UNIVERSITY OF CALIFORNIA, SAN DIEGO

Pronouns, Prosody, and the Discourse Anaphora Weighting Approach

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

in
Psychology

by
Jennifer Elaine Balogh

Committee in charge:

Professor David Swinney, Chair
Professor Jeffery Elman
Professor Victor Ferreira
Professor Andrew Kehler
Professor Timothy Rickard

2003
This dissertation of Jennifer Elaine Balogh is approved, and it is acceptable in quality and form for publication on microfilm:

Chair

University of California, San Diego

2003
DEDICATION

In memory of World War II fighter pilot Lieutenant Colonel Jay Marts who always encouraged me to fly.
TABLE OF CONTENTS

Signature Page ................................................................................................................................iii
Dedication........................................................................................................................................ iv
Table of Contents .................................................................................................................................. v
List of Figures ........................................................................................................................................ vii
List of Tables .......................................................................................................................................... viii
Acknowledgements .............................................................................................................................. ix
Curriculum Vitae .................................................................................................................................. xi
Abstract of Dissertation ..................................................................................................................... xv

I. Introduction and Background ..............................................................................................................1
   A. Theories of Pronoun Referent Resolution..................................................................................3
   B. Theory of Intonation ..................................................................................................................21
   C. Pronouns and Prosody ..............................................................................................................23

II. On-line Methodologies and Findings .................................................................................................32
   A. Methods Used in On-line Psycholinguistic Experiments ......................................................34
   B. On-line Processing of Pronouns .............................................................................................41
   C. Research on Prosody and Pronoun Processing .......................................................................73

III. Pitch Accent as Focus .......................................................................................................................77
   A. Pitch Accent and Focus ..............................................................................................................77
   B. Experiment 1: Focus Stress in Active Structures ...................................................................79
   C. Experiment 2: Focus Stress in Passive Structures ....................................................................88
   D. Experiment 3: Focus Stress in Conjoined Structures ............................................................95
   E. Experiment 4: Focus Stress with Source-Goal Theta Roles ................................................100
   F. Experiment 5: Focus Stress with Goal-Source Theta Roles .................................................105

IV. Pitch Accent as Contrastive Emphasis .............................................................................................113
   A. Experiment 6: Off-line Effects of Contrastive Stress ..........................................................118
   B. Experiment 7: On-line Effects of 2nd NP at Baseline and Pronoun ..................................123
   C. Experiment 8: On-line Effects of 1st NP at Baseline and Pronoun ..................................134
D. Experiment 9: On-line Effects of 2nd NP at 400 ms and 800 ms ........142
E. Experiment 10: On-line Effects of 1st NP at 400 ms and 800 ms .......146

V. A New Model – The Discourse Anaphora Weighting Approach ..............155
   A. Discourse Anaphora Weighting Approach (DAWA) .....................158
   B. Examples ..................................................................................................195
   C. Conclusion .................................................................................................216

Appendix A: Materials for Experiments in Chapter 3 ..........................219
Appendix B: Materials for Experiments in Chapter 4 ............................225
References .........................................................................................................232
LIST OF FIGURES

Figure 1: Example of network in Construction-Integration Model ..................71
Figure 2: Pronoun selection with focus stress in active constructions. ..............85
Figure 3: Pronoun selection with focus stress in passive constructions .............92
Figure 4: Pronoun selection with focus stress in conjoined phrases ..................98
Figure 5: Pronoun selection with focus stress and Source/Goal verbs ...............104
Figure 6: Pronoun selection with focus stress and Goal/Source verbs ..............109
Figure 7: Mean reaction times of probes related to the second NP and control probes when presented 800 ms before the pronoun and at the offset of the pronoun. ..........131
Figure 8: Mean reaction times of probes related to the first NP and control probes when presented 800 ms before the pronoun and at the offset of the pronoun. ..........137
Figure 9: Mean reaction times of probes related to the second NP and control probes when presented 400 ms and 800 ms downstream .....................................................144
Figure 10: Mean reaction times of probes related to the first NP and control probes when presented 400 ms and 800 ms downstream .....................................................149
Figure 11: Stages of the Discourse Anaphora Weighting Approach ..................161
Figure 12: Illustration of Encoding ....................................................................162
Figure 13: Illustration of mini-network in LTM copied into the CS ....................163
Figure 14: Illustration of identification mechanisms in Architecture 1 ..............173
Figure 15: Illustration of identification mechanisms in Architecture 2 ..............175
Figure 16: Illustration of identification mechanisms in Architecture 3 ..............177
Figure 17: Process of assessing an individual candidate’s activation level from Mechanisms 1 and 2. .................................................................184
LIST OF TABLES

Table 1: Mean Pitch, Duration and Intensity Measurements for Each Type of NP in Experiment 1..............................................................................................................................82

Table 2: Mean Pitch, Duration and Intensity Measurements for Each Type of NP in Experiment 2..............................................................................................................................89

Table 3: Mean Pitch, Duration and Intensity Measurements for Each Type of NP in Experiment 3..............................................................................................................................96

Table 4: Mean Pitch, Duration and Intensity Measurements for Each Type of NP in Experiment 4..............................................................................................................................102

Table 5: Mean Pitch, Duration and Intensity Measurements for Each Type of NP in Experiment 5..............................................................................................................................107

Table 6: How Theta Roles Influenced Pronoun Referent Resolution. .........................111

Table 7: Mean Pitch and Duration Measurements of the Pronoun in Experiment 6. 121

Table 8: Theories, Models, and Frameworks of Anaphora. .................................155

Table 9: Example of Lexical Feature Vector Information Tagged for Nouns.............162

Table 10: Units of Time Needed to Process Sentences for Architecture 1..............174

Table 11: Units of Time Needed to Process Sentences for Architecture 2.............176

Table 12: Units of Time Needed to Process Sentences for Architecture 3.............178
ACKNOWLEDGEMENTS

This dissertation would not have been possible without the help and support of many generous people in my life. First and foremost, I want to thank Debajit Ghosh who has been my greatest inspiration and source of encouragement. Whenever I ran into an obstacle, he would be there, debugging my scripts, piloting my experiments and even soldering wires to help me. I thank my loving family, Anthony Balog, Elaine and Dave Clow, Renee, Tim, Garrison, and Abigail Lizotte, for being there for me always. With much respect, I thank my advisor Dave Swinney, especially for his support of my decision to take a nonconventional path, and my committee members and professors, Andy Kehler, Tim Rickard, Vic Ferreira, Jeff Elman, Edgar Zurif and Yosef Grodzinsky, who have been the foundation of my education. I would also like to thank Michael Cohen who has always listened to my research with interest and who never runs out of stories that show he understands.

There are numerous others who have helped me with my experiments and/or who have offered morale support including Carrie Armel, Sondra Ahlén, Martha Senturia, Cathy Pearl, Amy Ulug, Elizabeth Strand, Su-Lin Wu, Trace Wax, Nino Walker and family, Karen Kaushansky, Charles Jankowski, Dan Burnett, Tony Sheeder, Margaret Urban, Michael Shafer, Jeff Shrager, Rajeev Agarwal, Shamitha Somashekar, Ann Thyme-Gobble, Dawn Morales, Nicole LeDuc, Mike Prince, Rebecca Nowlin Green, Carol Bleyle, Brian Strope, Lisa Falkson, Elizabeth Balog, Liz Wiesend and family, John Balog and family, Theresa Griffin and family, Helen and Jay Marts, James and Juanita Harris, Lea Hald, Tracy Love, Cassie Taylor, Steven
Williams, Judee Humberg, Carole Betzner and family, Anne Betzner, Jared Strawderman, James Riseman, Paul Constantinides, Milan Young, Jeff Bergan, Dave O’Connor, Christopher Nicols, Torsten Zeppenfeld, Will Lavery, Riley Willcox, Linda Walden, Jonathan Luk, Matthew Whipple, Eric Tober, Rod Golden, Russ Nixon, Helen Hsiu, Edwin Li, Eliot Rubinov, Hank Liao, James Giangola, Jeanne Hsu, Kevin Fong, Leo Haasbroek, Yasunori Morishima, Christine Nakatani, and Mary Constance Parks.

Some of the research reported in this dissertation was supported by NIH grant DC02984.
CURRICULUM VITAE

Jennifer E. Balogh

EDUCATION

Ph.D.  University of California, San Diego, Psychology, 2003 (ABD)
M.A.  University of California, San Diego, Psychology, 1998
B.A.  Brandeis University, English and American Literature (Highest Honors), 1994
Cognitive Science and Linguistics minor

PROFESSIONAL EXPERIENCE

Communications, Menlo Park, CA
2000 – 2001  Assistant Instructor, Spoken Language Understanding Systems, Dr.
Michael Cohen, Instructor. Stanford University, Palo Alto, CA
1996 – 2000  Teaching Assistant, Department of Psychology, University of
California, San Diego
1994 – 1996  Research Assistant, Aphasia Research Center, VA Hospital, Boston,
MA

HONORS AND AWARDS

Summa Cum Laude, Phi Beta Kappa, NSF Graduate Fellowship Honorable Mention,
Andrew Grossbardt Memorial Poetry Prize, Pushcart Prize Nominee, John P. Schaffer
Scholarship, Presidential Fitness Award.

GRANTS AND FELLOWSHIPS

Center for Research in Language Training Fellowship, Fall 1997 - Spring 1998.
Graduate Student Research Support Grant, Fall 1996.

INVITED LECTURES


PUBLICATIONS

Books

Refereed Journal Articles


Refereed Conference Papers and Posters


Book Chapters


PATENTS

ABSTRACT OF DISSERTATION

Pronouns, Prosody, and the Discourse Anaphora Weighting Approach

by

Jennifer Elaine Balogh
Doctor of Philosophy in Psychology

University of California, San Diego, 2003

Professor David Swinney, Chair

The focus of this dissertation is how prosody, in particular pitch accent, affects pronoun referent resolution. First, some influential theories of pronoun and reflexive referent resolution are reviewed, followed by separate work on the representation and meaning of prosody. Some proposals integrating the two are presented, but what becomes apparent is that there is a large collection of data for which these theories do not account. Further, the paucity of research devoted to prosody’s real-time influence on pronouns underscores the need for more research and a better theory recognizing prosody’s role.
Two sets of experiments are presented that investigate pitch accent’s effect on pronoun referent resolution. The first set explores the role of focus stress when placed on one of two potential referents of a pronoun. The results show that focus stress significantly impacts pronoun interpretation, but in subtle ways. Pitch accent serves as a secondary influence that interacts with other, more powerful forces including preference for the subject and theta roles. The first two experiments provide data to refine a hypothesis that predicts a specific pattern of results. The following three experiments validate the hypothesis using different sentence structures.

The second set of experiments deals with pitch accent used for contrastive stress. These experiments use a Cross-Modal Lexical Priming paradigm with three-sentence discourses to map out the time course of prosody and its impact on pronoun referent assignment. Based on data from four different probe points, the experimental results show that pitch accent as contrastive stress is a fast-acting cue that constrains potential referent activation at an early stage of processing. The results also suggest that contrastive stress triggers a second pass interpretation several hundred milliseconds after the pronoun.

Given these results, I propose a unified theory of the real-time processing of discourse anaphora in English, called the Discourse Anaphora Weighting Approach (DAWA). The DAWA takes into account the data up to this point and promotes the role of prosody as an important factor in coreference.
INTRODUCTION AND BACKGROUND

Pronouns are one of the most frequently used words in the English language, but the problem of how native speakers resolve who or what a pronoun refers to is a difficult problem for science. Even though much research has been devoted to the topic in linguistics, psychology, and computer science, there are often inconsistent viewpoints. Native speakers use a vast amount of information to resolve the pronoun’s referent, such as the type of pronoun used, consistency of gender, number and animacy, syntax, discourse context, prosody, and general knowledge of the world (to name a few). Consider the following:

(1) a. Rhonda hired Lisa immediately.

b. She was very impressed by her programming skills.

In this example, even though both Rhonda and Lisa are syntactically licensed referents of the pronouns she and her in 1b, English speakers will associate she with Rhonda and her with Lisa. Here, gender and number consistency cannot resolve the antecedents, but other cues, such as pragmatics, can. Listeners reason that the person in the position of hiring is the one who would be attentive to a candidate’s skills. If pragmatics were not enough to resolve the ambiguity, other cues such as the grammatical role of each entity can help listeners. Defining how these pieces interact during processing introduces many complexities for a unified theory of pronoun
resolution, and unfortunately, it is easy for researchers to selectively neglect some of the information used to identify a referent. In addition, few have addressed the question of how prosody, in particular pitch accent, interacts with this process.

What follows is a brief review of some influential theories of pronoun and reflexive referent resolution, followed by separate work on the representation and meaning of prosody. After this, some proposals integrating the two will be presented. What will become apparent is that there is a large collection of data involving prosody for which these theories do not account.

Chapter 2 continues with a description of on-line methodologies for probing the unconscious processing of anaphora. Using on-line techniques, researchers have discovered many things about how pronoun and reflexive referent resolution is represented in the mind, and the second half of this chapter provides a summary of these findings. Here again, however, the paucity of research devoted to prosody’s real-time influence underscores the need for more research and better theories recognizing prosody’s role.

In Chapter 3, I present several experiments that explore the role of prosody on pronoun referent resolution from a new perspective. Specifically, I show that focus stress impacts the interpretation of pronouns in subtle ways and that it interacts with other, more powerful forces. The findings suggest that there is more to prosody and its influence on pronouns than previously thought.
Chapter 4 deals with pitch accent used for a different purpose: contrastive stress. The experiments reported in this chapter focus on the time course of prosody and uncovers whether it is a fast-acting cue or one that appears in the later stages of processing.

Finally, in Chapter 5, I propose a unified theory of the real-time processing of discourse anaphora in English, called the Discourse Anaphora Weighting Approach (DAWA). The DAWA takes into account the data up to this point and promotes the role of prosody as an important factor in coreference.

Theories of Pronoun Referent Resolution

Referents Outside the Local Discourse

Pronoun resolution is influenced by many linguistic factors. At the most basic level, antecedents that are part of the same scene or the common ground between the speaker and listener are preferred (Anderson et al., 1983; Gordon & Scerace, 1995; Greene, Gerrig, McKoon & Ratcliff, 1994; Lesgold, Roth & Curtis, 1979). As evidence of this, Daneman & Carpenter (1980) discovered that the more sentence boundaries between the pronoun and its referent, the more difficult pronoun resolution was for the participant. Since pronouns offer no semantic content on their own, the shared understanding between the speaker and listener is that the antecedent was mentioned or implied as part of the local context of the discourse.
Number, Gender, and Animacy Agreement

Additionally, there are morphological and syntactic constraints imposed by the type of pronoun used. Although English is not a language rich in morphology, pronouns are marked with number and gender, and these features must be consistent with the pronouns’ referent. Animacy is also a feature that both a pronoun and its antecedent must share. Arnold (1998) offers the following examples for consistency of gender and animacy between a pronoun and its referent:

\[(2)\]
\[a. \text{Lisa, sat down next to Ben. She ordered a slice of pie.} \]
\[b. \text{Lisa sat down next to Ben. He ordered a slice of pie.} \]

\[(3)\]
\[a. \text{Lisa bought a slice of pie. She was very hungry.} \]
\[b. \text{Lisa bought a slice of pie. It was freshly baked and smelled wonderful.} \]

The examples above very simply illustrate how the pronoun can refer only to one of the two entities mentioned in the previous sentence given the pronoun’s morphology.

Binding Constraints

Given this basic assumption of pronoun/referent agreement, syntactic position also plays a role in constraining the potential antecedents of an anaphor. In *Lectures*
on Government and Binding (1981), Chomsky specified the rules that govern which syntactic relations license referents for anaphors such as reflexives (himself/herself), pronouns (he/she and him/her) and referring expressions or R-expressions (Dick/Jane/dog). The syntactic rules are limited to the domain of the sentence or clause. He formally defines the domain in terms of a binding category, where “Beta is a binding category for alpha if and only if beta is the minimal category containing both alpha and a SUBJECT accessible to alpha.”

In the following tree structure, the left-most node is the SUBJECT position.

```
S
   /\   /
 NP VP
  /   \ /
 Garrison V NP
   \   | |
       introduced Abigail (alpha)
```

Thus, S is the binding category for alpha because it contains both alpha and a SUBJECT. There are no other nodes in the tree for which both of these conditions are true. Now, subject position takes on a special role. If the subject is coindexed with anything within its sister node, the verb phrase, or any of the VP’s children, then the subject binds that entity. Nothing binds the subject position itself, so it is considered free.
Given that the binding category has been established (the sentence) and the status of each constituent has been defined as either bound or free, the following three principles define the syntactic constraints on coreference.

Chomsky’s Principles of binding theory:

(A) An anaphor is bound in its binding category
(B) A pronominal is free in its binding category
(C) An R-expression is free

Violations to these principles result in ungrammatical utterances as indicated by the examples with asterisks below. Subscript i indicates coreference:

(4) a. *Himself$_i$ introduced Garrison$_i$.
   b. Garrison$_i$ introduced himself$_i$.

In 4a, the anaphor is free because it is in subject position. Since, Principle A has been violated (an anaphor must be bound), the utterance is ungrammatical. 4b introduces no violations because the anaphor is in object position and is bound by the subject.

The opposite is true for pronouns:

(5) a. *Garrison$_i$ introduced him$_i$.
   b. Abigail introduced him.
The utterance in sentence 5a is ungrammatical because the pronominal in object position is bound by the subject, resulting in a violation of Principle B. In 5b, the pronoun refers to an entity outside the scope of the sentence. Since *him* is not bound by the subject, leaving the pronominal free in its binding category, the sentence is grammatical. R-expressions cannot not be bound either:

(6)  

(a. *Garrisoni introduced Garrisoni. 

(b. Abigail introduced Garrison.

Here, the R-expression in object position violates Principle C because it is bound by the subject. In the grammatical counterpart in 6b, the R-expression is free because it is not coindexed with the subject.

Even though syntactic coreference constraints specified by Chomsky’s binding principles limit the possible referents that are grammatically licensed for any given pronoun, these syntactic constraints tell only part of the story. Rarely are sentences spoken in isolation. Often a pronoun’s referent does not lie in the sentence itself but in the previous discourse. Thus, the discourse context must be taken into consideration. One very influential theory regarding the constraints of pronoun coreference in discourse is Centering Theory.
Centering Theory

In 1986, Grosz and Sidner developed a comprehensive theory of discourse structure in which a discourse consists of segments that have intentional structure, linguistic structure and attentional state (Grosz and Sidner, 1986). At any point in the discourse, attentional state models the listeners’ focus of attention by tracking the salient entities and relations at both global and local levels. The form of an entity (for example, whether it is presented as a pronoun or R-expression) signals its level of attentional processing. Specifically, R-expressions are associated with global focusing and pronouns with local focusing (Grosz, Joshi, & Weinstein, 1995). Centering is a mechanism that functions at the local level – it determines which entities are most accessible (Sidner, 1979). For any given utterance, entities are listed in a ranked set of forward-looking centers (C_f), (which are potential antecedents for future utterances). For English, the entities are partially ordered according to grammatical function, where subject is assigned the highest rank, direct object, the next highest, followed by other grammatical roles such as indirect objects and adjuncts. (Brennan, Friedman, & Pollard, 1987; Gordon & Chan, 1995; Grosz et al., 1995). The most salient entities are those in the highest positions on the C_f list. The entity that actually links the current utterance (U_n) to the previous utterance (U_{n-1}) is referred to as the backward-looking

1. Brennan et al. (1987) acknowledge that speakers do not use grammatical role exclusively for pronoun preferences, but they claim that the algorithm proposed is more successful than those that depend solely on recency or parallel grammatical function. Other researchers have enhanced this ordering scheme in order to account for more data. For example, Gordon et al. (1993) state that fronted constituents are higher ranked than subjects. By changing the ranking, they can include syntactic focus as a factor in pronoun referent resolution within the framework of Centering.
center (C_b). Usually, the entity most likely to be the C_b is the highest-ranked member of the C_f list of the previous utterance, or the preferred center (C_p).

