jewelry to the museum in an attempt to keep it from a despised sibling, and the collector of Post-WW1 artifacts did the same knowing that many would appreciate their beauty.
   R: offered  C: crowded

35. The graduate student set the books on a shelf to deal with them later, and the accountant from the office did too so he would know where to begin in the morning.
   R: placed  C: ranted

36. The rock star promised the proceeds to charity when touring through LA last year, and the millionaire from the marketing firm did too when trying to build a better public image.
   R: pledged  C: spanned

37. The dogcatcher gave a puppy to the child in the dirty dungeon, and the grandfather from a ranch in Montana did the same during a visit to the family.
   R: returned  C: mattered
38. The principal addressed the problems of the students during the assembly period in the auditorium, and the police officer from the Public Relations Department did the same for a discussion about drug abuse.
   R: discussed   C: suggested

39. The housekeeper loaned the recipe to a neighbor for a dinner party the following evening, and the actress in need of publicity did too after making a big deal about how fantastic the food was.
   R: borrowed   C: glimpsed

40. The astronaut gave the clothing to the space museum for a re-opening next week, and the designer for the Globe Theater did the same with some of the older priceless costumes.
   R: donated    C: thrived
Table 1. Means and Standard Deviations in terms of Response Time (ms) for Experiments 1 and 2.

<table>
<thead>
<tr>
<th>Probe Type</th>
<th>Probe Point</th>
<th>1 (n = 31)</th>
<th>2 (n = 31)</th>
<th>3 (n = 33)</th>
<th>4 (n = 33)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td></td>
<td>632</td>
<td>619</td>
<td>647</td>
<td>624</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(74)</td>
<td>(78)</td>
<td>(126)</td>
<td>(117)</td>
</tr>
<tr>
<td>Related</td>
<td></td>
<td>633</td>
<td>590</td>
<td>614</td>
<td>594</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(107)</td>
<td>(83)</td>
<td>(103)</td>
<td>(110)</td>
</tr>
<tr>
<td>Difference</td>
<td>-1</td>
<td>29*</td>
<td>32*</td>
<td>31*</td>
<td></td>
</tr>
</tbody>
</table>

Note. Difference = Control – Related. A positive difference indicates faster response times for related than control probes (i.e. priming) and a negative difference indicates slower response times for related than control probes. The difference does not always equal the subtraction of the control and related because of rounding. Paired t-tests based on a-priori hypotheses revealed that differences marked with a star (*) were significant by participants and items at $p < .05$ while those marked “ns” were not ($p > .05$). Probe points 1 and 2 were tested in Experiment 1 while Probe points 3 and 4 were tested in Experiment 2.
Chapter 6

General Discussion
The studies contained in this dissertation represent a multi-methodological investigation of the processing and interpretation of coordinate sentences involving verb phrase (VP) anaphors. These experiments exploit the advantages of a multi-methodological approach by obtaining converging evidence using a variety of tasks and dependent measures as well as both written and spoken stimuli. Furthermore, for each individual experiment, the most useful methodology was chosen based on the specific questions of interest in each study and the type of information most relevant to the hypotheses under investigation. First, the naturalness rating experiment described in Chapter 2 and the interpretative preference experiment described in Chapter 4 provided information about conscious preferences in relation to the two VP-anaphoric forms. Second, the eye-tracking experiment in Chapter 2 and the ERP experiment in Chapter 3 provided information about the unconscious real-time processing of these forms. Employing both eye-tracking and electrophysiological methodologies also served to provide complementary neurophysiological evidence about the direction of the processing effects as well as their timecourse and distribution. Finally, since coordinate sentences are one very natural environment for VP anaphors, the associated cross-modal lexical priming experiment in Chapter 5 revealed information about the processes and underlying representations involved in parallelism, an important related factor in the processing of coordinate sentences.
Effects of Anaphoric Form

In terms of conscious preference, the naturalness rating experiment in Chapter 2 (and the naturalness rating pretest in Chapter 3) revealed that, for contexts involving a single highly-accessible antecedent, sentences containing an overt VP-anaphoric phrase were rated as less natural than sentences containing a null VP-anaphoric phrase. Despite this observed preference for the null VP-anaphoric phrase, the results of the interpretative preference experiment in Chapter 4 revealed that comprehension accuracy was still extremely high for both forms. Furthermore, the two forms were interpreted similarly both in contexts that were fully ambiguous (i.e. both the strict and sloppy interpretations were available) as well as contexts where pragmatic information biased interpretation (i.e. the strict interpretation was blocked by characteristics of the VP). Thus, although comprehenders prefer the null VP-anaphoric phrase in these contexts, their performance on off-line measures probing final conscious interpretation is similar for both forms.