When a change in the attentional state (or a change in Center) takes place, it is referred to as a transition. There are different classes of transitions, and these depend on the amount of change. The Center could remain the same from the current utterance to the next; the Center could retain its role, but change in the next utterance since it has moved down the C_f list and is no longer preferred; or the Center could shift to a different entity in the current utterance and shift the focus of attention. If the new Center is the C_p, then the shift is smooth; whereas, if the new center is not the C_p, the shift is considered rough (Brennan et al., 1987). The local coherence of the discourse depends on the speaker’s transitions, where the most coherent discourses maintain the same Center, and incoherent ones contain frequent shifts.

The following is an example of how the Centering algorithm results in pronoun referent resolution using Brennan et al’s algorithm.

\[(7)\]

\(a.\) Karen, likes Cathy.

\(b.\) She often invites her to lunch.

\(c.\) She likes chatting about snowboarding.

In this example, we will predict what the most coherent referent of She is in 7c. To do this, we consult the C_f list of the previous utterance. The C_f(U_{n-1}) list of sentence 7c is (Karen, Cathy). The pronouns from 7b have been expanded to the R-expressions they
refer to. In sentence 7b, Karen (represented by She) is in subject position, therefore, Karen is higher ranked in the list and takes on the role of C_p.

For the current utterance, we have to figure out what the C_f list is. All the possible C_b-C_f combinations for the current utterance are generated to help narrow down the choices. (NIL represents the possibility that no C_b can be identified for the utterance):

<table>
<thead>
<tr>
<th>Candidate</th>
<th>C_b</th>
<th>C_f list</th>
</tr>
</thead>
<tbody>
<tr>
<td>i.</td>
<td>Karen</td>
<td>Karen, snowboarding</td>
</tr>
<tr>
<td>ii.</td>
<td>Karen</td>
<td>Cathy, snowboarding</td>
</tr>
<tr>
<td>iii.</td>
<td>Cathy</td>
<td>Karen, snowboarding</td>
</tr>
<tr>
<td>iv.</td>
<td>Cathy</td>
<td>Cathy, snowboarding</td>
</tr>
<tr>
<td>v.</td>
<td>NIL</td>
<td>Karen, snowboarding</td>
</tr>
<tr>
<td>vi.</td>
<td>NIL</td>
<td>Cathy, snowboarding</td>
</tr>
</tbody>
</table>

Then filters are applied to eliminate any combinations that violate basic syntactic constraints on coreference. Since She is both the backward-looking Center and the first item in the C_f list of the current utterance, and since She cannot refer to two entities at the same time, candidates ii, iii, v and vi are removed. This leaves i and iv as the final candidates. At this point, the transitions are ranked, and the C_b-C_f combination that provides for the most coherent discourse will become the preferred interpretation. When Karen is the C_b, the transition is Continuing in that Karen remains the Center in the current utterance and will most likely be the Center in the
following utterance. In contrast, when Cathy is the $C_b$, the transition is Shifting. Since Continuing is preferred over Shifting, candidate (i) in which Karen is the referent for *She* is selected.

In addition to helping to define which referent would be preferred for any given pronoun, Centering Theory makes stipulations about how an element should be realized in subsequent utterances:

Rule 1: If any element of $C_t(U_n)$ is realized by a pronoun in $U_{n+1}$, then the $C_b(U_{n+1})$ must be realized by a pronoun also.

Rule 1 captures the strangeness of the third sentence in the following example from Grosz et al. (1995):

(8) a. He has been acting quite odd. [$C_b = John = \text{referent ("he")}]$

b. He called up Mike yesterday. [$C_b = John = \text{referent ("he")}]$

c. John wanted to meet him urgently.

Using *he* in place of *John* in 8c makes the discourse more coherent, a preference that Centering Theory addresses in Rule 1, since John is the $C_b$ rather than Mike. In fact, psycholinguistic data have shown that reading times tend to be slower when R-expressions are used in a discourse after pronouns in subject position have been used
to refer to the same entity in previous sentences\(^2\) (Chambers & Smyth, 1998; Gordon et al., 1993).

Studies have shown that prior topicalization makes an antecedent more salient, as is consistent with Centering Theory (Anderson, Garrod & Sanford, 1983; Sanford, Moar & Garrod, 1988). Centering Theory can also account for preferences in which a pronoun’s referent appears in subject position (although Centering does not require the referent to be the subject). The tendency to select the subject as the referent is quite strong and has been observed by a number of researchers (Garnham, Traxler, Oakhill & Gernsbacher, 1996; Gernsbacher & Hargreaves, 1988; Stevenson, Nelson & Stenning, 1993). However, Centering does not account for all the data (Kehler, 1997), so other researchers have offered different theories.

**Pragmatics and Theta Roles**

Pragmatics influences interpretations in ways not predicted by Centering Theory. In the following example adapted from Winograd (1972), the referent of the pronoun *they* changes depending on the verb.

\(^2\)Sidner (1983) claims that a pronoun is more likely to have a pronoun as its antecedent than an R-expression as in the following examples:

\[(9)\]

a. I got a new hat and I decorated it with a big red bow.
   b’. I think the bow will brighten it up a lot.
   b’’. I think it will brighten up the hat a lot.
   c. If not, I think I’ll still use it.

The intuition here is that *it* looks to the pronoun as its antecedent in the previous sentence. Therefore *it* in 9c preceded by 9b’ would refer to a *new hat*, and *it* in 9c preceded by 9b’’ would refer to a *big red bow.*
(10) a. The city council denied the demonstrators a permit because they feared violence.

b. The city council denied the demonstrators a permit because they advocated violence.

If the verb is *feared* as in 10a, the pronoun *they* refers to the *city council*; whereas if the verb is *advocated* as in 10b, the pronoun refers to the *demonstrators*. Here, it is not grammatical function that is driving informants to make different interpretations. The fact that the syntax in the two versions of the sentence is identical implicates the semantics of the respective verbs and general world knowledge about which entity might fear versus advocate violence. Evidence for preferences based on pragmatics has been shown in many psycholinguistic studies (Caramazza, Grober, Garvey & Yates, 1977; Caramazza and Gupta, 1979; Ehrlich, 1980; Garvey and Caramazza, 1974).

The problem with varying pragmatics in discourse is that often the strength of pragmatics as a cue to the listener to switch the pronoun referent is unknown until the experiment has been conducted. In order to predict the direction of the verb’s bias consistently, some researchers have focused on the effect of thematic roles[^3] on

[^3]: Every verb assigns the roles that each of its arguments plays in relation to one another. These are called theta roles. For example, in the sentence, “The boy feared the girl,” *the boy* receives an Experiencer theta role because he experiences the fear; whereas *the girl* is the Theme or Stimulus, namely the object of the fear. In contrast the theta roles in the sentence “The boy frightened the girl,” are the opposite - *the boy* is the Theme and *the girl* is the Experiencer. Thus, theta roles can be different even though the syntactic structure of the sentences is identical (Belletti and Rizzi 1988).
pronoun referent resolution. Consider the following examples from Stevenson, Crawley & Kleinman (1994):

(11)  

   a. Ken admired Geoff. He …
   b. Ken impressed Geoff. He …

In 11a, Ken is assigned the theta role Experiencer and Geoff the theta role Stimulus (or Theme); whereas in 11b, the opposite theta roles are realized. Interestingly, Stevenson et al.’s results show that He tends to corefer with the constituent assigned the Stimulus theta role, regardless of grammatical position. Likewise, with the thematic roles Goal and Source, informants tend to prefer the entity assigned Goal as in the following examples:

(12)  

   a. John seized the comic from Bill. He …
   b. John passed the comic to Bill. He…

In 12a, John is the Goal and therefore, native speakers coindex He with John. The preferences are not as symmetrical, however. In 12b, Bill is the Goal (but in object position), and here, informants choose one referent just as often as the other.

Arnold (1998) corroborated Stevenson et al’s results. In a production experiment with Source/Goal verbs, she found that participants who continued stories with a pronoun referred to Goals more often than Sources. She found the same
patterns in a corpus analysis as well, although in both cases there was a strong preference for referents to be in subject position.

Hobbs (1979) developed a coherence-driven theory to account for how pronoun resolution is intimately tied to the semantics of an utterance. In simple terms, the referent of a pronoun naturally falls out from world knowledge and general reasoning when interpreting an utterance. Take sentence 10a, presented here as 13:

(13) The city council denied the demonstrators a permit because they feared violence.

In his book, Kehler (2002) steps through the reasoning required to interpret (13) with an Explanation coherence relation (indicated by because). The identification of the pronoun’s referent results as a by-product of this reasoning process.

Pronouns are considered free variables that are bound during the process of making inferences. World knowledge is represented by axioms in first-order logic. Kehler offers the following axiom:

\[ \text{fear}(X, V) \land \text{advocate}(Y, V) \land \text{enable_to_cause}(Z, Y, V) \supset \text{deny}(X, Y, Z) \]

Through an inference process, the variables are bound to entities: X is bound to city council and V to violence. The because clause establishes a link between the two clauses of the sentence. The predicate of the because clause has its own representation which uses the variable T in place of the pronoun they: fear(T, violence). When this
representation is used to identify an antecedent from the relation already established, T becomes unified with X. It follows, then, that since X is bound to city council, T is also bound to city council. Through this process of using discourse coherence, inference and logic, pronouns become coindexed with the semantically appropriate referent.

Even though these accounts handle both subject assignment and pragmatics, there are yet other factors that influence pronoun referent assignment such as parallelism.

**Parallelism**

Pronouns in parallel syntactic structures introduce more complexity to the problem of pronominal reference resolution. Take the following examples from Oehrle (1981):

\[(14) \quad \begin{align*}
&\text{a. Felix}_i \text{ hit Max and then he}_i \text{ hit Bill.} \\
&\text{b. Felix hit Max}_j \text{ and then Bill hit him}_j.
\end{align*}\]

In 14a, *Felix* is in subject position and also shares the same grammatical function as the pronoun *he*; therefore, it makes sense that *Felix* is identified as the referent of *he*. In 14b, the referent in subject position (*Felix*) and the referent sharing the same grammatical function as the pronoun (*Max*) are different. In this case, informants prefer the referent in object position to corefer with the pronoun.
Psycholinguistic researchers have found evidence for these preferences (Chambers & Smyth, 1998; Corbett and Change, 1983; Grober, Beardsley & Caramazza, 1978; Springston, 1975; Smyth, 1994). In one experiment, Smyth (1994) used parallel structures such as the following:

(15) Robert bullied Peter and Melanie attacked him.

(His examples were based on the materials from Crawley, Stevenson, & Kleinman (1990), with the intention of arguing against Crawley et al.’s conclusions.) He found that participants interpreted the referent of the object pronoun to be the noun phrase in object position of the previous clause. The less parallel the clauses of the sentence, the less robust the effect. Consider the following two examples:

(16) Sarah visited Cathy at home and Charles phoned her at work.
(17) Ben talked to Steven about the exams and Sue kissed him.

Subjects were much more prone to select Cathy as the referent in sentence 16 with complete parallel structure as opposed to 17 in which both grammatical function and additional adjunct phrases are not exactly parallel. From these results, Smyth proposed his Extended Feature Match Hypothesis in which pronouns in parallel constructions (as in conjoined clauses) are resolved by a search mechanism that identifies an antecedent with the same grammatical role as the pronoun. Parallelism is
determined by the number of features shared by the pronoun and each potential referent. Smyth posits that a strategy based on subject assignment is the default when a threshold for ‘nonparallelism’ is reached.

A different approach, which attempts to account for the same data, is Stevenson et al.’s (1995) competition model in which the subject assignment and parallel function strategies each produce a candidate interpretation. These interpretations battle against one another until one is declared the final interpretation of the utterance.

Taken together, the theories of pronoun resolution mentioned above account for first-mention biases of coreference and preference for the subject, parallelism, and pragmatics\(^4\). However, none of the theories discussed so far account for all three factors. Some researchers including Kameyama and Kehler have explored how these influences interact with one another (or do not interact, as the case may be).

**Dynamic Preference Model**

Kameyama (1996) developed a Dynamic Preference Model that defines how parallelism and general world knowledge interact with Centering preferences. Kameyama defines context as a multi-component data structure that includes Attentional State, a Logical Form (LF) Register (which is the preferred interpretation of the previous utterance), and a Discourse Model. Attentional State draws from

\(^4\) Other factors such as distance, syntactic focus and pronoun chaining come for free in Centering Theory.
Centering Theory and defines when entities are salient based on ranked centers. The Logical Form Register keeps in its buffer the Logical Form (LF) of the previous utterance. Kameyama states that the LF of the previous utterance seeks maximal parallelism with the current utterance. Thus, when the salience-based preference from the Attentional State is indeterminate, a preference based on parallelism wins out. According to Kameyama, each component generates its preferred interpretation independently, and these compete and ultimately result in a final preference. Interpretations from one component override contradictory interpretations from another component lower on a preference hierarchy shown below (where > indicates that the component shown on the left can override the component on the right):

Syntax and Semantics > World Knowledge > Attentional State > Logical Form Register

Competition approaches in which candidate interpretations compete for the status of final interpretation have been challenged by Kehler (2002) in his recent book on discourse coherence. His view is that there are qualitative differences in the materials used to measure pronoun resolution. Listeners constantly reconcile how one utterance is related to another in a discourse, and even make assumptions to establish coherence. There are different ways that utterances can be connected, and these relations between utterances have been categorized into classes. Kehler posits the existence of three classes: Contiguity, Resemblance and Cause-Effect. He shows how the coherence relations inferred for the stimulus items in the pronoun interpretation
literature explain some discrepancies in how participants resolve pronouns. Thus, theories in which candidate preferences compete with one another for the role of final interpretation are obviated. Participants derive interpretations based on the coherence relations suggested by the stimuli. Once a coherence relation such as Cause-Effect has been established, there is no need to apply an inappropriate grammatical function heuristic.

There is evidence to support Kehler’s claim. In Stevenson et al.’s (1995) third experiment looking at the effect of thematic structure on pronoun referent resolution, they introduced because to the materials and found that the patterns changed. Taking Kehler’s perspective, the experiments without because most likely suggested an Occasion coherence relation whereas the third experiment with because promoted an Explanation relation. As a result, different results were observed. As further evidence, Ehrlich (1980) used clauses conjoined with the words because, but and and. When because was used, the effects of implicit causality on pronoun resolution were strong. With the conjunction but, the effects were weaker but still observable, and with and, no effect was measured. Similarly Grober, et al. (1978) found that in sentences such as “John must scold Bill because/but he …” participants continued the sentence with different intended referents as a result of the connective. With because, there were more assignments of the second NP compared to when the connective was but. This makes sense given that but suggests different coherence relations (Contrast, Violated Expectation) than does because (Explanation).
Despite all of the work on the factors that affect pronoun coreference and how they interact with one another, many of the theories neglect another factor: prosody, or the way that the utterance is spoken. A working definition of prosody is the acoustic patterns of F0, duration and amplitude and the higher-level, abstract structure that can account for these patterns in a separate component of the grammar (Shattuck-Hufnagel & Turk, 1996). To understand how prosody intermingles with pronoun resolution, a brief explanation of a theory of prosody, specifically intonation, is presented.

Theory of Intonation

According to Pierrehumbert’s (1980) model of intonational structure, the phonological representation of intonation is autosegmental, meaning that it is represented on a separate tier, apart from the segmental and/or syllabic properties of the utterance. It consists of a linear sequence of atomic tonal units, with three units being essential to the intonational contour. These tonal units are pitch accent, phrase accent and boundary tones.

Pitch accent is associated with a primary stressed syllable, and in English, its placement is important to the notion of focus. The stressed syllable is identified by either a metrical grid or strictly layered prosodic tree with weak/strong nodes. For each Intonational Phrase there must be at least one pitch accent. The phrase accent is realized within a certain time period after the final, (nuclear) pitch accent. And finally, boundary tones are associated with the initial and final syllable of the
Intonational Phrase. Boundaries have strengths (0-6), with strong boundaries exhibiting lengthening of the final accented syllable and a pause before the next word.

Each of these tonal entities is represented by two tonal levels, high (H) and low (L). Pitch accent is represented with an asterisk and consists of one or two tones: a main tone and an optional leading or trailing tone. The following pitch accents have been identified: H*, L*, L*+H, L+H*, H+L*, H*+L, and H*+H. The phrase accent is a single tone, occurring after the final pitch accent. The boundary tone is also a single tone, and is written as either H% or L%. The following example shows each of these tones over an utterance taken from Pierrehumbert (1980):

(18) *Legumes are a good source of vitamins.*

H* L H%

H* under the first syllable of the utterance is the pitch accent, L the phrase accent and H% the boundary tone. These tonal units make up the intonational structure of the utterance.

In addition, the tonal units provide meaning to the utterance. The study of intonational meaning in English consists of two types: focus structure and expressiveness (Halliday, 1967; Crystal, 1969; Bolinger, 1965; Ladd, 1980). In light of Pierrehumbert’s theory, focus structure is the meaning associated with pitch accent and is linked to the information structure of the utterance. Constituents with pitch accent are considered to be focus marked (Selkirk, 1984). Expressiveness is the
meaning associated with the types of tones used across the Intonational Phrase, for example, questioning intonation versus statements of fact. Expressiveness gives some indication of the belief system of the speaker regarding the content of the utterance.

So, how does pitch accent come into play with pronoun coreference? The most well studied cases are those that involve contrastive stress and parallelism.

Pronouns and Prosody

As was illustrated in the examples of parallelism, when a pronoun appears in a sentence or clause that is syntactically parallel to the previous one, native informants often assume the referent is in the same grammatical position as the pronoun. When contrastive stress is added to the mix, however, interpretations change. Lakoff (1971) presented the following example as an illustration of how pitch accent as contrastive stress on a pronoun can shift the interpretation of the referent:

\[(19)\]

\[\begin{align*}
\text{a. } & \text{John}_i \text{ called Jim}_j \text{ a Republican, then he}_i \text{ insulted him}_j. \\
\text{b. } & \text{John}_i \text{ called Jim}_j \text{ a Republican, then HE}_j \text{ insulted HIM}_j.
\end{align*}\]

19a follows the typical preferences associated with parallelism. However, in 19b, *He* is coreferenced with *Jim* and *him* with *John*, meaning Jim interpreted John’s label, Republican, as an insult and retaliated with his own.
Contrastive stress gives prominence to certain elements within the discourse. The default position of normal or nuclear pitch accent in English tends to come towards the end of an expression. Sometimes a speaker may choose to override this default pattern and deliberately place the stress on a different portion of the sentence to direct the listener’s attention to information that may not be anticipated (Mansfield, 1997). For example, contrastive stress on the pronoun in 19 above indicates that the speaker’s intended referent is ‘atypical.’ Contrastive stress informs the listener that the pronouns are not going to be assigned to the referents with parallel function as would be expected, but rather will be coindexed with the other possible antecedent.5 Kerto (1991) provided evidence of these judgments. Using sentences with parallel structure and a subject pronoun in the second clause, she found that adult subjects typically selected the first NP as the referent of the pronoun; however, with contrastive stress on the pronoun, subjects tended to choose the second NP. Results have yet to be reported with contrastive stress on an object pronoun.

Theories that Bring Together Prosody and Pronoun Referent Resolution

Several researchers have attempted to incorporate prosody into accounts of pronoun coreference. Cahn (1995) captured the interaction of pitch accent and pronoun referent resolution in the framework of Centering Theory. In her paper, she claims that a contrastively stressed pronominal may force a shift in Center. First, she makes a distinction between attentional salience and propositional salience, where

5. How this interpretation unfolds in real time will be explored in Chapter 4.
attentional salience is expressed as the ranking of items in the $C_f$ list, and propositional salience is the status of an entity based on its relation to mutual beliefs assumed during the discourse. Pitch accents can augment either attentional salience or propositional salience, depending on the form on which they are realized. Pitch accents on R-expressions, which are rich in semantic content, enhance mutual beliefs and therefore function at the propositional level. In contrast, pitch accents on pronouns, which are semantically impoverished, function at the attentional level. When the specific pitch accent pattern L+H* is realized on a pronoun, it forces a reordering of the $C_f$ list such that a $C_b$ other than the one predicted by Centering Theory is selected. Of course, an alternate entity must be available for the shift in Center to take place. The thrust of Cahn’s argument, then, is that pitch accent indicates a Center shift.

Nakatani (1997) found support for pitch accent being associated with a shift of attention. In a spontaneous narrative monologue she discovered that presence of a pitch accent on an entity systematically indicated changes in attentional state. In the monologue, when subjects’ pronouns were stressed this usually meant a local shift to a forward-looking Center that was not the current $C_b$. From these results, she posited an algorithm for processing pronominal expressions. If a subject pronoun is stressed and if the referent is in the $C_f$ of the previous utterance but is not the $C_b$, then the referent is promoted to $C_p$ in the current utterances’ $C_f$ list. If an object pronoun is stressed, then the referent is placed just under the subject in the current utterance’s $C_f$ list. In this way, pitch accent shuffles around the rank order of the constituents in the $C_f$ list, potentially shifting the attention of the discourse.
In a slightly different approach, Beaver (2002) uses pronoun stress in an example of how discourse anaphora in Centering Theory can be reformulated in terms of Optimality Theory (Prince & Smolensky, 1993). The input to Beaver’s model is form, or syntactically analyzed sentences. The output is meaning, for example, a mapping of pronouns to their antecedents. The optimal interpretation results in the fewest and weakest violations of a rank-ordered set of constraints. As an example, the constraint AGREE prevents a pronoun’s antecedent from being a different number or gender from the pronoun; COHERE prefers the topic of the current sentence to be the topic of the previous one (an embodiment of the Continuing transition in Centering Theory) and ALIGN prevents the topic of an utterance from being in a place other than subject position. Of course, not all the constraints will always be met, and from those mentioned above, AGREE is the strongest and ALIGN the weakest.

When considering stress, Beaver adds another constraint, Avoid F, or “avoid focus” which is appropriate for pronouns since the information pertaining to the referent is already ‘given’ in the discourse. Consider the following example (slightly simplified from Beaver’s original example):

(20)  a. Fred$_i$ was eating.  
b. He$_i$ saw Jim$_j$. 

 c’. He$_j$ winked.  
 c’’. HE$_j$ winked.
Capitals indicate focal stress on the pronoun. When presented in a table with the rank-ordered constraints, the following pattern results (only relevant constraints are shown here for simplicity):

<table>
<thead>
<tr>
<th>Context</th>
<th>AGREE</th>
<th>COHERE</th>
<th>ALIGN</th>
<th>Avoid F</th>
</tr>
</thead>
<tbody>
<tr>
<td>20a, b</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Meaning:

Winked(j)

<table>
<thead>
<tr>
<th>20c’ He winked.</th>
<th>*</th>
</tr>
</thead>
<tbody>
<tr>
<td>20c’’HE winked.</td>
<td>*</td>
</tr>
</tbody>
</table>

The asterisks show which constraints are violated by the candidate in that row. This example shows 20c’ as the preferred realization since it has the fewest constraint violations. However, winked(j) in which Jim is doing the winking is not the optimal meaning for He winked. Rather native informants would conclude from the discourse that Fred was doing the winking. Therefore, stress on the pronoun in this case results in an interpretation that is not ‘bi-directionally’ optimal. To solve this problem, yet another constraint is introduced: *BLOCK. This constraint is violated if the meaning of the utterance does not derive the optimal form (as in the example above). With this additional constraint, the following revised table is produced:

<table>
<thead>
<tr>
<th>Context</th>
<th>AGREE</th>
<th>*BLOCK</th>
<th>COHERE</th>
<th>ALIGN</th>
<th>Avoid F</th>
</tr>
</thead>
<tbody>
<tr>
<td>20a, b</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Meaning:
Here 20c’’ does not violate the *BLOCK constraint; therefore, it becomes the optimal meaning (with the fewest constraint violations). Thus, with only a few constraints, Beaver shows how a representation of Centering Theory reformulated into Optimality Theory can account for pronoun coreference.

So far, three views of how prosodic stress and pronoun resolution come together have been presented. Interestingly, all three use Centering Theory as the framework for choosing from potential referents. But, as discussed above, a limitation of Centering Theory is that it does not capture the influences of object pronouns in parallel constructions (or pragmatics). This is ironic, given that both Cahn and Beaver use parallelism examples in support of their approaches. The reason the examples do not snag their theoretical reasoning is that they both use subject pronouns – thus, even though the parallel construction is there, the pronoun shares the same grammatical function as the referent in subject position, and so the prediction from a parallel grammatical function strategy remains consistent with the predictions of Centering Theory. Thus, none of these proposals account for the following:

(21) a. The butcher hit the baker$_i$. Then the candlestick maker hit him$_i$. 
b. The butcher hit the baker. Then the candlestick maker hit HIM.

Centering Theory predicts that the most coherent interpretation for the pronoun in 21b is *butcher* because it is the most salient entity. Theories incorporating prosody within the Centering framework would say *baker* is the most felicitous for 21b. However, the opposite is true. Thus, the preferred interpretations in 21 go beyond the scope of Centering Theory, and as a result, none of the explanations can account for pronoun resolution with stress.⁶

Kameyama (1996) realized that there are more forces at work in resolving the antecedent of a pronoun; therefore, she proposed a hypothesis of how prosody affects pronoun resolution, building from her Dynamic Preferences Model. In her Complementary Preference Hypothesis, she posits that there is a systematic relation between interpreting the antecedent of an unstressed and stressed pronoun, such that the preferred value of a stressed pronoun takes the complementary preferred value of its unstressed counterpart. First the partially ranked set of potential antecedents from the previous utterance is listed. Then preference for the unstressed counterpart is

---

⁶ Smyth (1994) mentions how stress on a pronoun changes the default behavior of his Extended Feature Match Hypothesis. In a parallel sentence construction, if a pronoun’s grammatical role matches that of a potential antecedent’s, then the pronoun selects that antecedent; however, if the pronoun is stressed, this selection is blocked. Unfortunately, Smyth’s proposal does not account for those cases where only one antecedent is being considered (in which case the pronoun would not be assigned a referent). Also, the Extended Feature Match Hypothesis itself does not cover pragmatic factors in pronoun referent resolution.
computed; during this step in the algorithm, other components of the context such as parallel structure influence the final preference. Next, the complementary preference order for the stressed pronoun is computed. Coherence of the current utterance with the previous is then established. If the interpretation of the unstressed pronoun is indeterminate, then the resolution of the stressed pronoun remains ambiguous as well. Also, if the referent of the pronoun is unambiguous, then the same referent will be chosen for the stressed pronoun since both the stressed and unstressed pronouns draw from the same set of possible antecedents.

Kameyama’s approach accounts for the most data. Her hierarchy captures different factors that influence pronoun resolution including preference for subject position, parallel grammatical function, and pragmatics. Thus, her account of how stress interacts with these factors is also more complete. Unfortunately, the data do not always fit the model. For example, in Venditti, Trueswell, Stone and Nautiyal (2003)’s picture eye tracking experiment, contrastive stress did not always sway interpretations. There are other shortcomings of Kameyama’s theory as well. For one, it does not predict why certain factors override others. We are again reminded of Kehler’s insight that competition between factors is obviated when coherence relations are considered. Another question is when do these effects take place. As a listener is trying to construct meaning from the utterance, what processes are happening in real time that help bring the listener to a specific interpretation?
What is clear at this point is that there are many factors that influence coreference including number, gender and animacy agreement, syntactic position, grammatical function, thematic roles, pragmatics, and discourse coherence relations. Additionally there is still some room to consider a broader role of pitch accent. In the next chapter the topic of how all these constraints and preferences come into play as the listener is hearing a sentence/discourse in real time is discussed in detail.
ON-LINE METHODOLOGIES AND FINDINGS

The previous chapter explored some discourse-related preferences and constraints on coreference. Although formal definitions provide us with a better understanding of pronouns, they do not tell us anything about the mechanisms underlying these preferences and constraints. Thus, there is an entire discipline dedicated to learning how comprehension unfolds in real time. With many linguistic and computational linguistic approaches, there is an assumption that the interpretation is based on a complete utterance. For the psycholinguist, sentences are often continuing as processing takes place, and therefore not all the information is immediately available. The terms ‘off-line’ and ‘on-line’ are often used to distinguish at what stage measurements of the comprehension process are being gathered. If information is recorded off-line, then participants are consciously pondering their interpretations. Such measurements are simple question-and-answer tasks aimed at determining how the subject understood the stimuli. When measurements are gathered on-line, ideally, participants are blissfully unaware of their comprehension processes. During on-line processing, implicit understanding is taking place, measured by subtle differences in indirect metrics. These slight differences, for example of reaction time, convey a vast amount of information about what is going on in the human mind as the listener is undergoing the act of comprehension.

There are many assumptions about how information reaches different processing components and what these components might be. Some of these
assumptions will be put forth here before addressing the more in-depth issue of processing pronouns and prosody.

First, sounds of speech hit the human ear, and a signal is relayed to the neurons in the audio cortex of the brain. From here there is a complex interplay between brain regions involved in language processing, such as Wernicke’s area (Broadmann’s area 22 of the temporal lobe), which assists semantic processing (Goodglass & Geschwind, 1976; Hillis, et al. 2002), and Broca’s area (Broadmann’s areas 44 and 45 of the frontal lobe), which is critical for syntactic processing (Zurif, Swinney, Prather, Solomon & Bushell, 1993; Swinney, Zurif, Prather & Love, 1996) and important for anaphora resolution (Avrutin, 1999; Grodzinsky, Wexler, Chien & Marakovitz, 1993; Love, Nicol, Swinney, Hickok & Zurif, 1998). At a conceptual level, the listener segments the string of phones into syllables and words. The words are retrieved from a lexicon and are tagged according to grammatical categories. Words are grouped into constituents and placed into a hierarchical syntactic structure. If pronouns are encountered, anaphoric processing takes place.

There has been a formidable amount of research on the mechanics underlying the processing of pronouns. However, before going in the details of what is known about how that anaphoric processing happens, a review of the different techniques used to measure on-line processing is presented.
Methods Used in On-Line Psycholinguistic Experiments

The challenge of probing the unconscious processes of language comprehension is that all the measurements are indirect. When asking for someone’s interpretation of an utterance, that person will respond with the end-result of the completed processes with no ability to articulate how that interpretation came to be. Over the years, researchers have devised creative techniques to record comprehension as it is happening in real time. This next section delves into these different techniques, and describes some of the more common ones used in psycholinguistic research.

Reading and Listening Times

In self-paced reading and listening tasks, participants control how fast they are presented with information. Sentences are revealed section-by-section (usually word-by-word or phrase-by-phrase). The participant simply clicks a button connected to a computer set-up to proceed to the next section. In a reading time experiment, as one section appears on the screen, the previous one disappears. The time the participant takes on each section is recorded and analyzed. The materials in such experiments (as with all psycholinguistic studies) need to be crafted carefully so that key sections of the sentences can be minimally compared with those of other sentences varying only with regard to a single manipulation. If the reading or listening times for a section of one type of sentence are significantly slower than for the other type, the researcher can conclude that processing the former is more cognitively taxing. If the difference is not
observed until a later section in the sentence, then the effect does not take place immediately, but rather happens ‘downstream’ from when the pronoun was heard.

The drawback of reading time experiments is that they cannot tap into auditory comprehension processes. This deficiency is most notable when trying to investigate phenomenon unique to speech such as prosody. Although listening time experiments overcome this problem, the materials in these studies need to be artificially segmented in order to be able to measure listeners’ reaction times. The fact of the matter is that we do not normally listen to speech in this way, and so the measurements are based on artificial scenarios.

**Eye Tracking**

Several variations of eye tracking methods exist in the literature. Some measure reading times while others assess semantic processing through the use of picture viewing. In eye tracking experiments, a device is fitted onto the participant’s head that shines light into the retina and measures the light’s reflection. The eye-tracking device allows researchers to record eye movements and the time spent in specific regions of a sentence. Since readers naturally look back to previous sections of a sentence, these movements can be tracked as well. The assumption is that the longer the participant spends on a particular area of a sentence, and the more regressive eye movements to previous sections, the more cognitively demanding those sections are to process. In addition to the inability to measure effects of prosody, another drawback of using eye tracking to deduce processing load from reading times is that eye
movements might very well precede the time it takes to process the information (Ehrlich & Rayner, 1983) or the subject’s gaze might gloss over function words such as pronouns, which might be the focus of a given experiment (Ehrlich, 1983). As a result, it is unclear whether delayed differences in reading times are due to later processing mechanisms or asynchrony between the eye and brain.

In another variation of eye tracking, researchers measure semantic processing through the use of pictures. Subjects fitted with the eye-tracking device listen to sentences or stories while looking at drawn figures or pictures. The more time participants spend gazing at a specific figure in the picture, the more salient that figure is to the content of the information being processed by the participant. This tendency was observed by Cooper (1974) who found that when an item was mentioned in a spoken narrative, participants fixated on that item more often then control items unrelated to the story. Fortunately, auditory stimuli can be presented with this variation, and as such it has been used to understand anaphora and pronoun referent resolution (Dahan, Tanenhaus & Chambers, 2002; Vendetti, Stone, Nanda & Tepper, 2002). The more a subject looks at a particular figure while hearing a pronoun, the more likely the subject interprets that entity as being the referent of the pronoun.

Probe Verification

In probe verification studies, participants are presented with sentences or stories either visually or auditorily. At a predetermined point during each sentence or story, a word appears on a computer screen. The subject must press a button indicating whether he
or she had encountered the word previously in the current sentence or story. Subjects’ reaction times are recorded and a comparison between pairs of minimally-different sentences is made. Instead of processing load, the reaction times reflect the relative salience of given entities. When considering pronoun processing, the techniques measure the relative prominence of potential pronoun referents. A drawback of this approach is that participants must rely on their memories to perform the task as opposed to simply comprehending the sentences. They also must make conscious decisions about the presence or absence of particular entities, which may taint the normal, unconscious processing mechanisms of interest.

Cross Modal Lexical Priming

In Cross Modal Lexical Priming (CMLP) (Swinney, Onifer, Prather & Hirshkowitz, 1979), participants are presented with sentences or stories auditorily. At a predetermined point in a sentence, the listener sees a string of letters (or probe) appear on a computer screen and presses a button as quickly as possible indicating whether that letter-string forms a real word or not (often referred to as a lexical decision). Generally, half of the trials present non-words and half real words. To ensure participants are paying attention to the material presented auditorily and not just focusing on the visually-presented letter strings, comprehension questions are administered randomly throughout the trial.

As participants are hearing sentences, they are tacitly undergoing sentence processing, thereby activating concepts and entities presented in the materials. During
some of the sentences, the word that flashes up on the computer screen is actually a semantic associate of a target word in the sentence. A facilitation in processing is observed when a probe is responded to faster when preceded by semantically related material. Because of semantic priming (Neely, 1991), participants are faster to decide that the semantic associate is a word compared to an unrelated control word (encountered in the same sentence, but by another group of subjects). Control words are carefully selected to match the related words with regard to number of letters and frequency, and they elicit similar average reaction time responses as the related words when presented in isolation, without the auditory stimuli. The priming effect provides us with clues about what entities are activated at a given moment during language comprehension. With regard to pronoun processing, CMLP has been used to determine which potential referents of pronouns are activated, and when they are activated. Specifically, if a semantic associate of a potential referent shows priming at the pronoun, this referent is considered to be ‘activated’ at the pronoun. For example, consider the following sentence:

(22) The cowboy pushed the robber into the chairs by the bar, and the waiter pushed him * into the poker table by the staircase.

The word bank is a semantic associate of the pronoun’s antecedent robber. Thus, if bank appeared as a probe word immediately after him was heard in the sentence, it should elicit a response faster than the unrelated word, rice.
Additionally, the timing of the probe presentation during the sentence can be manipulated. In order to see if the referent had been ‘reactivated’ at the pronoun, a position several milliseconds before the pronoun is usually tested as a baseline position. If priming is not observed at the baseline position, but is seen at the pronoun, then researchers deduce that the referent has been reactivated. Probe points several hundred milliseconds after the pronoun (often referred to as a downstream position) tap later stages of processing.

A variation of Cross Modal Lexical Priming is a Cross Modal Naming. Here, participants are asked to read probe words aloud that flash up on the computer screen. Reaction times of how fast participants start to say the words are recorded. Due to semantic priming, naming should be faster for probes related to a word just heard or activated in the auditorily presented sentence, as compared to an unrelated probe. As with CMLP, priming effects are measured by comparing the difference in reaction time of the related versus control probes.

The benefit of Cross Modal Lexical Priming and Naming paradigms is that the auditory sentences are never artificially partitioned, modified or stopped as a consequence of presenting the probe words. Also the decisions made about the visual probe words never involve metalinguistic judgments as probe verification tasks do. In this way, the unconscious comprehension processes are not disrupted by the task. In addition, probe words are presented during the presentation of the sentences, thereby avoiding end-of-sentence integration effects.
Event-Related Potentials (ERPs)

The billions of neurons in the brain communicate with each other using electrochemical signals. This electrical activity can be recorded with an electroencephalogram, or EEG, in which electrodes that are placed on different areas of the scalp. Although the signals are extremely small, they can be amplified by using specialized equipment. During a psycholinguistic experiment, a participant is exposed to stimuli (usually presented aurally) causing the neurons in the brain to respond to these ‘events’. When enough neurons react in concert, the electrical activity registers on the scalp and is recorded as an event-related potential (ERP).

Researchers have found patterns of electrical changes in response to controlled stimuli and are able to use them to make inferences about the time-course of processing (and to a lesser extend its localization). For example, a positive change at about 500 to 600 ms after an event is called a P600. This effect is consistently observed when participants hear an anomaly of syntactic structure. In contrast, a negative deflection called an N400 is associated with semantic anomalies.

The benefit of using ERPs is that this technique is completely non-invasive, and the measurements do not interfere with the participant’s unconscious language processing. The disadvantage is that it is difficult to draw conclusions about the processing mechanisms underlying the observed reactions to anomalies.

Although there are many other methods used in the field, the five techniques listed above are some of the most common in probing the unconscious processes
underlying language comprehension. Using these paradigms, researchers have
discovered much about the real-time resolution of pronoun referent assignment. The
following section summarizes some of these findings.

On-Line Processing of Pronouns

Although prosody has never played a significant role in the investigation of
pronoun coreference, much is already known about the activation of the pronoun’s
antecedent as it occurs in real time.

Referents Outside the Local Discourse and Recency

First, when considering which entities might be a potential referent, listeners
find comprehension much easier when the antecedent is part of the local discourse. As
evidence of this, Badecker & Straub (1992) found a significant slowing of reading
times when an entity not mentioned in the discourse was the referent of a pronoun.
Additionally, the further the antecedent is in the discourse, the more taxing the task is for native listeners. When the antecedent is separated from the pronoun with
intervening boundary clauses, researchers have found a significant slowing of reading
times (Clark & Sengul, 1979; Ehrlich & Rayner, 1983). These results suggest that
native speakers first look to the local discourse for a viable referent. Only when they
do not find one do they extend their search to other, more obscure entities.
Number, Gender and Animacy Agreement

Number

As discussed in the previous chapter, some basic constraints on coreference are number, gender and animacy agreement. Given this, psycholinguistics have asked the question of whether these constraints occur immediately upon hearing the pronoun or whether other potential antecedents that violate agreement constraints are entertained for a few hundred milliseconds until a later phase of processing when these ungrammatical referents are shed away.