Although off-line measures are important tools for establishing conscious preferences, these measures cannot reveal subtle variation in processing as it occurs throughout the sentence. Thus, to investigate processing variation related to anaphoric form, it is necessary to employ on-line measures that are sensitive to real-time processing. Along these lines, two experiments were conducting using methodologies with excellent temporal resolution: eye-tracking and electrophysiological recordings. The experiment in Chapter 2 recorded participants’ eye-movements as they read sentences involving the two anaphoric forms. Longer first-pass and total reading times
were observed for the anaphor region when it contained an overt as compared to a null VP-anaphoric phrase; however, this effect was reversed in the post-anaphor region where marginally longer total times were observed for the null as compared the overt VP-anaphoric condition. Similarly, the ERP experiment in Chapter 3 recorded electrical activity in the brain as participants listened to sentences involving the two anaphoric forms. Two effects related to anaphoric form were observed: an early centro-posterior negativity that peaked approximately 100ms after the onset of the overt anaphor (i.e. *it*) in the overt VP-anaphoric condition and an anterior positivity between 300 and 500ms after the onset of the word following the anaphor site (i.e. *too*) for the null VP-anaphoric condition.

When taken in the context of the off-line preference studies, the results of the eye-tracking and the ERP experiments suggest that an unexpected continuation (i.e. the overt VP anaphor) initially elicits additional processing. This fits with other ERP studies that suggested that expectation plays a critical role in the processing of constructions involving null anaphors (Kaan, Wijnen, & Swaab, 2004; Streb, Hennighausen, & Rösler, 2004; Walenski, 2002). In addition, the results of the eye-tracking and ERP experiments suggest that processing resources are deployed at different times during sentences involving null and overt VP-anaphoric phrases. Essentially, these resources are deployed as soon as the anaphoric relationship is detected; however, this occurs later in the case of the null VP-anaphoric phrase because the anaphor site cannot be detected until the next word (i.e. *too*). This variation in the timecourse of resource allocation is reflected in the reversal pattern
observed in both the eye-tracking and ERP experiments. Initially, processing is increased for the overt as compared to the null VP anaphor, as reflected by the increased reading times in the anaphor region and an early centro-posterior negativity (i.e. possibly an N100) following the overt anaphor, *it*. Later, processing is increased for the null as compared to the overt VP anaphor, as reflected by increased reading times in the post-anaphor region and an anterior positivity following the onset of the word after the anaphor (i.e. a delay in the onset of the contingent negative variation (CNV)).

Overall, then, the results of the naturalness rating, eye-tracking, and ERP experiments described in Chapters 2, 3, and 4 suggest a sensitivity to VP-anaphoric form that is observable both in off-line measures that reflect conscious preference as well as in on-line measures that reflect unconscious processing. The complementary on-line and off-line results in these studies support the following conclusions: a) in contexts where multiple anaphoric forms are grammatical and there is only one highly-accessible antecedent, comprehenders have an expectation for the least informative anaphoric form possible, b) violations of this expectation elicit additional processing at the anaphor site, c) variation in when the anaphoric relationship can be detected results in variation in processing.

The effects observed throughout this dissertation do not support the predictions of differentiation theories (e.g. Frazier & Clifton, 2001; Sag & Hankamer, 1984) because, overall, there was no evidence of qualitatively different processes for null and overt VP anaphors. In particular, the waveforms for the null and overt conditions
in the ERP study were very similar, suggesting that the processing of the two forms likely involves the same underlying processing mechanisms, at least in contexts where both are grammatical. Along these lines, the results are generally more consistent with unification theories (e.g. Ariel, 1990, 1991; Givón, 1983; Murphy, 1985) since processing variation due to anaphoric form was related to preferences for one anaphoric form over another in a given context.