Nicol (1988) used the Cross Modal Lexical Priming technique to investigate the on-line patterns of number agreement as it relates to pronoun processing. Sentences were presented to subjects aurally; immediately after the subject heard the pronoun, the subjects had to make a lexical decision about a visually displayed probe word that appeared on the computer screen. Reaction times to words semantically related to a possible antecedent were compared to unrelated control words. She used sentences such as the following, in which one of two potential referents was plural vs. singular:

(23) The boxers told the skier that the doctor for the team would blame him for the injury.

Nicol found that with a singular pronoun, she observed priming only for the singular referent. However, when she used a plural pronoun such as *them*, she found
reactivation only for the plural antecedent. Her study provides strong evidence for activation of only referents that match the pronoun in number.

**Gender**

Numerous studies have used gender to force a specific antecedent to be the referent of a pronoun. These results have been compared to ambiguous referents which share the same gender as the pronoun. In reading time studies, the sentences with an ambiguous pronoun are read slower than those with an unambiguous pronoun differentiated by gender (Badecker & Straub, 1994; Ehrlich, 1980; Garrod, Freudenthal & Boyle, 1994; Matthews & Chodorow, 1988). Unfortunately, not all of these studies track word-by-word reading times and therefore do not tap into language processing as it is happening in real time.

Nicol (1988) investigated gender congruence for real-time antecedent activation. Using the Cross-Modal Priming paradigm, she compared reactivation of potential referents that agreed or did not agree with the gender of the subsequent pronoun, as in the following:

(24) The ballerina told the skier that the doctor for the team would blame him for the recent injury.

Materials were counterbalanced such that the congruent referent (in this case, *skier*), appeared in subject position for half the sentences and object position for the other
Nicol found that it was only the referent that agreed in gender that was reactivated at the pronoun (*skier*). Likewise, when she inserted the female pronoun, *her*, in place of the male pronoun *him*, she found priming for both *ballerina* and *skier*, since neither clashed with the gender of the pronoun.

An effect of gender was also found in a study by McDonald & MacWhinney (1995). In several probe verification experiments, they reported statistically significant slowing when sentences contained two potential referents with the same gender compared to sentences with characters of different gender. Although the authors downplay the results, they do point out in the discussion of the last experiment that both NPs were activated at the pronoun when they both agreed with the pronoun in gender. This was not the case when the pronoun was congruent with only one of the possible antecedents.

Unfortunately, some counterevidence exists in the literature. Based on findings from several probe verification studies, MacDonald & MacWhinney (1990) claimed that the gender effect is delayed. On the one hand, they stated that gender information was not available until 500 ms after the pronoun because this was when facilitation of the referent and inhibition of the non-referent occurred. The pattern of the results for the ambiguous and non-ambiguous sentences at 0 ms after the pronoun were almost identical. (Although the pattern was similar, the data clearly show a difference between the non-referent and referent in the unambiguous condition; however, this statistical comparison was not reported in the paper.) On the other hand, they described a significant slowing in the sentences with an ambiguous pronoun
because of increased difficulty of finding a referent without the aid of gender cues. This latter observation begs the question of whether gender is truly a delayed effect. Moreover, the findings are increasingly difficult to reconcile since the results of the unambiguous condition reported in the article are completely different in the two experiments even though the same materials and methodology were used. In Experiment 1, MacDonald and MacWhinney actually found inhibition of the correct antecedent at 0 ms after the pronoun – a finding that no other studies have corroborated.

Other researchers have also found that the gender effect is delayed rather than occurring directly after the pronoun (Gernsbacher, 1989; Stevenson & Vitkovitch, 1986). In an attempt to clarify the disparate results in the literature, Garnham, Oakhill and Cruttenden (1992) manipulated the design of their experiments to understand its effects on performance. Specifically, they varied the frequency of asking comprehension questions about who the pronoun was referring to in the stimuli. They found that when comprehension questions followed each item, the gender effect was immediate, but when the comprehension questions were interspersed, the effect was not obtained.

In order to address these discrepancies in the literature, Arnold, Eisenband, Brown-Schmidt, and Trueswell, (2000) used eyetracking to investigate the real-time effects of gender. They used the eyetracking technique in which participants look at a picture of characters and listen to a short discourse about them as in the following:
(25) Donald is bringing some mail to Mickey/Minnie while a violent storm is beginning. He’s carrying an umbrella…

They found that 200 ms after the offset of the pronoun, participants used both gender and advantage for the first-mentioned entity to disambiguate the pronoun. In conditions where the pronoun was not ambiguous because of gender cues, the participants fixated on the correct referent immediately. Also, when both referents were of the same gender, but the referent was the first-mentioned NP, subjects also looked at the correct referent almost immediately after the pronoun. This was not the case, however, when the referent was ambiguous and the context of the story (the one holding the umbrella) implicated the second NP as the referent. In these conditions, the eye fixations were split between the two potential referents. Even 1000 ms after the pronoun, the split between the two referents continued to be 60 percent for the referent to 40 percent for the non-referent.

Taken together, these studies suggest that gender is a cue that can be exploited in the early stages of anaphoric resolution. Since priming effects and processing load (as reflected in reading time and probe verification latency) both point to the activation of multiple potential referents when both match the gender of the pronoun as compared to when only one agrees in gender, it seems that gender does assist in early processing of pronoun referent resolution. The fact that some experiments have not captured this effect suggests that not all methods or experimental designs measure the same thing (as underscored by Garnham et al.’s study). Such discrepancies between
findings offered by different methods will continue to arise in the following discussion of other on-line results.

Animacy

Pronoun agreement entails number, gender, but also animacy congruence. The animacy of a pronoun supplies the listeners with important clues about whether the referent is an animate human or an inanimate object such as a chair.

Shillcock (1982) investigated the role of animacy agreement and its constraints on activation of potential pronoun referents. The following sentences were presented to participants:

(26) The teacher did not board the train, for the simple reason that he was not going to the South Coast of England.

Comparing reaction times for related versus control words, Shillcock found that teacher was reactivated at the pronoun he; whereas in sentences with the inanimate pronoun it, in place of he, no priming effect was observed.

However, Marslen-Wilson & Tyler (1980) found conflicting results in their Cross-Modal Priming experiment. One important aspect of Marslen-Wilson & Tyler’s experiment was that their materials compared human vs. non-human (but animate) referents such as sailor and cat respectively. Since they did not find differential facilitation for semantic associates of the referents after the pronouns he and it, they
concluded that a featural property such as animacy does not constrain coreference immediately. A problem with this conclusion is that speakers often refer to non-human animate entities such as cats as *he* or *she*. Thus, the experiment does not tease apart the human/non-human distinction from the animate/non-animate one. Shillcock’s results still hold, then, demonstrating that only antecedents that agree with the pronoun with regard to animacy are accessed at the pronoun.

In sum, the research on agreement illustrates the important role that these features have on immediately constraining activation of potential referents upon hearing the pronoun. The on-line research indicates that the morphology of the pronoun limits the accessibility of entities mentioned in the discourse at an early stage of processing.

**Binding Constraints**

If features of morphological agreement can constrain pronoun referent activation, the question arises as to whether the same is true for structural constraints. As mentioned in the previous chapter, there are Chomskian principles limiting the grammatically licensed referents of a pronoun. These principles are supported by the utterances native speakers find grammatical in a language. As some psycholinguistics have shown, these principles influence the range of potential antecedents accessed immediately upon hearing the pronoun.
Research by Nicol (1988) suggests that only syntactically licit referents are activated in real time at the moment the pronoun is heard. Nicol used sentences such as the following:

(27) The boxer told the skier that the doctor for the team would blame him for the recent injury.

Here, the referent of the pronoun *him* is ambiguous between *boxer* and *skier* since both possible antecedents are outside the pronoun’s binding category (Chomsky, 1981). However, if the pronoun were to be coindexed with the noun, *doctor*, this would result in an ungrammatical sentence. Activation of each of these NPs was investigated. Nicol found priming effects for both of the syntactically grammatical antecedents, but not for the NP *doctor*, which would result in ungrammatical coindexation. However, when a reflexive appeared in place of the pronoun in the same sentences, Nicol observed a priming effect for *doctor* and not the other two ungrammatical referents.

These findings demonstrate that all and only syntactically grammatical antecedents become accessed immediately after a listener has heard the pronoun.

Nicol’s results have been corroborated by other research as well (Nicol, 1993; Nicol & Swinney, 1989; Swinney, Ford, Bresnan, Frauenfelder & Nicol, cited in Swinney, Ford, Frauenfelder & Bresnan, 1987; Foder, McKinnon & Swinney, cited in Fodor, 1993).
As with the gender studies, though, counterexamples appear in the literature. The research producing these contradictory results involve the same methodologies that caused opposition in the literature on gender; the CMLP paradigm results in an immediate effect whereas probe verification uncovers a delayed effect. Specifically, Badecker & Straub (1999) used probe verification to investigate pronoun activation levels and how they relate to the binding principles. They used sentences such as the following:

\[(28) \quad a. \text{ John thought that Ben owed him another chance to solve the problem.} \\
 b. \text{ John thought that Ben owed himself another chance to solve the problem.} \\
 c. \text{ John thought that Ben owed Jim another chance to solve the problem.} \]

By comparing participants’ response times to a visual probe (John or Ben) in sentence 28a and the control sentence (28c), Badecker & Straub make claims about the level of activation for each of these entities as referents of the pronoun him. Likewise, accessibility of each referent is deduced for the reflexive, himself, when comparing reaction times to sentence (28b) and the control. The results of the probe verification experiments tell a different story than those of Nicol’s. Instead of immediate facilitation of the syntactically licensed antecedent, Badecker & Straub’s experiments
showed faster reaction times for both referents at the pronoun or reflexive, and selective facilitation of the syntactically correct referent 500 ms later.

This discrepancy merits further analysis of the experimental methodology. In probe verification, the control sentence replaces the pronoun or reflexive with something else, and this replacement is usually another entity. As O’Brien, Duffy, and Myers (1986) illustrate, when a new topic is introduced into a discourse, response times to probe words are slowed, not because of facilitation, but because of interference effects introduced by the shift in topic. As further evidence of this, Gernsbacher’s (1989) last experiment demonstrated how newly introduced entities in the discourse actually suppress the other entities. Thus, instead of referent activation at the pronoun/reflexive, the method is more likely detecting interference and suppression at the introduction of the new entity, Jim, in the control sentence. In CMLP, probes are controls instead of sentences, and therefore, no confounding interference effects from new topics are introduced into the design.

Taken together, then, both agreement and binding constraints seems to impact the suite of potential antecedents that are activated – and this activation seems to take place immediately at the point a pronoun or reflexive is encountered in a sentence.

**Gender and Binding Constraints**

Some researchers have gone one step further by experimenting with combinations of agreement and binding constraints. For example, Badecker and Straub (1992, 1994) presented participants with sentences containing two entities, but
only one syntactically licit antecedent. In one variation, the incorrect antecedent was of a different gender from the correct antecedent (incongruent with the pronoun), and in the other variation, the incorrect antecedent was of the same gender, as in the following:

(29)  
    a. Bill thought that John owed himself another opportunity to solve the problem.  
    b. Beth thought that John owed himself another opportunity to solve the problem.  

(30)  
    a. Bob thinks that John will give him a better cut of venison next year.  
    b. Bob thinks that Joan will give him a better cut of venison next year.  

Badecker and Straub measured reading times. Activation of multiple pronouns tends to slow reading times, so a difference between the two sentences suggests that more antecedents are being considered in the sentence that is read slower. In the materials above, the difference in reading times is not expected. The subject of the matrix clause is not coindexed with the reflexive pronoun; it is the subject of the embedded clause that *himself* is referring to. Therefore, only the syntactically correct antecedent (*John*) should be activated and the reading times between the two sentences should be the same. Similarly, the reading times for sentences 30a and 30b should be identical since only *Bob*, the syntactically licensed referent should be activated at the pronoun. However, this is not what Badecker and Straub found. Instead they noted that reading
times for sentences with two entities of the same gender were slower – this means that even though the subject of the matrix clause is ungrammatical as the antecedent for a reflexive, it was considered anyway. Additionally, even though the subject of the embedded clause is illicit as the referent for a pronoun, it too was considered regardless. However, multiple referents were only accessible in cases where the incorrect antecedent agreed with the pronoun or reflexive in gender. Badecker and Straub went on to claim that gender congruent antecedents are initially considered as potential antecedents, and that syntactic information comes into play later by filtering out structurally irrelevant entities.

However, syntactically illicit antecedents do not always seem to be activated, even if they share the same gender as the pronoun. One such example is from an experiment by Badecker & Straub (1997, cited in Nicole and Swinney (forthcoming)) themselves in which the gender-congruent entity was not in subject position, but rather was a modifying possessive, as in the following:

\[(31) \quad \text{a. John thought that Bill’s cousin owed herself another opportunity to solve the problem.} \]
\[\text{b. John thought that Beth’s cousin owed herself another opportunity to solve the problem.} \]

Like the previous set of sentences, there is only one syntactically correct antecedent, cousin. There is also another entity that agrees with the reflexive in gender; however,
this time, the other entity is not the subject, it is the possessive modifying the antecedent cousin. In this study, no difference in reading times was found between the two sentence types. This finding suggests that the activation of Beth as the potential antecedent of herself did not occur; otherwise, these sentences would have been read slower than those with the male counterpart.

Another study by Nicol (1997) corroborates the finding that syntactic constraints influence activation of potential referents just as fast as gender. In Nicol’s study, participants were exposed to sentences such as the following:

(32) a. My aunt / heard that / the congresswoman / would contact / her about / the complaint.

b. My aunt / heard that / the congressman / would contact / her about / the complaint.

(33) a. My aunt / heard that / the congresswoman / would contact / me about / the complaint.

b. My aunt / heard that / the congressman / would contact / me about / the complaint.

The slash indicates phrases that were presented in a self-paced reading task. In these materials, the matrix subject was either congruent with the embedded subject (as in 32a and 33a), or incongruent (as in 32b and 33b). The object was either a third person pronoun or another type of NP such as a first-person pronoun, second-person pronoun
or R-expression. Unlike previous experiments, this study required participants to answer a question after every single item to gauge comprehension. This became an important element of the experiment when considering the results. When all the data were included in the analysis, Nicol showed an interaction between type of object and congruence, in that reading times were fastest when there was the third-person pronoun and the embedded subject was incongruent with the pronoun. Although this finding implies that the syntactically illicit embedded subject was being considered as the referent of the pronoun, the results changed when only the items that were comprehended correctly were included in the analysis. When incorrect interpretations were eliminated, the results did not reveal an interaction for reading times. This observation implies that only grammatical referents are accessed. The confusion for participants could have easily arisen from the fact that both of the other same-gendered NPs were subjects, leading to some extra computation of what the limits of the binding categories in the sentence were. The added processing load could have encouraged some struggling listeners to rely on the cue of gender when it was available. It is not unreasonable to think that the information encoded in the anaphor’s form would be more quickly available. What becomes apparent is that these factors as well as others including prosody are not always applied to the same extent by all native speakers of a language.
Centering Theory and the Advantage of the First-Mentioned Participant

Sometimes agreement and binding constraints are not enough to limit the selection of potential antecedents. In these cases, the discourse plays an important role in helping to narrow down the referent of a pronoun.

In Brennan et al. (1987)’s interpretation of Centering Theory, it is the syntactic position in the previous sentence of a discourse (or clause of a sentence (Kameyama, 1998)) that elevates the salience of potential antecedents, where the subject position is the most salient, object position next most salient, and so on. Numerous studies have shown a processing preference for the subject as the antecedent in psycholinguistic studies.

Using a self-paced reading paradigm, Matthews and Chodorow (1988) investigated sentences such as the following:

(34) a. After the bartender served the woman, he got a big tip.
    b. After the bartender served the woman, she left a big tip.

(35) a. After the bartender served the patron, he got a big tip.
    b. After the bartender served the patron, he left a big tip.

End of sentence comprehension time was marginally faster when the referent was the subject of the fronted embedded clause as opposed to the object. When breaking down the results by each sentence type, the effect of grammatical position did not arise in the unambiguous gender condition at all, but played a very significant
role in the gender ambiguous sentences. Namely, when the pragmatics of the sentence forced coreference with the second NP, overall reading times were greatly slowed.

General slowing was also observed when antecedents were grammatical positions embedded more deeply in the sentence such as a possessive or object of a prepositional phrase (e.g., Because the owner of the restaurant prepared the food, it was always crowded.) Together these studies suggest that preference for an entity in a more salient grammatical role facilitates processing but does not immediately constrain the activation of other potential referents.

Gernsbacker & Hargreaves (1988) showed a striking preference for the subject of the sentence using a visual probe verification task. Word-by-word on a computer screen, participants were presented with sentences such as the following:

(36) Tina beat Lisa in the state tennis match.

Gernsbacker & Hargreaves noted that subjects were faster to verify that Tina had appeared previously in the sentence than they were to verify Lisa. Gernsbacker (1989) corroborated the advantage of the first-mentioned entity when pronouns were incorporated into the dialog.

Interestingly, Gernsbacker & Hargreaves’ (1988) sixth and seventh experiments showed that it is order of mention rather than grammatical function that gives the entity an advantage. They changed the experimental sentences such that the
first-mentioned entity appeared in a fronted clause and the second-mentioned participant was the subject:

(37) Because of Tina, Lisa was evicted from the apartment.

They observed the same facilitation of the first-mentioned entity, as opposed to the sentence’s syntactic subject. From a cognitive standpoint, then, it is the first-mentioned entity that is the most salient as opposed to the subject of the sentence.

The preference for the first-mentioned entity has been pitted against other factors in an attempt to understand its relative strength and general behavior. For example, Badecker and Straub (1994) looked at how gender affects reading time when syntactic information does not constrain pronoun referent resolution. They used sentences in which gender forced a shift in the Center such that the object was the referent of the pronoun, as opposed to the subject.

(38) a. Kenny assured Lucy that he was prepared for the new job.
    b. Julie assured Adam that he was prepared for the new job.

They found that sentences forcing the object to be the pronoun’s referent were read slower than those that forced coreference with the subject. Not surprisingly, other studies including research by Ehrlich & Rayner (1983), Garrod et al. (1994), and
Hudson, Tanenhaus & Dell (1986) also found longer latencies when the context forced a rough shift in Center.

Badecker and Straub also manipulated the gender of the potential referents such that both agreed with the pronoun:

\[(39) \quad \text{Kenny assured Adam that he was prepared for the new job.}\]

In this study, slower reading times were measured after the pronoun in the sentences with an ambiguous referent compared to those in which gender selected for the subject of the previous clause. These results suggest that gender has an immediate effect on pronoun resolution, but also that the discourse context does not suppress activation of potential referents. If this were true, then the subject, Kenny, would be the only available referent and the two sentences would have been read equally fast.

Recall, also, Arnold et al.’s (2000) eye-tracking experiment in which the pronoun referred to two potential referents. When the picture was consistent with the first NP, participants fixated on that referent almost immediately. However, when the picture indicated the second NP, subjects looked at both the first and second entities equally as often. It seems that when saliency converged with gender cues, the selection of the antecedent becomes committed early-on.
Pragmatics and Theta Roles

From the evidence so far, it seems that some factors such as agreement features and binding constraints limit the suite of potential referents activated for a given pronoun; whereas other factors, namely, order of mention do not (although they may influence the ultimate interpretation). When considering pragmatics and theta roles, the same question of when these effects take place during processing has been asked by psycholinguists.

At first glance, it appears that pragmatics plays a critical role early-on. Hirst and Brill (1980) found that when participants were exposed to a lead-in sentence with a contextual bias for an antecedent, reaction times for determining the antecedent (for an unambiguous pronoun) were faster depending on the strength of the bias.

However, other evidence has suggested that when a referential ambiguity is present, pragmatics does not seem to play such an important role immediately. In an experiment looking at pragmatics, Ehrlich (1980) used gender to disambiguate the referent of a pronoun in half the target materials as in the following:

(40)  a. Steven blamed Jane because she spilt the coffee.
     b. Steven blamed Frank because he spilt the coffee.

She found that participants were faster to decide that Jane was the antecedent in 40a versus Frank in 40b. From this she concluded that in an immediate processing
phase, only the gender of the pronoun is used; pragmatic information is summoned only in a later stage if the assumed antecedent is infelicitous.

The problem with this study is that the pragmatic information disambiguating the referents does not become available until after the pronoun, and the measurement is of reading an entire sentence. To reconcile these problems, Gernsbacher (1989) created materials such that two potential referents were introduced in one sentence, and then pragmatic information was revealed in a second sentence that resolved the ambiguity of the subject pronoun as in the following discourse:

(41)  

(41)  

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)

(41)
pragmatics of the target sentence. For example, they used materials such as the following:

(42) *Flying to America.* Joan wasn’t enjoying the flight at all. The dry air in the plane made her really thirsty. Just as she was about to call her/him, she noticed the stewardess/steward coming down the aisle with the drinks trolley.

(43) a. Right away, she ordered a large glass of Coke. (ambiguous)
b. Right away, she poured a large glass of Coke. (ambiguous)
c. Right away, she ordered a large glass of Coke.
d. Right away, she poured a large glass of Coke.
e. Right away, he ordered a large glass of Coke.
f. Right away, he poured a large glass of Coke.

(44) Joan finished it in one go and ordered another one.

The continuations in which Joan is pouring a large glass of Coke violates the schema of flights in which stewards and stewardesses serve passengers drinks. Likewise, those with the stewardess or steward ordering a Coke do the same. Garrod et al. (1994) found that, overall, participants read sentences without pragmatic conflict faster than those with it. When the pragmatics did not make sense (for example, when the steward was ordering a Coke), the reading times slowed. Critically, Garrod et al. analyzed first-pass eye fixations in addition to overall reading times to assess the
immediacy of the effect of pragmatics. What they found was that the reading times were the same in the plausible and implausible cases, in all but one situation. Pragmatics did reveal an immediate effect when the pronoun was not ambiguous and the gender was consistent with the most salient antecedent in the local discourse (sentence 43c compared to 43d). From this result, it appears that when the antecedent was unambiguous (from cues of gender agreement and discourse focus), pragmatic effects are immediate. Thus, an important distinction to make is the process of actually choosing a referent versus integrating a committed referent into the context of the sentence.

Theta roles often enter the picture when considering pragmatic information. To help precisely predict how semantics of a discourse changes interpretations, researchers have varied verbs with different but complementary thematic structures. McDonald and MacWhinney (1995) conducted a series of on-line experiments to study the effect of theta roles in the real time resolution of pronoun antecedents. They claim that theta roles have an immediate effect on pronoun resolution. As stimulus materials, they used both Experiencer/Stimulus and Stimulus/Experiencer verbs as in the following:

(45)  
\begin{itemize}
  \item a. Gary \textit{admired} Alan time after time because he was so talented.
  \item b. Gary \textit{amazed} Alan time after time because he was so talented.
\end{itemize}
Probe words appeared at different points during the sentence and subjects’ response times were measured. The results showed that the first-mentioned referent was accessible at most points during the sentence. Reaction times to Gary were consistent in both sentences; however, the Stimulus referent, Alan was equally accessible in sentences like 45a at the pronoun and at the end of the sentence. McDonald and MacWhinney’s experiment suggests that implicit causality is something that is used in pronoun activation immediately upon hearing the pronoun. However, instead of disambiguating the referent, their results show that both referents are accessed at the pronoun for Experiencer-Stimulus verbs. In this study it is difficult to tease apart how much gender affected the results, since for half of the materials, the gender was different between the two referents. It would not be surprising if when a participant heard the sentence: “Gary amazed Ellen time after time because he …” that the participant responded faster to Gary than Ellen, not because of the verb, but because of gender agreement. If the findings were collapsed in the same way as in the Gernsbacher paper, the resulting picture would be similar: both antecedents active at the pronoun, and the contextually irrelevant entity suppressed downstream. Given this, it seems that pragmatics (including theta roles) do not constrain potential referents, but rather help resolution at a later stage of processing.

Parallelism

Most research dealing with parallelism has focused on off-line interpretations. Many have used reading times to demonstrate that the effect exists, but unfortunately,
many of these studies analyzed latencies for entire sentences instead of reporting word-by-word measurements. For example, Chambers & Smyth (1998) showed that overall reading times of target sentences slowed when gender forced the referent of a pronoun to be in a nonparallel position. They used discourses such as the following:

(46) A group of celebrities were having car trouble. Martin Miles told Liz Lovejoy to check the oil.

(47) a. Then he told Dean Morgan to inspect the coolant.
    b. Then Dean Morgan told her to inspect the coolant.
    c. Then she told Dean Morgan to inspect the coolant.
    d. Then Dean Morgan told him to inspect the coolant.

(48) The engine had overheated.

Referents that were impelled by gender but that did not match the grammatical function of the referent in the previous sentence (47c and 47d) resulted in longer reading times. In a similar vein, increases in reading times have been reported for sentences in which an antecedent was the grammatical object of a previous sentence as compared to reading times when the antecedent was the subject (Hudson et al., 1986).

In a truly on-line experiment, Frisch (2003) used event-related potentials (ERPs) and found that parallel grammatical function indeed has on-line implications for pronoun resolution in German. When there is grammatical function parallelism with the subject but the actual antecedent referred to the object, participants showed a
P600 reanalysis effect (the same effect observed for syntactic anomaly). Likewise, when the object was parallel but the actual referent was the subject, participants again showed a P600 effect. This result is a bit surprising given that grammatical function in the context of the advantage for the first-mentioned and centering theory experiments showed delayed effects. Again, one difference is the function – on the one hand there is selecting a referent when there is more than one potential antecedent; and on the other, there is integrating a committed antecedent into the interpretation of the sentence. In the experiment above, the referent was selected, but this antecedent could not be integrated seamlessly. The next section offers clues that it is the overarching coherence relation that does influence pronoun referent resolution immediately.

Coherence Relations

In a direct investigation of the role of coherence relations on pronoun processing, Wolf, Gibson & Desmet (2003) used a self-paced reading paradigm to measure differences between materials with Cause-Effect versus Resemblance coherence relations. The materials were minimally different with regard to the syntax. In the Cause-Effect materials, the phrase “and so” joined the two clauses; whereas in the Resemblance, or parallel case, the phrase “and similarly” intervened between the two clauses. (The verbs in the first clause were also different to make sense with the coherence relation.):
(49)  a. Fiona complimented Craig and similarly James congratulated him, so they ...
       b. Fiona complimented Craig and similarly James congratulated her, so they ...

(50)  a. Fiona defeated Craig and so James congratulated him, so they …
       b. Fiona defeated Craig and so James congratulated her, so they …

The difference in gender forces the interpretation to follow the subject or object of the previous clause. Wolf et al. found that in the sentences with Resemblance, the reaction times at the pronoun and following word were faster when the pronoun was consistent with the object referent in gender; and in Cause-Effect sentences, reaction times were faster when the pronoun was consistent with the subject. The findings suggest that different coherence relations influence listener’s strategies for assigning a referent to the pronoun.

Reflecting back on all the studies discussed up to this point, it seems that there are certain factors that have immediate effects and others that are delayed. First, listeners access appropriate referents that exist within the local discourse. This makes sense given humans’ limited memory capacity. Second, agreement factors including number, gender, and animacy all constrain the suite of possible antecedents that a pronoun can refer to and these factors do so almost immediately upon encountering the pronoun. Binding constraints can also impact activation of referents immediately.
Although the advantage for the first-mentioned participant does not prevent an antecedent from being activated at the pronoun, it can help the system to converge on a single referent in an early stage of processing when it is consistent with other factors such as gender agreement. In contrast, pragmatic information seems to help disambiguation in a later stage of processing. Surprisingly, parallel function (or rather coherence relations in the Resemblance category) behaves more like a quick syntactic reflex like a binding constraint, rather than the slower-acting pragmatic information.

Different factors contribute to pronoun disambiguation at different times, and this lends itself to a model of discourse comprehension that embodies two phases of processing.

**Construction-Integration Model**

One framework that models discourse processing and gives specific attention to anaphora is Kintsch’s Construction-Integration Model (Kintsch 1988). In this model, there are two phases, Construction and Integration. In the first phase, Construction, initial processing takes place. It is bottom up and generates candidate elements such as word meanings, propositions and inferences without regard to the discourse context. In fact, where appropriate, it generates multiple candidates with the correct one being a member. Then, these generated elements are fed into the second phase, Integration. Integration strengthens contextually relevant elements through a spreading activation process and creates a coherent, integrated interpretation.
In this model, general knowledge is conceptualized as an associative network, where nodes are concepts or propositions with different activation levels. As an utterance unfolds, a word identification process and parser extract concepts and propositions directly corresponding to the input. This begins the Construction phase. The location of these concepts and propositions are identified within the general knowledge network and copied to a memory buffer (such as one used for verbal short-term memory). A small number of closely associated neighbors of these propositions are selected and copied from the general knowledge net as well. The algorithm for retrieving these associated nodes is based on probabilities calculated from the strength of the connection between the input node and the associated node. Then an inference mechanism generates additional connections given general understanding of the world. Finally, activation levels and connection weights between the different propositions are assigned – these are inherited directly from the general knowledge network.

Integration is based on connectionist processes, which have been used to exclude unwanted elements from a representation (Rumelhart & McClelland, 1986). The organization is assumed to be in cycles corresponding to short sentences or phrases. In each cycle, a new network is constructed including the networks stored in the short-term buffer from the previous cycle. During the Integration process, activation spreads around until the system stabilizes. The process is relatively rapid, but if the Integration process fails to provide a coherent interpretation, new constructions are added to the net and integration is attempted again. The output of
the Integration phase is an activation vector with high activation values for some of
the nodes. These nodes with the highest values are the final interpretation or discourse
representation.

To demonstrate a simulation with anaphora disambiguation, Kintsch used an
example from Garrod et al. (1994)’s study. The materials used are discourses that
present several characters and a final sentence using a pronoun:

(51) *Flying to America.* Jane[sic] wasn’t enjoying the flight at all. The dry
air in the plane made her really thirsty. Just as she was about to call her,
she noticed the stewardess coming down the aisle with the drinks
trolley.

(52) a. Right away, she ordered a large glass of Coke.

b. Right away, she poured a large glass of Coke.

Of interest is the time course of how readers interpret *she* in the continuation sentences
in 52. The following is a simplified diagram of the network used in the simulation for

*Right away she ordered a large glass of Coke:*
The filled diamonds represent the nodes constructed during reading, which are in long-term memory. The open diamond (order) is the test node. There are two possible interpretations of Continuation 1. Given that the antecedent is in the immediate discourse, either Jane ordered a large glass of Coke or the stewardess did. These two possibilities are represented by the test cases (open circles). These nodes are linked to one another with connections varying in strength. For this example, all lines are equal to +1, except for the dotted line connecting the two circles which equals −1.

According to the process defined above, the propositions mentioned in the text input are identified in the general knowledge network, and associative nodes are copied into a memory buffer (the network diagram above shows the simplified entry into the memory buffer). Then the inference mechanism creates additional connections based on general understanding of the world. In the example above, the
line from the test case order(Jane,glass) is connected to the node order because logically Jane is a passenger and passengers order drinks. The connection is not drawn for the test case order(stewardess,glass) since stewardesses do not typically order drinks on flights. The final step in the Construction phase is to assign activation levels and weights to each of the connections. For this example, the two antecedent alternatives are assigned 0 for their activation levels and have a link of –1 between them (since only one antecedent will be selected). All the other propositions inherit their activation levels from the general knowledge network and have connection weights of 1.

During the Integration phase, a spreading activation process is applied to the network and over the course of a series of cycles activation levels change until the final interpretation is realized at about 10 cycles. For the sentence with order, the entity most salient in the discourse and most acceptable regarding pragmatics both favor Jane. In the simulation, Jane exhibits strong activation levels early-on and throughout all cycles. In contrast, the stewardess never shows much activation. For the continuation with pour, Jane originally shows high activation levels, but then these trail off as the pragmatics of who generally pours drinks during a flight factor into the calculations, especially during the later cycles. Alternatively, the stewardess starts with an activation level a bit lower, but this gradually increases as the pragmatics become more influential in the later integration cycles.

As demonstrated by this example, the pragmatics dictating which entity is more likely to perform a given action is not integrated until several cycles later. It is
only after this integration occurs that the contextually irrelevant candidates are eliminated from the discourse representation. Thus, the CI Model provides insight into how the subconscious mind undergoes the process of pronoun referent resolution.

Even so, the CI Model only deals with textual input. It contains no mechanism to represent spoken utterances or capture the influence of prosody on interpretation. Prosody does play an important role. Although more work has focused on pronouns, the topic of prosodic processing and its time course has been addressed in psycholinguistics, especially in recent years.

Research on Prosody and Pronoun Processing

Most of the research focusing on prosody and sentence processing has examined local syntactic attachment ambiguity (Ferreira, Anes & Horine, 1996; Marslen-Wilson, Tyler, Warren, Grenier & Lee, 1992; Pynte & Prieur, 1996; Schafer, Carter, Clifton & Frazier, 1996; Speer, Kjelgaard & Dobroth, 1996; Watt & Murray, 1996). The on-line studies have suggested that prosody influences interpretations immediately. However, prosody surfaces at many different levels of a sentence, so this type of prosody is merely one piece of the puzzle. Prosody also influences pronoun interpretation. As previously discussed, off-line preferences indicate that the accent on the pronoun shifts the interpretation of the referent from the entity with parallel grammatical function to the one with incongruent grammatical function. The question is, at what phase of processing does the prosody influence interpretations? Is
it early in the Construction phase as previous experiments on prosody suggest? Or is it in the Integration phase when context suppresses implausible interpretations?

Vendetti and colleagues recently researched the effects of pitch accent on pronoun interpretation during on-line processing. Venditti, Stone, Nanda, & Tepper (2001, 2002) used semantic eye-tracking to understand the on-line processing of pronoun resolution in discourses such as the following (although these were analyzed post hoc since they were included as fillers):

(53) The animals were playing out near the barn when something unexpected happened. The lion started going ballistic. He hit the alligator with a long wooden rake. Then he/HE hit the duck. A big fight ensued and it was a terrible scene.

Although no statistical analyses were performed on the data, it seemed that in the unaccented condition, only the first NP was activated, whereas in the accented condition, the second NP was activated. Additionally, however, the discourses with stressed pronouns showed activation of the first NP for about 40 ms. After this, the activation level trailed off while participants continued to gaze at the second NP. The results suggest that prosodic information is not used immediately. However, conclusions are difficult to reach when no statistical analyses are available. Also, if the effect only lasts for 40 ms, then the question arises as to whether it should really be
considered as part of the Integration phase, which does not take effect until about 400 ms later.

Venditti et al. postulated that the temporary ambiguity might have been the result of an ambiguous coherence relation. If the participant presupposes that the sentence is narrative, then he or she will prefer the first NP; however, if a Resemblance relation is assumed, the participant will most likely choose the NP with parallel grammatical function. Although accent on the pronoun has a strong effect when in a Resemblance coherence relation, it loses its power in narrative discourses where it fails to reliably shift the interpretation of the pronoun. Using an eye-tracking experiment, Venditti, Trueswell, Stone and Nautiyal (2003) tested whether the coherence relation had an effect on pronoun interpretation with contrastive stress. For this study, they used materials with either a Narrative or Parallel coherence relation:

(54) The kangaroo tapped the elephant with a magic wand.

(55) a. Then HE tapped the raccoon.

b. Then HE looked into the crystal ball.

As in the other experiment, participants fixated on both the first and second NPs early on. Then at the offset of the verb in the parallel clause, a statistically significant preference for the second NP emerged. Venditti et al. propose that the verb assists the listener in resolving the coherence relation that aids in the interpretation of the stressed pronoun. Since the different coherence relations lead to divergent results, and since
the verbs always appear after the pronoun in these experiments, it is not clear whether the prosodic information is used immediately or not.

In sum, researchers have discovered many factors that contribute to the referent of a pronoun. The psycholinguistic data indicate that some of these factors, such as gender, might be available immediately; whereas others, such as pragmatics, might be incorporated in a later stage of processing. Unfortunately, different methods often lead to contradictory results, making determinate conclusions difficult. Although models exist of how different factors might contribute to an interpretation at different times during processing, no model to date has captured the important role of prosody. With regard to when prosody interacts with pronoun referent resolution, recent experiments have not been able to say definitively whether prosody is used immediately or not since presentation of the pronoun has always been accompanied by a temporarily ambiguous coherence relation. Thus, more research is needed to truly understand the time course of the prosodic processing and its relation to pronouns.
PITCH ACCENT AS FOCUS

Chapters 1 and 2 defined the many factors that influence coreference including number, gender and animacy agreement, syntactic position, grammatical function, thematic roles, pragmatics, discourse coherence relations and prosody. Up to this point, prosody has been considered only when realized on the pronoun. Another possible utility of pitch accent is to place focus stress on the referent. The aim of Chapter 3 is to explore this broader role of pitch accent in pronoun referent resolution.

Pitch Accent and Focus

Researchers have long associated prosodic accenting with focus (for a review, see Ladd, 1996). In Selkirk’s view (1984) the important element for placing a constituent in focus is the pitch accent. When a pitch accent occurs over a constituent, it is considered to be focus marked. According to Jackendoff (1972), focus is a hypothetical construct that has a phonological correlate (accent), a semantic-pragmatic correlate (the new information) and a syntactic correlate (feature [+F]). Thus a word that is accented is often assumed to introduce new as opposed to given information.

Focus does not necessarily require information to be new, though. Bard and Aylett (1999) reported that in the spontaneous dialogs they analyzed, entities mentioned previously in the discourse were not consistently deaccented in spontaneous dialogs. Cohan (2000) discovered that in corpus data from natural speech, focus seemed to flag alternatives inherent in the discourse or implied by the
speaker. Pierrehumbert & Hirschberg (1990) claim that pitch accent offers
information about the status of a referent and how the speaker ranks its relative
salience in the discourse. Thus, it seems that pitch accent does not require the
constituent to be new, but rather marks it as being more salient than other entities. By
extension, then, pitch accent on a potential pronoun referent may bring it into focus
and make it more salient as the chosen referent.

The notion that pitch accent affects pronoun interpretation by making a
referent more salient has existed in previous work, but has never been explicitly tested.
For example, in her dissertation on discourse, Arnold (1998) included a table listing
factors that could influence reference form, including intonational focus on the
referent, but she never discussed or developed this topic further. Schafer et al. (1996)
reported experimental results suggesting that pitch accent does indeed change
sentence-level interpretations. Although they did not investigate pronoun coreference
specifically, they found that a pitch accent on one of two nouns modified by a relative
clause affects informants’ interpretation of which noun is being modified. In
particular, the informant is more likely to assume that the relative clause modifies the
noun that is stressed. Maynell (1999) replicated these results on materials controlled
for nuclear pitch accent placement and intonational phrase breaks.

Since pitch accent affects relative clause attachment, it suggests that pitch
accent should also affect pronoun referent assignment. Even with suggestions in the
literature, however, there is no direct evidence to date. Thus, the question of whether
pitch accent on a referent does indeed affect pronoun resolution was addressed in the following series of experiments.

**Experiment 1: Focus Stress in Active Structures**

The purpose of Experiment 1 was to investigate whether pitch accent on potential referents affects pronoun assignment in subsequent utterances. Two hypotheses were pitted against each other. The first was based on Pierrehumbert & Hirschberg’s (1990) claim that accent influences how the speaker ranks a constituent’s relative salience in the discourse. The idea was that prosody serves as a primary force that promotes constituents in the rank-ordered set of potential antecedents when that constituent is focus-marked:

**Focus Stress as a Primary Influence (FSPI) Hypothesis:** When pitch accent is placed on a potential referent of a pronoun, native informants will select the prosodically focus-marked entity to be the referent.