Parallelism

The experiments in Chapters 2, 3, and 4 confirmed that variation in anaphoric form was associated with variation in the processing of sentences involving VP-anaphoric phrases. In natural language, the specific VP-anaphoric phrases tested (i.e. *did too* vs. *did it too*) are typically used to express parallelism because the word, *too*, implies parallel events. Along these lines, these phrases are often used in *and*-conjoined sentences where the conjunction, *and*, also implies parallelism. Accordingly, the study in Chapter 5 investigated the role of this related factor (i.e. parallelism) in the processing of coordinate sentences such as those tested throughout this dissertation. Inspired by the natural linking between parallelism and anaphora, this study explored the possibility that similar processes and representations might be involved in the two phenomena. In particular, two processes well-known to be operative in the processing of anaphoric constructions (i.e. reactivation and maintenance) were proposed to also be involved in the processing of parallel constructions.
Along these lines, Chapter 5 evaluated a specific hypothesis about parallelism (i.e. the sustained reactivation hypothesis) that claims that, following a conjunction that implies parallelism (e.g. *and*), material from the first conjunct is reactivated and sustained until it can be integrated with material from the second conjunct. The cross-modal lexical priming experiments in Chapter 5 tested this hypothesis by presenting coordinate sentences and evaluating the activation of the verb from the first conjunct at four points throughout the two conjuncts. The results revealed that the activation of the verb was no longer detectable approximately 1600ms after its initial presentation in the first conjunct; however, then, reactivation occurred at a point immediately following the conjunction. Moreover, this reactivation was sustained at least until the end of the second conjunct (i.e. approximately 2500ms after the conjunction). This pattern of results supports the sustained reactivation hypothesis and fits with a previous study that indicated an important role for the conjunction in eliciting parallelism effects (Knoeferle, 2007). Thus, the sustained reactivation hypothesis, as supported by the results of the study in Chapter 5, suggests that so-called parallelism effects (i.e. facilitated processing for the second of two parallel conjuncts) can be linked to the reactivation of material from the first conjunct during the processing of the second conjunct.

**Conclusion**

In terms of the relationship between parallelism and anaphora, the findings in Chapter 5 suggest that one reason that less informative anaphors are typically used to
refer to parallel antecedents is that the reactivation of parallel material renders this material highly-accessible at the anaphor site. Figure 1 depicts this reactivation within a model of the real-time processing of coordinate sentences involving VP anaphors. Essentially, this model views language comprehension as a process of educated guessing. Prior experience is used to generate probabilistic expectations about upcoming material. In the case of anaphora, prior experience incorporates statistical information about the associations between different anaphoric forms and different types of antecedents (as captured by the generalizations expressed in unification theories). As linguistic input is processed, these probabilistic expectations are refined. Thus, in the case of the processing of coordinate sentences, prior experience and the processing of a conjunction that implies parallelism (e.g. *and*) lead to an expectation for parallel material and less informative anaphors. As compared to material that is consistent with expectations, material that violates expectations elicits additional processing (as reflected by, for example, increased reading time). Thus, parallel material and less informative anaphoric forms are processed more easily than non-parallel material and more informative anaphoric forms.

The model proposed here is neurophysiologically-plausible as the brain is known to be sensitive to probabilistic expectation (e.g. the N400 reflects the context-specific frequency of a particular word, see Kutas, Federmeier, Coulson, King, & Münte, 2000). In addition, this model fits with other theories that posit a critical role for probabilistic expectations in sentence processing (e.g. surprisal theories, see Hale, 2003; Levy, 2008). Moreover, although the schematic in Figure 1 illustrates the
processing of coordinate sentences involving VP anaphors, the model as construed in a
broader sense makes straightforward predictions about the processing of other
coordinate constructions and other types of anaphors. More generally, this model
underscores the importance of future research on the delicate interplay between prior
experience and immediate context in generating expectations that have observable
effects on processing.

Overall, the work contained in this dissertation makes a number of important
contributions to the field of language processing. First, since VP anaphors are
relatively unexplored in psycholinguistic and neurolinguistic research, these studies
extend previous research on anaphoric processing into a new domain. Second, by
comparing null and overt VP anaphors in contexts where both are grammatical, these
studies contribute to the growing psycholinguistic and neurolinguistic literature on the
processing of null anaphoric constructions and elucidate the effects of anaphoric form
on real-time processing. Finally, by suggesting a novel hypothesis about the processes
underlying parallelism effects and a model of the processing of coordinate sentences
involving VP anaphors, this work links research on parallelism with research on
anaphora. Thus, the conclusions presented here emphasize the importance of exploring
connections between related phenomena and suggest an important new set of
questions for future research. In addition, this dissertation highlights the utility of a
multi-methodological approach. In the end, successful wide-scope theories of
language must account for behavior in terms of conscious usage and interpretation as
well as unconscious real-time processing (including the underlying neural bases of
such processing). Developing and refining such theories will require all the methodological tools at our disposal.
References


Figure 1. A model of the processing of coordinate sentences involving verb phrase anaphors.