In the second hypothesis, accent functions in the same way, but does not hold the same amount of strength. Pitch accent affects coreference without impelling a preference for one or the other entity. The motivation behind this view was that a speaker needs to indicate an entity’s level of salience in a discourse and does so by using different cues. Some of these cues, such as placing a particular entity in subject position, are powerful indicators of salience and can greatly influence interpretations (Garnham et al., 1996; Gernsbacher & Hargreaves, 1988; Stevenson et al., 1993).
Other cues, such as pitch accent, may not be as powerful indicators of salience. Data from Schafer et al.’s (1996) pitch accent and attachment studies suggest that prosody does not dictate which noun was selected as being the one modified, but rather increases the likelihood that participants would select the stressed noun. Thus, the following hypothesis for pronoun resolution claims that a pitch accent placed on one of two potential referents only increases the likelihood that participants will prefer the one that is accented:

**Focus Stress as a Secondary Influence (FSSI) Hypothesis:** Pitch accent placed on a potential referent of a pronoun increases the likelihood that native informants will select the prosodically focus-marked entity as the referent.

Experiment 1 compared these two hypotheses by using pronouns and potential referents presented in the context of short discourses with intra-sentential anaphora.

**Method**

**Participants**

Participants were 24 volunteers recruited from the community. Eighteen of the participants were college students at UC San Diego, Stanford, and UC Berkeley. The remaining six were working professionals ranging in age from 25 through 54. Fourteen were male and ten were female. All were native English speakers.
Design and Materials

Materials consisted of 32 three-sentence discourses that were recorded and presented aurally. In the 16 experimental discourses, the first sentence introduced the scenario and a main character (“Wendy was taking a cruise to Alaska during her spring break.”). Then, in the second sentence, the main character encounters two other people. The sentence structure is such that the subject of the embedded clause is the Agent of some action, while the object is the Patient (“As she was going to her room to grab a book, she noticed a paramedic calling the captain down the hall.”). For each discourse, two versions of the second sentence were recorded. In one version, a pitch accent was placed on the subject of the embedded clause and in the other, a pitch accent was placed on the object. The third sentence included the main character and a third-person object pronoun that referred to one of the two entities mentioned in the previous sentence (“Later Wendy bumped into him and asked if everything was okay.”). The remaining 16 discourses were fillers with varying sentence structures. Appendix A provides a list of the experimental materials.

Discourses for all three experiments were recorded by a 30-year-old female who had a graduate degree in psycholinguistics and experience recording voice messages for speech applications.

An acoustic analysis was performed on the materials to ensure that the delivery of the second sentence achieved the intended prosodic contour. Using Praat software (Boersma & Weenink, 2002), the maximum pitch accent over each noun was extracted. Although duration and intensity are not directly motivated by
Pierrehumbert’s theory of intonation, these measurements were gathered for completeness since they are factors associated with stress (Ladd, 1996). Table 1 shows the mean scores for each condition, where first NP and second NP refer to the subject and object of the embedded clause, respectively.

Table 1.

Mean pitch, duration and intensity measurements for each type of NP in Experiment 1.

<table>
<thead>
<tr>
<th>NP Type</th>
<th>Pitch (Hz)</th>
<th>Duration (ms)</th>
<th>Intensity (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1\textsuperscript{st} NP Accented</td>
<td>315</td>
<td>783</td>
<td>67</td>
</tr>
<tr>
<td>1\textsuperscript{st} NP Unaccented</td>
<td>252</td>
<td>519</td>
<td>69</td>
</tr>
<tr>
<td>2\textsuperscript{nd} NP Accented</td>
<td>317</td>
<td>749</td>
<td>66</td>
</tr>
<tr>
<td>2\textsuperscript{nd} NP Unaccented</td>
<td>210</td>
<td>472</td>
<td>68</td>
</tr>
</tbody>
</table>

Statistical analysis of the prosodic measurements showed that across sentences all accented first NPs were delivered with significantly higher pitch than all unaccented first NPs ($t_{15} > 10$, $p < .001$) and all accented second NPs had significantly higher pitch than all unaccented second NPs ($t_{15} = 10$, $p < .001$). Likewise, when comparing the NPs that appeared in the same sentence, all accented first NPs were higher in pitch than all unaccented second NPs ($t_{15} = 10$, $p < .001$), and all unaccented first NPs were presented with lower pitch than accented second NPs ($t_{15} = 10$, $p < .001$). For duration, all accented first NPs were longer than unaccented first NPs ($t_{15} = 6.44$, $p < .001$) while accented second NPs had longer duration times than unaccented second NPs ($t_{15} =$
When comparing duration measurements within the sentence, accented first NPs were significantly longer than unaccented second NPs ($t_{15} = 10, p<.001$), and unaccented first NPs were shorter than accented second NPs ($t_{15} = 4.78, p<.001$).

Finally, with regard to intensity, accented first NPs were louder than unaccented first NPs ($t_{15} = 3.01, p=.01$), and likewise for second NPs ($t_{15} = 2.37, p=.03$). However, accented first NPs were not significantly louder than unaccented second NPs ($t_{15} = 1.62, p=.13$), although there was a difference between unaccented first NPs and accented second NPs ($t_{15} = 3.95, p<.001$). The results validate the delivery of the materials in that the NPs that were intended to be accented were higher in pitch and longer in duration. For the most part, accented NPs were louder than unaccented NPs, although this measure was not as strong an indicator of accent as the other two.

The experimental design was within-subjects. Each participant heard eight experimental discourses with pitch accent on the first potential referent (first NP) and eight with a pitch accent on the second (second NP). Two scripts were created to counterbalance the target items. Across scripts, the discourses appeared in the same pseudo-random order. A discourse with the pitch accent realized on the Agent in one script would have the pitch accent on the Patient in the other, and vice versa.

An answer sheet consisting of 32 questions (one per discourse) was created for the experiment. Each question asked to whom the main character did a particular action, for example, “Who did Wendy bump into?” Two of the three choices were the potential referents mentioned in the second sentence of the discourse, and the third

---

7. This is not unusual. As Werner and Keller (1994) note, stress is often characterized by only one or two of the classic three parameters.
was a foil response. The order in which the choices appeared on the answer sheet was
counterbalanced so that the Agent, Patient, and foil appeared in the first, second and
third positions an equal number of times.

Procedure

The experimenter provided the participant with the answer sheet of 32
questions. The experimenter explained that the participant would be listening to short
stories or scenarios of about three sentences each, and that each story would be played
only once. The participant was instructed to answer a question about each story by
circling one of three choices on the answer sheet. The experimenter explained that the
answer would always be one of the choices given and that the participant was to
choose one and only one of the choices. Participants were also encouraged to guess if
they did not know the answer (as opposed to making no selection at all).

The experimenter provided the participant with a portable cassette player and
headphones and started up a tape, which included one of the two scripts. A four-
second pause followed the presentation of each discourse to allow participants enough
time to make a selection. The participant heard all 32 discourses, with half of the
target items having a pitch accent on the Agent and half on the Patient. After the tape
had finished, the experimenter asked the participant what he or she thought the
experiment was about and tried to elicit any strategies the participant might have used
to make decisions.
Results

Answers to the questions were tallied for each participant. The range of correct choices for the 16 filler questions was 14 to 16, with an average of 15.4, thus demonstrating that participants were attending to the task.

In order to measure the effect of pitch accent placement on pronoun resolution, a measure of how many times each participant chose the first NP (subject of the embedded clause), second NP (object of the embedded clause), and foil as the referent was calculated for both conditions. Figure 2 shows the averages for discourses with a pitch accent on the first NP, and discourses with the pitch accent on the second NP.

![Figure 2. Average number of times the first NP, second NP, and foil in an active construction were selected as the pronoun referent. Since there were eight discourses per condition, each subject’s score could range from 0 to 8. The average was calculated across participants.](image-url)
To determine whether there was an effect of pitch accent, two t-tests were performed. The first t-test compared the average selection of the first NP when pitch accent was placed on the first versus the second NP. The t-test revealed a statistically significant difference between pitch accent placement on the first NP (M=6.5) and pitch accent on the second NP (M=5.3) ($t_{23} = 3.91, p < .001$). For the second t-test, the selection of the second NP was compared for both conditions. Since none of the participants chose the foils, the results were the exact inverse of the first NP selection. Participants chose the second NP significantly less often when pitch accent was placed on the first NP (M=1.5) compared to when pitch accent appeared on the second NP (M=2.7) ($t_{23} = 3.91, p < .001$).

Discussion

The results of Experiment 1 showed a significant effect of pitch accent on pronoun referent resolution. Thus, the suggestion in the literature that appoints pitch accent as another factor influencing anaphora resolution is supported by the present findings.

Although, the effect of pitch accent exists, it sits in the background of the more powerful bias of choosing the subject as the pronoun’s referent (in this case, the entity selected is the subject of the embedded clause). Because listeners seemed to look to the subject position to determine which entity was most salient in the discourse more often than pitch accent, the findings of Experiment 1 support the FSSI Hypothesis in
that pitch accent did not impel participants to select the prosodically focus-marked entity. Rather, prosody increased the likelihood that participants would favor the focus-marked entity when the two cues (subject position and pitch accent) marked different entities as being salient. Thus, focus stress serves as a secondary, as opposed to a primary, cue when promoting a constituent’s level of salience in the discourse.

In the sentence construction used in Experiment 1, the first NP not only occupied the subject position of the embedded clause, but also assumed the theta role, Agent. As mentioned before, theta roles play an important role in pronoun referent resolution (Stevenson et al., 1994). In general, speakers tend to prefer Patients to Agents as referents of a pronoun. However, theta roles do not solely dictate the outcome of pronoun resolution. Again, a speaker can use many cues to indicate to a listener that an entity is salient in a discourse. In English, a very powerful cue is to place an entity in subject position. In active sentences, the subject is salient given that the speaker placed this entity in the most prominent grammatical position, but the Patient is also salient since it is the recipient of some action. Unlike a strict hierarchy as Kameyama postulates, then, it seems that the relation between the forces influencing pronoun resolution interact in a more complex way. The purpose of Experiment 2 was to investigate how pitch accent comes into play when considering both subject position and theta roles.
Experiment 2: Focus Stress in Passive Structures

Experiment 2 tested the FSSI Hypothesis and pitch accent’s interaction with other cues of salience, namely, thematic structure. With passive constructions with transitive verbs such as “The cat was chased by the dog,” the Patient is in subject position and the Agent appears in a by-phrase. The speaker is explicitly promoting the Patient to a prominent grammatical position and is emphasizing the salience of this entity. If pitch accent is placed on the subject, then the cues of salience are consistent, and the listener should choose the subject as the referent. If pitch accent appears on the Agent (or the second NP in the embedded clause), then the cues are inconsistent. The FSSI predicts that pitch accent on the Agent will increase the likelihood that listeners will choose the Agent in the by-phrase as the referent, even though the passive construction suggests to the listener that it should be the subject.

Method

Participants

Participants were 24 volunteers recruited from the community. None of the subjects had participated in Experiment 1. The profile of the participants was similar to those in Experiment 1, the majority being college students from UC San Diego and Stanford and the remaining being professionals ranging in age from 27 to 52. All were native English speakers. Twelve were male and twelve were female.
Design and Materials

Materials consisted of the same three-sentence discourses from Experiment 1, except that in the 16 experimental discourses, the second sentence was structured in the form of a passive instead of an active sentence. Thus, the subject of the embedded clause was the Patient of the action, while the second NP appearing the by-phrase was the Agent (“As she was going to her room to grab a book, she noticed a paramedic being called by the captain down the hall.”). As with Experiment 1, two versions of the second sentence were recorded. In one version, a pitch accent was placed on the first NP (Patient) and in the other, a pitch accent was placed on the second NP (Agent).

To ensure the integrity of the prosodic delivery of the materials, the audio recordings underwent the same acoustic analysis as in Experiment 1. Table 2 shows the mean scores of pitch, duration and intensity for each condition.

Table 2.
Mean pitch, duration and intensity measurements for each type of NP in Experiment 2.

<table>
<thead>
<tr>
<th>NP Type</th>
<th>Pitch (Hz)</th>
<th>Duration (ms)</th>
<th>Intensity (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st NP Accented</td>
<td>325</td>
<td>764</td>
<td>72</td>
</tr>
<tr>
<td>1st NP Unaccented</td>
<td>287</td>
<td>580</td>
<td>72</td>
</tr>
<tr>
<td>2nd NP Accented</td>
<td>325</td>
<td>704</td>
<td>70</td>
</tr>
<tr>
<td>2nd NP Unaccented</td>
<td>239</td>
<td>552</td>
<td>71</td>
</tr>
</tbody>
</table>
As with Experiment 1, statistical analyses found a significantly higher pitch measurement for accented first NPs compared to unaccented first NPs ($t_{15} = 3.70$, $p<.001$), with the same pattern for accented and unaccented second NPs ($t_{15} = 7.00$, $p<.001$). Likewise, when comparing the NPs that appeared in the same sentence, accented first NPs were delivered with higher pitch than unaccented second NPs ($t_{15} = 8.34$, $p<.001$), and unaccented first NPs were presented with lower pitch than accented second NPs ($t_{15} = 3.09$, $p=.01$). For duration, accented first NPs were longer than unaccented first NPs ($t_{15} = 3.44$, $p<.001$) while accented second NPs had longer duration times than unaccented second NPs ($t_{15} = 3.35$, $p<.001$). When comparing duration measurements within the sentence, accented first NPs were significantly longer than unaccented second NPs ($t_{15} = 5.70$, $p<.001$), but the duration of unaccented first NPs versus accented second NPs only showed a trend ($t_{15} = 2.08$, $p=.06$). Finally, with regard to intensity, none of the comparisons revealed statistically significant differences between the accented and unaccented NPs. The results validate that the accented NPs had higher pitch and, for the most part, longer duration times compared to unaccented NPs.\footnote{8}

The experimental design was the same as in Experiment 1, in which each participant heard eight target discourses with pitch accent on the first potential referent (first NP) and eight with a pitch accent on the second (second NP). Two scripts were created to counterbalance the target items.

\footnote{8. The lack of significant differences in intensity between these conditions, even when uttered by an experienced speaker, reinforces the notion that intensity is not as reliable as cue of accent as pitch and duration, at least for focus stress.}
The same answer sheet from Experiment 1 was used here, since all the characters were identical and since the pseudo-random ordering of the items was retained.

Procedure

The current experiment used the same procedure as in Experiment 1.

Results

Similar to Experiment 1, the average number of correct choices for the filler questions was 15.8 out of 16. Scores ranged from 15 to 16 correct out of the total 16.

Figure 3 illustrates how often each participant chose the first NP, second NP, or foil as the referent for both pitch accent conditions.
Figure 3. Average number of times the first NP, second NP, and foil in a passive sentence was selected as pronoun referent.

Neither t-test revealed a statistically significant difference between pitch accent placement on the first NP versus the second NP. The rate of selecting the first NP in both pitch accent conditions was very similar (M=6.2 with pitch on the first NP; M=6.0 with pitch on the second NP) (t_{23} = 0.68, p=.50). Only one foil item was chosen in the experiment; therefore, the results for second NP selection are almost the exact inverse: (M=1.8 on first NP and 1.9 on second NP) (t_{23} = 0.50, p=.62).

In order to assess whether there was a significant interaction between the results with active constructions in Experiment 1 and passives in Experiment 2, a repeated measures ANOVA with stress as the within-subjects factor and sentence type as the between-subjects factor was calculated (using the average number of times the
second NP was selected). The results show a significant interaction ($F_{(1,46)} = 7.67$, $MSE = 7.59$, $p = .01$).

**Discussion**

The results of Experiment 2 showed that pitch accent did not affect pronoun referent resolution when the potential referents were presented in a passive construction. These results, then, do not support the FSSI hypothesis, which states that pitch accent increases the likelihood that the accented constituent will be chosen as the pronoun’s referent. The only difference between the discourses in Experiment 1 and Experiment 2 was the transformation of actives to passives – shifting the theta roles of the two constituents with the use of a by-phrase. Order of presentation of the two potential referents remained the same. Even though the experiment resulted in a null result, the interaction between active and passive constructions was significant, showing that there was a difference in how pitch accent interacted with the sentence constructions.

Of note was the fact that the subject was selected a majority of the time, as in Experiment 2. Therefore, regardless of who is doing the action, participants favor the subject as the pronoun’s referent. A possible reason why pitch accent had no effect in passive constructions is offered by Kehler (2003): Passive constructions mark the Patient as the topic – thus, the listener has no motivation to abandon such a strong cue of salience, despite the fact that a pitch accent appears on the other constituent. In contrast, active constructions do not unambiguously mark the topic.
Given the results from Experiment 2, the FSSI Hypothesis requires a modification:

**Focus Stress as a Secondary Influence Hypothesis, Revised (FSSIr):** Pitch accent placed on a potential referent of a pronoun increases the likelihood that native informants will select the prosodically focus-marked entity as the referent, unless another entity is explicitly marked as salient.

Explicit marking of salience includes selecting a sentence construction or verb that allows an entity with a thematic role not canonically appearing in subject position to be realized as the subject (as in passive constructions).

In Experiment 1, which used active constructions, the subject was not explicitly marked as salient, and pitch accent had an effect. With passives in Experiment 2, the passive sentence construction brought the Patient into the subject position, and explicitly marked this entity as being salient. Here, pitch accent had no effect because the sentence construction was the cue that listeners heeded when considering which entity was more salient in the discourse.

Although the modified FSSIr Hypothesis correctly predicts the results of Experiments 1 and 2, it should be tested further given the addition of a new clause. In Experiment 3, a different construction was used to create a condition in which the pitch accent should again influence pronoun resolution.
Experiment 3: Focus Stress in Conjoined Structures

As stated in the FSSIr Hypothesis, pitch accent should affect pronoun assignment when no other entity is explicitly marked as salient. If both potential referents were assigned the same theta role, then, pitch accent should again encourage informants to choose the accented referent more often than when it is not accented. One way to equalize the theta roles for both potential referents is to place them in a conjoined phrase such that the two entities do not interact with each other. The FSSIr Hypothesis predicts that when the referents are presented in the conjoined phrase, pitch accent will increase the likelihood that the accented entity will be selected as the referent. Since no other factors such as theta role preference is increasing the salience of the first NP, pitch accent should influence coreference by making the accented NP more salient to the listener.

Method

Participants

Participants were 24 volunteers recruited from the community. All of the participants were different from those run in Experiments 1 and 2. Eighteen were college students from UC Berkeley. Six were working professionals ranging in age from 26 to 55. All were native English speakers. Ten were male and fourteen were female.
Design and Materials

Materials consisted of the same three-sentence discourses from Experiments 1 and 2, except that in the 16 experimental discourses, the second sentence presented the first and second NPs in a conjoined noun phrase. Thus, both assumed the same theta role since no verb intervened between them; however, both remained in object position in the matrix clause (“As she was going to her room to grab a book, she noticed a paramedic and a captain down the hall.”). As with the previous two experiments, two versions of the second sentence were recorded. In one version, a pitch accent was placed on the first NP in the conjoined phrase and in the other, a pitch accent was placed on the second NP.

To ensure the integrity of the prosodic delivery of the materials, the audio recordings underwent the same acoustic analysis as in the previous experiments. Table 3 shows the mean scores of pitch, duration and intensity for each condition.

Table 3.

Mean pitch, duration and intensity measurements for each type of NP in Experiment 3.

<table>
<thead>
<tr>
<th>NP Type</th>
<th>Pitch (Hz)</th>
<th>Duration (ms)</th>
<th>Intensity (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st NP Accented</td>
<td>332</td>
<td>757</td>
<td>67</td>
</tr>
<tr>
<td>1st NP Unaccented</td>
<td>268</td>
<td>549</td>
<td>69</td>
</tr>
<tr>
<td>2nd NP Accented</td>
<td>310</td>
<td>675</td>
<td>67</td>
</tr>
<tr>
<td>2nd NP Unaccented</td>
<td>246</td>
<td>495</td>
<td>67</td>
</tr>
</tbody>
</table>
As with the previous two experiments, statistical analyses revealed a significant difference between the pitch measurements of accented first NPs versus unaccented first NPs ($t_{15} = 4.01, p<.001$), with the same pattern of increased pitch for accented versus unaccented second NPs ($t_{15} = 3.72, p<.001$). The same patterns emerged when comparing the accented and unaccented NPs that appeared in the same sentence, (for accented first NPs and unaccented second NPs: $t_{15} = 8.34, p<.001$; for unaccented first NPs and accented second NPs: $t_{15} = 3.09, p=.01$). Accented first NPs had longer duration times than unaccented first NPs ($t_{15} = 4.01, p<.001$), with the same results for accented and unaccented second NPs ($t_{15} = 3.72, p<.001$). When comparing duration measurements within the sentence, accented first NPs were significantly longer than unaccented second NPs ($t_{15} = 6.48, p<.001$), but the duration of unaccented first NPs versus accented second NPs only showed a trend ($t_{15} = 2.05, p=.06$). As with Experiment 2, none of the comparisons of intensity resulted in statistically significant differences between the accented and unaccented NPs. The analysis of the materials again validate that the accented NPs had higher pitch and, for the most part, longer duration times compared to unaccented NPs. (Note again that intensity is not a strong indicator of accent in this context.)

The experimental design was the same as in the previous two experiments, in which each participant heard eight target discourses with pitch accent on the first potential referent and eight with a pitch accent on the second.

The same answer sheet from Experiment 1 was used here.
Procedure

The current experiment used the same procedure as in Experiment 1.

Results

For the filler questions, the average number of correct choices was 15.4, with a range of 13 to 16 out of the total 16.

The average number of times each participant chose the first NP of the conjoined phrase, second NP and foil as the referent were calculated for both conditions, as shown in Figure 4.

Figure 4. Average number of times the first NP, second NP, and foil in a conjoined phrase were selected as the pronoun referent.
T-tests did reveal statistically significant differences between conditions when comparing pitch accent placement. Participants chose the first NP significantly more when pitch accent was placed on the first NP (M=4.0) compared to when pitch accent appeared on the second NP (M=3.4) ($t_{23} = 2.46$, $p=.02$). Likewise, participants selected the second NP (M=4.4) more often when pitch accent emphasized the second NP compared to when it emphasized the first NP (M=3.9), ($t_{23} = 2.08$, $p=.049$).

**Discussion**

As predicted, the results of Experiment 3 show that pitch accent on one of two potential referents increases the probability of selecting the accented referent. Therefore, the findings support the FSSIr Hypothesis.

Debriefing interviews revealed that some participants always chose the last-mentioned entity as the referent of the pronoun. This contrasts with the first two experiments in which participants reported selecting the subject, or first-mentioned entity. This is why the average for selecting the second NP is higher in this experiment compared to when the referents were placed in active or passive constructions.

Although Experiment 3 illustrated that the prosodic effect returns when the materials do not explicitly mark one entity as the more salient, the experiment did not demonstrate how the effect can return and then disappear again, as with actives and passives. Another test of the Hypothesis could do just this by manipulating both the presence and absence of the effect through a variation in the materials.
Previous research has shown that Source and Goal theta roles embody a similar contrast to Agents and Patients (Stevenson et al., 1995; Arnold, 1998). The drawback of creating an Agent-Patient contrast by using actives and passives is that a by-phrase is introduced in the passive construction. The benefit of using a Source-Goal / Goal-Source contrast is that the syntactic structure remains identical between the two conditions. Experiments 4 and 5 used materials with Source and Goal theta roles to determine whether the presence and absence of the effect of pitch accent can be manipulated as it was with Agent and Patient theta roles, according to the FSSIr Hypothesis.

Experiment 4: Focus Stress with Source-Goal Theta Roles

The FSSIr Hypothesis states that pitch accent increases the likelihood that the stressed candidate will be selected unless another entity is explicitly marked as salient. The hypothesis can be put to the test with discourses containing two potential referents with Source and Goal theta roles. In Arnold’s (1998) corpus work, she showed that for event verbs, comprehenders focus on the consequences of the action. For these verbs, Goals tend to be more salient than Sources. According to the FSSIr, when the Goal appears as the second NP and the pitch accent also appears on the second NP, then the probability that participants will break from the default preference for the subject should increase. However, when the Goal appears as the subject, then both the subjecthood of the first NP and the theta role act as cues explicitly marking the entity as salient, and pitch accent should have no effect. Experiment 4 used materials with
Source-Goal theta roles; whereas Experiment 5 used discourses with Goal-Source.
The hypothesis predicts that prosody will have an effect in Experiment 4 (Source-Goal), but not in Experiment 5 (Goal-Source).

Method

Participants

Participants were 24 volunteers recruited from the community. None of the subjects had participated in the previous experiments. All were native English speakers. Ages ranged from 18 to 49 with an average of 30.8. Participants represented a range of occupations including students, engineers, and office managers. Nine were male and 15 were female.

Design and Materials

Materials consisted of the same three-sentence discourses from Experiments 1, 2 and 3, except that in the second sentence, the first and second NPs had Source and Goal theta roles respectively (“As she was going to her room to grab a book, she noticed a paramedic giving a preserver to a captain down the hall.”). As with the previous two experiments, two versions of the second sentence were recorded. In one version, a pitch accent was placed on the first NP, and in the other, a pitch accent was placed on the second NP.
To ensure the integrity of the prosodic delivery of the materials, the audio recordings underwent the same acoustic analysis as in previous experiments. Table 4 shows the mean scores of pitch, duration and intensity for each condition.

Table 4.

Mean pitch, duration and intensity measurements for each type of NP in Experiment 4.

<table>
<thead>
<tr>
<th>NP Type</th>
<th>Pitch (Hz)</th>
<th>Duration (ms)</th>
<th>Intensity (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st NP Accented</td>
<td>324</td>
<td>773</td>
<td>66</td>
</tr>
<tr>
<td>1st NP Unaccented</td>
<td>278</td>
<td>543</td>
<td>68</td>
</tr>
<tr>
<td>2nd NP Accented</td>
<td>361</td>
<td>707</td>
<td>65</td>
</tr>
<tr>
<td>2nd NP Unaccented</td>
<td>245</td>
<td>462</td>
<td>66</td>
</tr>
</tbody>
</table>

As with the previous two experiments, statistical analyses revealed a significant difference between the pitch measurements of accented first NPs versus unaccented first NPs ($t_{15} = 2.63$, $p<.05$), with the same pattern of increased pitch for accented versus unaccented second NPs ($t_{15} = 4.34$, $p<.01$). The same patterns emerged when comparing the accented and unaccented NPs that appeared in the same sentence, (for accented first NPs and unaccented second NPs: $t_{15} = 3.28$, $p<.05$; for unaccented first NPs and accented second NPs: $t_{15} = 3.83$, $p=.01$). Accented first NPs had longer duration times than unaccented first NPs ($t_{15} > 5.36$, $p<.01$), with the same results for accented and unaccented second NPs ($t_{15} = 7.30$, $p<.01$). When comparing duration measurements within the sentence, accented first NPs were significantly
longer than unaccented second NPs ($t_{15} > 10$, $p<.001$), and unaccented first NPs were shorter than accented second NPs ($t_{15} = 3.70$, $p=.01$). Only one of the comparisons of intensity resulted in statistically significant differences between the accented and unaccented NPs. Unaccented first NPs measured significant higher in decibels (lower intensity) compared to accented second NPs in the same sentences ($t_{15} = 2.94$, $p<.05$). The analysis of the materials again validate that the accented NPs had higher pitch and longer duration times compared to unaccented NPs.

The experimental design was the same as in the previous two experiments, in which each participant heard eight target discourses with pitch accent on the first potential referent and eight with a pitch accent on the second potential referent.

**Procedure**

Participants were seated in front of a computer and were fitted with a pair of headphones. All audio was played by the computer. Participants advanced through the experiment by clicking a button on the screen. Instead of an answer sheet, participants saw a question related to the discourse they were hearing and three choices. Participants made a selection by clicking a radio button in front of their chosen answer. Questions and choices appeared exactly as they did on the answer sheet in the previous experiments except that the participant only saw the question and choices relevant to the discourse they were hearing. Participants advanced to the next item by clicking a “next” button under the question and choices.
Results

For the filler questions, the average number of correct choices was 15.5, with a range of 14 to 16 out of the total 16.

The average number of times each participant chose the first NP (subject of the embedded clause), second NP and foil as the referent were calculated for both conditions, as shown in Figure 5.

![Average Number of Times Each NP (SOURCE / GOAL) Was Selected as the Pronoun Referent (n=24)](image)

Figure 5. Average number of times the first NP, second NP, and foil were selected as the pronoun referent. Theta roles for potential referents were Source and Goal.

T-tests revealed statistically significant differences between conditions when comparing pitch accent placement. Namely, participants chose the first NP significantly more when pitch accent was placed on the first NP (M=7.3) compared to when pitch accent appeared on the second NP (M=6.3) \((t_{23} = 2.94, p<.01)\). Likewise,
participants selected the second NP (M=1.7) more often when pitch accent emphasized the second NP compared to the first NP (M=0.7), (t_{23} = 3.03, p<.01).

Discussion

For Experiment 4, the high scores on the filler questions verified that participants were attending to the task. When potential referents took on the theta roles Source and Goal, the FSSI held – focus stress on one of two NPs increased the likelihood that participants would select that NP. As with active structures with Agent and Patient theta roles, the size of the effect was quite small in comparison to the preference for the subject. Thus, the results here corroborated those of Experiment 3 in that focus stress functioned as a secondary influence. This was observed when the order of the thematic roles did not explicitly mark the subject as the salient entity. The true test of the FSSI is to cause the effect to disappear by using materials in which the positioning of the theta roles does mark the subject as the most salient. Experiment 5 did just this.

Experiment 5: Focus Stress with Goal-Source Theta Roles

The purpose of Experiment 5 was to test the last clause of the FSSI restated here:

Focus Stress as a Secondary Influence Hypothesis, Revised (FSSIr): Pitch accent placed on a potential referent of a pronoun increases the likelihood that
native informants will select the prosodically focus-marked entity as the referent, unless another entity is explicitly marked as salient.

The presence of the effect was illustrated by the significant results of Experiment 4. Now, the absence of the effect must be shown for cases in which the ordering of the theta roles marks the subject as the most salient. The factor exploited in the experiment was the theta roles. In contrast to verbs that project Source / Goal theta roles on to their arguments, Experiment 5 used verbs that reversed the roles to Goal / Source. The hypothesis predicts that pitch accent will have no effect in this case.

Method

Participants

Participants were 24 volunteers recruited from the community. None of the subjects had participated in the previous experiments. All were native English speakers. They represented a range of occupations including students, designers, office managers, analysts and curators. The average age was 28.5, ranging from 18 to 41. Five were male and 19 were female.

Design and Materials

Materials consisted of the same three-sentence discourses from Experiments 4, except that in the second sentence, the first and second NPs had Goal and Source theta roles respectively (“As she was going to her room to grab a book, she noticed a
paramedic taking a preserver from a captain down the hall.”). As with the previous experiments, two versions of the second sentence were recorded. In one version, a pitch accent was placed on the first NP (the subject of the embedded clause), and in the other, a pitch accent was placed on the second NP.

To ensure the integrity of the prosodic delivery of the materials, the wave files underwent the same acoustic analysis as in the previous experiments. Table 5 shows the mean scores of pitch, duration and intensity for each condition.

### Table 5.

Mean pitch, duration and intensity measurements for each type of NP in Experiment 5.

<table>
<thead>
<tr>
<th>NP Type</th>
<th>Pitch (Hz)</th>
<th>Duration (ms)</th>
<th>Intensity (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt; NP Accented</td>
<td>347</td>
<td>740</td>
<td>66</td>
</tr>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt; NP Unaccented</td>
<td>271</td>
<td>513</td>
<td>67</td>
</tr>
<tr>
<td>2&lt;sup&gt;nd&lt;/sup&gt; NP Accented</td>
<td>341</td>
<td>727</td>
<td>65</td>
</tr>
<tr>
<td>2&lt;sup&gt;nd&lt;/sup&gt; NP Unaccented</td>
<td>202</td>
<td>444</td>
<td>65</td>
</tr>
</tbody>
</table>

Statistical analyses revealed a significant difference between the pitch measurements of accented first NPs versus unaccented first NPs ($t_{15} = 6.15$, $p < .01$), with the same pattern of increased pitch for accented versus unaccented second NPs ($t_{15} > 10$, $p < .001$). The same patterns emerged when comparing the accented and unaccented NPs that appeared in the same sentence, (for accented first NPs and unaccented second NPs: $t_{15} > 10$, $p < .001$; for unaccented first NPs and accented
second NPs: $t_{15} = 6.85, p<.01$). Accented first NPs had longer duration times than unaccented first NPs ($t_{15} = 6.26, p<.01$), with the same results for accented and unaccented second NPs ($t_{15} = 7.30, p<.01$). When comparing duration measurements within the sentence, accented first NPs were significantly longer than unaccented second NPs ($t_{15} > 10, p<.01$), and the duration of unaccented first NPs was shorter than accented second NPs ($t_{15} = 5.10, p<.01$). None of the comparisons of intensity resulted in statistically significant differences between the accented and unaccented NPs.

The experimental design was the same as in the previous experiments, in which each participant heard eight target discourses with pitch accent on the first potential referent and eight with a pitch accent on the second potential referent.

The same computer interface used in Experiment 9 was used here.

**Procedure**

The current experiment used the same procedure as in Experiment 9.

**Results**

For the filler questions, the average number of correct choices was 15.7, with a range of 15 to 16 out of the total 16.

The average number of times each participant chose the first NP, second NP and foil as the referent were calculated for both conditions, as shown in Figure 6.
Figure 6. Average number of times the first NP, second NP, and foil were selected as the pronoun referent. Theta roles of potential referents were Goal and Source.

T-tests did not reveal statistically significant differences between conditions. Participants chose the first NP when pitch accent was placed on the first NP (M=6.4) just as often when pitch accent appeared on the second NP (M=6.6) \((t_{23} = .46, \ p = .65)\). Likewise, participants selected the second NP (M=1.4) equally as often when pitch accent emphasized the second NP compared to when it emphasized the first NP (M=1.5), \((t_{23} = .36, \ p = .72)\).

Because of the null results, an ANOVA was calculated to assess whether there was a significant interaction between pitch accent and the Goal-Source and Source-Goal sentence types. (The calculation used number of times the second NP was selected as the referent.) A repeated measures ANOVA with prosody as the within-
subjects factor and sentence type as the between subjects factor revealed a significant interaction ($F_{(1,46)} = 5.14$, MSE = 7.12, $p = .03$).

Discussion

As expected, Experiment 5 produced a null result. The materials were identical to those used in Experiment 4 aside from the verb. Thus, it was the manipulation of the theta roles realized on the pronouns’ potential referents that caused the effect to appear in one arrangement and disappear in another. When the ordering of the thematic roles promoted the subject as the most salient entity, pitch accent was not enough to sway interpretations.

Taken together, the findings from the five experiments above validated pitch accents’ influence in pronoun referent resolution. Instead of a determinate force, pitch accent serves as a secondary influence behind the more powerful cues of syntax and semantics. The experiments also revealed pitch accents’ intricate interaction with some of these other cues. Most notably, pitch accent lost its potency when the salience of one of the candidates was bolstered by the arrangement of theta roles.

Table 6 shows how different cues marked each potential referent as salient, and how these cues interacted with the presence or absence of the effect of pitch accent.
Table 6.

How Theta Roles Influenced Pronoun Referent Resolution.

<table>
<thead>
<tr>
<th>Order of Theta Roles</th>
<th>Did the Most Salient</th>
<th>Effect of Accent?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Thematic Role</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appear as Subject?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agent-Patient (Actives)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Patient-Agent (Passives)</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Patient-Patient (Conjoined Phrase)</td>
<td>N/A</td>
<td>Yes</td>
</tr>
<tr>
<td>Source-Goal</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Goal-Source</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

For actives, nothing explicitly marked one entity as being the most salient, and the results showed an effect of pitch accent. For passives, the Patient was placed in subject position, and explicitly marked the subject as being the topic of the sentence. Here pitch accent showed no effect because it could not override the strong cue inherent in the sentence construction. In a conjoined phrase, both potential referents had the same theta role, and therefore nothing explicitly indicated that one entity should be more salient than the other. This offered pitch accent the opportunity to sway interpretations again. Similar to Agent and Patient theta roles, Source and Goal produced the same pattern of results. Pitch accent affected interpretations when the most salient thematic role appeared as the object, but had no effect when the order of the thematic roles suggested that the subject was most salient entity.
The modified FSSIr Hypothesis correctly predicts the findings in this chapter. That said, there are many possibilities for testing the hypothesis further. For example, other cues such as general world knowledge and parallelism, which involve other coherence relation categories such as Contiguity and Resemblance respectively, would pose interesting tests for the hypothesis. Currently, the FSSIr assumes a Contiguity relation reflected through the strong selection of the subject. Since parallelism can be a powerful determinate of referents in parallel structures, the hypothesis might need to be adjusted for other coherence relations such as Resemblance.

The experiments in this chapter contributed empirical evidence in support of a hypothesis involving stress on potential referents – a different type of prosodic cue than those reported previously.

In the next chapter, the more canonical utilization of pitch accent as contrastive stress was addressed, but from a new angle. In particular, contrastive stress on pronouns was used to probe the time course of prosodic processing. The end goal was to discover how pitch accent interacts with pronouns in new ways (as seen in this chapter) and in real-time (the next chapter) so that an accurate representation of prosodic processing could be integrated into a theory of discourse anaphora.
PITCH ACCENT AS CONTRASTIVE EMPHASIS

Pitch accent that conveys contrastive stress on a pronoun has received much more attention in the literature than pitch accent as focus stress. Contrastive stress in this context has even inspired several researchers to propose ways to represent pitch accent in current theories of anaphora (Beaver, 2002; Cahn, 1995; Kameyama, 1997, Nakatani, 1997; Smyth, 1994). The problem with many of these approaches, though, is that they make assumptions that have been shown to be incomplete, or, at worst, untrue.

As mentioned in Chapter 1, the proposals by Beaver, Cahn and Nakatani all started with Centering Theory as the framework for how to choose between potential referents. In all three cases, the researchers began with Centering Theory and then modified it to incorporate prosody. The main problem with this approach is that the predictions from Centering do not account for parallelism preferences with object pronouns. Consider the following discourse:

(56) A butcher, baker, and candlestickmaker were having a fight as they were floating in the ocean on a raft. The butcher hit the baker; and then the candlestick maker hit him; …

Here, the C_f (if addressing each clause of the conjoined sentence separately), would be (butcher,baker), with the preferred Center being butcher. For the current utterance,
the possible $C_b$-$C_f$ combinations (after applying filters reported in Brennan et al. (1987)) would be the following:

i. $< \text{butcher}, (\text{candlestickmaker, butcher})>$

ii. $< \text{baker}, (\text{candlestickmaker, baker})>$

The transitions for these two possible combinations are different, where (i) is Retaining and (ii) is Shifting. In this case, Centering would select (i) since Retaining transitions are more coherent than Shifting. However, most informants would actually select \textit{baker} as the referent because of preferences based on parallelism.

Thus, augmenting Centering Theory so that it accounts for prosody does nothing to help account for a similar sentence with contrastive stress:

(57) A butcher, baker, and candlestickmaker were having a fight as they were floating in the ocean on a raft. The butcher$_j$ hit the baker and the candlestick maker hit HIM$_j$ …

The irony is that both Cahn and Beaver used parallelism examples as evidence of their proposals. However, their parallel structures used subject pronouns instead of object pronouns. When using subject pronouns, the preferences of Centering Theory are the same as those of parallelism.

The fact of the matter is that parallelism is necessary for contrastive stress to be so effective. As Venditti et al. (2003) showed, the coherence relation of the
materials drastically changes the effect of prosody. In materials with Resemblance
realized through parallelism, participants exhibited more gazes to the second NP (with
subject pronouns), whereas in discourses with Cause-Effect, contrastive stress did not
bring participants to change their eye fixations to the second NP.

Venditti et al.’s results are rather damaging to Kameyama’s (1997)
Complementary Preference Hypothesis, which states that the preferred value of a
stressed pronoun is the complementary value of its unstressed counterpart. Although
stress encourages a complementary interpretation of the referent in parallelism
examples, this pattern does not generalize to other coherence relations where the
interpretation usually stays with the most salient entity. Smyth (1994) also makes a
sweeping generalization about prosody with regard to his Extended Feature Match
Hypothesis. In a parallel sentence construction, if a pronoun’s grammatical role
matches that of a potential antecedent’s, then the pronoun selects that antecedent;
however, if the pronoun is stressed, this selection is blocked. Unfortunately, Smyth’s
proposal does not account for cases outside parallelism either. Thus, as Kehler (2002)
points out, theories that rely on Centering Theory cannot account for parallelism
examples and theories that count on parallelism effects cannot explain patterns of
other coherence relations.

In addition there are still open questions that have yet to be answered: (1)
When is prosod ic information used by the system: is it in an early stage of processing
or does it get integrated much later? (2) How does prosod ic information affect
processing load: does stress increase cognitive load or alleviate it?
The first gap in the research is that not much is known about how prosody is processed in real time. Although many researchers have made the claim that prosodic information is used immediately, most of this research has explored syntactic attachment ambiguity and not pitch accent (Ferreira et al., 1996; Marslen-Wilson et al. 1992; Pynte & Prieur, 1996; Schafer et al., 1996; Speer et al., 1996; Watt and Murray, 1996).

Recently, there have been a few studies looking at the effects of pitch accent on pronoun interpretation during on-line processing using eye tracking (Vendetti et al., 2001, 2002). The results suggest that in the unaccented condition (parallelism), only the NP with congruent grammatical function (the first NP) is activated, whereas in the accented condition, both the first and second NPs are activated momentarily and then the participants continue to gaze at the second NP. Unfortunately, the authors were reporting on filler sentences without statistical analyses because none of the experimental sentences had parallel structure. Venditti et al. (2002) postulated that the temporary ambiguity might have been the result of an ambiguous coherence relation and tested this hypothesis in a follow-up experiment (Venditti, 2003). For this study, they used materials both with narratives and parallelism (repeated here from Chapter 2):

(58) The kangaroo tapped the elephant with a magic wand.

(59) a. Then HE tapped the raccoon. (Parallel)

b. Then HE looked into the crystal ball. (Narrative)
As in the other experiment, participants fixated on both NPs early on. Then at the offset of the verb in the parallel clause, participants started to show a preference for the second NP. The conclusion drawn was that the verb assisted the listener in resolving the coherence relation, which aided in the interpretation of the stressed pronoun.

Unfortunately, the results do not get us any closer to understanding when prosody is used to constrain interpretations. Since the verb introduced ambiguity of the coherence relation, we do not know whether prosody would have had an immediate effect otherwise.

Another open question is whether contrastive stress actually makes processing easier or more difficult. The answer to this question can inform the way prosody is represented in future models of pronoun referent resolution. If the presence of contrastive stress taxes processing, then it is more likely that it introduces additional steps or requires more processing resources.

This chapter presents five experiments that address the question of when prosodic information is used by the system. Experiment 6 established a baseline for the effect of contrastive stress when using object pronouns (as opposed to subject pronouns). Using the same materials, Experiments 7 and 8 explored whether prosodic information is available during real-time language processing immediately at the pronoun. Experiments 9 and 10 examined reactivation of potential referents
downstream from the pronoun and addressed the question of whether stress increases processing load or alleviates it.

**Experiment 6: Off-line Effects of Contrastive Stress**

The purpose of Experiment 6 was to test the effect of contrastive stress with object pronouns. Since so many experiments have focused on subject pronouns, the goal was to show that the same effect occurs with object pronouns. There is a need to break free from the standard use of the subject pronouns in parallel structure. Object pronouns should be used for two important reasons: (a) Object pronouns do not confound parallelism with preferences for the first-mentioned entity as subject pronouns do. (b) Object pronouns occur after the verb and therefore the parallel structure in the materials is already available to the listener. This second point is important since researchers (Venditti et al. 2001) have drawn conclusions about the time course of processing based on the fact that the verb is not available to listeners until after the pronoun is heard. With object pronouns, the verb has already been processed.

The hypothesis is that preferences will follow the pattern of parallelism in which the pronoun is associated with the entity in the same grammatical position in the conjoined clause – in this case the second NP. When a pitch accent is placed on the pronoun, this indicates to the listener that the canonical interpretation should be overridden, and the referent should switch to the first NP.
In addition to contributing more data for object pronouns, Experiment 6 also validated materials to be used in the on-line studies. When manipulation of stress in the materials results in significant differences in interpretation, this verifies that the experimental sentences are delivered with the intended prosodic contour.

Method

Participants

Sixteen volunteers from the community participated in the study. Participants were working professionals with an average age of 27.5. Eight were male and eight were female. All had learned English fluently by age six.

Design and Materials

The stimulus materials consisted of 48 target discourses. Each discourse contained three sentences. The first sentence introduced the scenario. The second was always a conjoined sentence, with the first clause presenting an Agent and a Patient, and the second clause describing a different Agent performing the same action on a pronoun. Because a pronoun was used, it was ambiguous to whom the action in the second clause was being done. Both the first and second NPs of the previous clause were potential referents of the pronoun. The third sentence concluded the scenario. The following is an example of a discourse used in the experiment:
Excited by their costumes for the Halloween play, some of the third graders started rough-housing back stage. An alien pinched an acrobat just behind the curtain and a ghost pinched her near the backdrop. Soon the whole audience heard the giggling back stage.

Discourses were digitized using SoundForge software. Two versions of the second sentence were recorded: one with a canonical pitch accent contour, and another with contrastive accent over the pronoun. Discourses were spliced together so that there were two separate audio files for each discourse labeled with or without stress on the pronoun.

An acoustic analysis was performed on the materials to ensure that the delivery of the second sentence achieved the intended prosodic contour. Using Praat software (Boersma & Weenink, 2002), the maximum pitch accent over the pronoun was extracted. Duration was also calculated, although this measure was not directly motivated by Pierrehumbert’s prosodic theory. Table 7 shows the mean scores for each condition.
Table 7

Mean pitch and duration measurements of the pronoun in Experiment 6.

<table>
<thead>
<tr>
<th>Pronoun Type</th>
<th>Pitch (Hz)</th>
<th>Duration (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accented Pronoun</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>331</td>
<td>495</td>
</tr>
<tr>
<td>SD</td>
<td>25</td>
<td>57</td>
</tr>
<tr>
<td><strong>Unaccented Pronoun</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>207</td>
<td>271</td>
</tr>
<tr>
<td>SD</td>
<td>22</td>
<td>22</td>
</tr>
</tbody>
</table>

Statistical analysis of the prosodic measurements showed that across sentences all accented pronouns were delivered with significantly higher pitch than all unaccented pronouns ($t_{47} > 10$, $p<.001$). For duration, all accented pronouns were longer than unaccented pronouns ($t_{47} > 10$, $p<.001$). The results validate the delivery of the materials in that the pronouns with contrastive accent were higher in pitch and longer in duration than those that were not accented.

Items were divided into four scripts with a total of 34 discourses per script. Script 1A contained 12 experimental discourses with an unstressed pronoun, 12 with a stressed pronoun, and 10 filler discourses. Script 2A consisted of the same discourses except the accent on the pronoun for the experimental items was counterbalanced – a discourse with an unaccented pronoun in Script 1A appeared with contrastive accent on the pronoun in Script 2A, and vice versa for those with contrastive accent in Script
1A. Scripts 1B and 2B contained 24 different experimental discourses with the same fillers. Half of the experimental discourses had contrastive accent on the pronoun and half did not. The two scripts were counterbalanced for contrastive accent on the pronoun. A complete list of materials is listed in Appendix B.

Procedure

Prior to testing, each subject was told that he or she would be listening to short discourses of about three sentences each. Subjects were instructed to decide to whom the pronoun was referring and to circle that noun on the answer sheet provided. The answer sheet listed two nouns for each of the 34 discourses. For experimental items, the two nouns were the potential referents of the pronoun. By listing the potential referents, participant responses were not confounded with memory ability.

Results

When participants heard sentences with an unstressed pronoun, they selected the second NP as the referent 85 percent of the time, but when they heard sentences with a stressed pronoun, they selected the second NP only 20 percent of the time. A t-test revealed a statistically significant difference between the two sentence types, \( t_{15} = 6.71, p < .001 \). The effect was also observed when comparing averages across items \( t_{47} > 10, p < .001 \).
Discussion

The results showed that contrastive stress influences the ultimate choice of a referent for an ambiguous pronoun. In both conditions, listeners had a definite preference for one of the antecedents, but this bias differed depending on the pitch accent over the pronoun. For sentences without stress, listeners preferred the second NP as the antecedent, confirming the parallel function hypothesis (Smyth, 1994); and with contrastive stress, listeners showed a bias for the first NP. Since most studies have used subject pronouns, the present experiment offers complimentary results for object pronouns with contrastive stress in parallel structures.

Since the materials were validated (showing that prosody biases pronoun referent assignment), the following experiments investigated when the effect of the prosodic cue takes place during language processing.

Experiment 7: On-line Effects of 2nd NP at Baseline and Pronoun

In Experiment 7, the CMLP paradigm was used to investigate the time course of activation of possible pronoun referents during on-going language comprehension. As discussed in Chapter 2, the paradigm uses semantic priming to assess when the pronoun’s referent has been activated. The participant listens to discourses over headphones. During each discourse a string of letters flashes up on a computer screen and the subject decides whether it is an English word or not.
In this experiment, the focus was on the activation of the second NP. As the results from Experiment 6 indicated, the second NP is the preferred referent for sentences with an unstressed pronoun, but not with a stressed pronoun. Thus, in the CMLP experiment, some of the words that flashed up on the screen were related to the second NP, and others were matched controls.

There were several potential outcomes of the experiment. No priming at the pronoun in either condition might indicate delayed activation; however, the null result would have to be viewed with some skepticism since experimental power could have been the culprit. Priming for the second NP with unstressed pronouns but not stressed pronouns would suggest immediate use of the prosodic cue in processing and would reflect the off-line results. The opposite would be unexpected given the pattern of the off-line results. Finally, activating the second NP for both the stressed and unstressed pronouns would suggest a delay in processing pitch accent.

Method

Participants

Forty-eight subjects participated in the experiment. Participants were students from UC San Diego, UC Berkeley and Stanford. Two of the participants were professionals recruited from the community. UC San Diego students were given course credit for experimental participation. Berkeley and Stanford students were paid $15 for the session. The professionals participated for free. All had learned English
fluently by age six and had normal or corrected-to-normal hearing and vision. Of these, 14 were male and 34 were female. The average age was 19.8.

Materials

Discourses

The same 48 discourses from Experiment 6 were used in Experiment 7, in addition to 48 filler discourses. Although most of the fillers were random in syntactic structure and content, some of them closely resembled the format of the experimental discourses. The reason for this was that all the fillers (except for two of the six practice discourses at the beginning of the experiment) were coupled with non-words. Thus, having some of the non-word discourses mimic the experimental ones discouraged participants from associating a certain style of discourse with real-word probes.

Probes

Each experimental discourse was assigned four probe words, one related to the second NP of the sentence, one that served as a control, and two others to be used in Experiments 9 and 10 (a related and control probe for the first NP). To select related probes, semantic associates to the NPs in each experimental discourse were collected from 50 UC San Diego undergraduate psychology students who participated for course credit. Subjects gave their first associate to a list of words read aloud, and the most frequently mentioned associate was used as the related word for that sentence. For example, the second NP for the example discourse in Experiment 1 is *acrobat*. 
The semantic associated most often chosen for *acrobat* was *circus*. Thus, *circus* became the related probe word for this discourse.

Control words were selected from a database consisting of average reaction times (42 < n < 52 per word) on over 1700 words. The control word assigned to a given discourse was matched to the related word for number of letters, lexical decision reaction time, and frequency (Francis & Kučera, 1982). For example, the word *subway* was selected as the control word for *circus* since it matched most closely with regard to number of letters, reaction time and frequency compared to any other word in the database. Related and control words were matched a priori, so that any differences in reaction time would be attributable to experimental manipulation. Each filler discourse was assigned a non-word that followed normal phonological rules of English but was not entered in the English dictionary used by Microsoft Word 2000, for example, *chamots*.

**Design**

The experiment was a 3-factor within subjects design, with two levels per factor. The factors were Prosody (stress on pronoun x no stress on pronoun), Probe position (before pronoun x at the pronoun), and Probe type (related x control).

Since eight conditions were being tested, eight different scripts were prepared to counterbalance all conditions across discourses. The scripts were represented by eight input files to be used by the reaction time software. The input files contained the names of the audio files to be played, and timing information regarding what probes
should be flashed up on the computer screen and when in each discourse the probe should appear. 

For each script, 48 of the discourses were experimental and 48 were fillers. Half of the experimental discourses had unstressed pronouns and half stressed pronouns. For both the unstressed and stressed discourses, half of the probe positions were set 800 ms before the onset of the pronoun in the second sentence, and half the probe sites were located immediately after the offset of the pronoun. For the filler sentences, probe positions were placed in random places. For half of the experimental discourses with a probe site at an unstressed pronoun, the probe word was a semantic associate of the second NP, and for the other half, the probe was the control for the semantic associate of the second NP. Likewise for all the unstressed pronouns with probe positions 800 ms before the pronoun, half of the probes were related and half were control words. The same counterbalancing was done for stressed pronouns with probes points before and at the pronoun.

For each discourse, the three conditions (prosody, probe position and probe type) were rotated across scripts. For example, in Script 1, the first experimental discourse had an unstressed pronoun with a control probe appearing at the pronoun; in Script 2, that same discourse had an unstressed pronoun with a related probe appearing at the pronoun. In Scripts 3 and 4, the probe appeared before the pronoun, and in Scripts 5 through 8, the pronoun was stressed. Within each script, a subject responded to six discourses in each condition.
For all scripts the same pseudo-random order was used. No sentence or probe word was seen or heard more than once for any given participant.

**Apparatus**

The software program Tempo 1.2.4, developed by the Language and Cognition Lab at University of California, San Diego, operated the auditory presentation of the sentences and the visual display of the probe words. The program accessed individual files with digitized sentences and played each sentence from the computer in an order designated by the experimenter. Separate files with information about the position of the probe presentation told the software when to display the letter string on the monitor and simultaneously started a clock timer. Probes appeared on the screen for 500 ms. Once the software detected that one of two buttons on a button box was pressed, the timer was deactivated and a reaction time was recorded.

**Procedure**

Each subject was seated in a separate testing area in front of a computer display. Subjects were told that they would be listening to sentences over headphones and at some point during each discourse, a letter string would flash up on the computer screen. Subjects were instructed to press the right button on a button box if they thought the string of letters constituted a real English word, or the button on the left if it was not. During testing, discourses were played in stereo over a set of headphones. Participants were given six practice discourses to get comfortable with the task.
A simple comprehension task ensured that subjects were attending to the sentences. At random intervals, participants were asked to answer a multiple-choice question about the discourse they just heard. Instead of a probe word, the sentence, “Please answer question #1,” for example, would appear on the screen. This signaled to the subject that he or she was to circle on an answer sheet one of three choices regarding the topic of the previously heard discourse. Ten comprehension questions were asked during the course of the experiment. A filler discourse always followed a comprehension question so that circling answers on the answer sheet would never interfere with reaction times for target discourses.

A short off-line followed the on-line experiment, to help gauge the participant’s sensitivity to prosody. The on-line task was presented first so that participants would not consciously think about pronoun referent assignment and reactions times would reflect subconsciously processing. Following their participation in the on-line task, subjects were asked to make a judgment about the pronoun’s referent in ten discourses that were previously presented in the on-line experiment. Eight of these were experimental discourses and two were fillers. Four of the experimental discourses were versions recorded with canonical prosodic contour and four were the versions with stress over the pronoun. The same version of the sentence presented in the on-line experiment was presented in the off-line. The condition of the discourse was rotated across scripts.
Participants were told that they would be listening to some of the discourses again. After each discourse, participants heard a question asking to whom an action was being done. Two choices were presented on the computer screen side-by-side; namely, the first and second NPs. The presentation of each NP was varied such that the first NP was presented on the left side of the screen for half of the questions and on the right for the other half. The button presses were recorded and tallied during analysis.

Results

For the off-line questions presented at the end of the experiment, the average proportion of the time participants selected the second NP was calculated – the first NP was the exact inverse since the choices were mutually exclusive. For unstressed pronouns the mean was .68; for stressed pronouns it was .39. There was a statistically significant difference between the two conditions ($t_{47} = 4.44, p<.01$).

For each participant, eight means were calculated, one for each condition. Average mean reaction times across all subjects for each condition are shown in Figure 7 below.
Figure 7. Mean reaction times of probes related to the second NP and control probes when presented 800 ms before the pronoun and at the offset of the pronoun.

Data were analyzed using an ANOVA for a repeated measures design, with the following three factors: Prosody (unstressed x stressed), Probe position (before pronoun x at pronoun), and Probe type (related x control). The ANOVA revealed a significant main effect for Probe type ($F_{(1,47)} = 7.285, MSE = 41,729, p = .01$). There was also a significant interaction for Prosody x Probe position x Probe type ($F_{(1,47)} = 4.312, MSE = 20,548, p = .04$)

Planned comparisons (two-tailed, alpha level = .05) motivated by the probe type main effect were carried out separately for each probe site and sentence type. Four comparisons were made, each contrasting reaction times for related and control
probes in four conditions (unstressed before, unstressed at, stressed before and stressed at). The results revealed significant priming for related probes at the baseline position in discourses with an accent ($F_{(1,47)} = 7.353$, $MSE = 35,037$, $p = .01$). There was also significant priming for related probe words presented at the pronoun for unstressed sentences ($F_{(1,47)} = 5.402$, $MSE = 25,742$, $p = .02$). No other priming effects were observed.

Several two-way ANOVAs were calculated to determine whether there were significant interactions between the baseline and pronoun positions. For these tests, the factors were Probe position (before pronoun x at pronoun) and Probe type (related x control). Prosody was not a factor because separate ANOVAs were calculated for each sentence type. For unstressed pronouns, the interaction was not significant but showed a trend ($F_{(1,47)} = 2.19$, $MSE = 8,125$, $p = .14$). The same was true for discourses with stressed pronouns, but in the opposite direction ($F_{(1,47)} = 2.10$, $MSE = 12,675$, $p = .15$).

**Discussion**

The off-line profiling showed the same pattern of results as Experiment 6; namely, the majority of the time, participants chose the second NP as the pronoun’s referent when no stress was realized on the pronoun, but with stress, the first NP was the preferred choice.

The on-line results show a priming effect at the pronoun in discourses with an unstressed pronoun, but not with a stressed pronoun. For the unstressed pronoun
condition, priming was not observed before the pronoun (with an interaction approaching significance); thus, the effect cannot be attributed to general facilitation of the related word, but rather indicates a reactivation of the semantic associate of the antecedent at the pronoun. The condition where facilitation was observed before the pronoun was with stressed pronouns. For this condition, the priming effect disappeared at the pronoun.9

The findings suggest that prosodic information does indeed constrain activation of potential referents. Even though the second NP was available for unstressed pronouns, when contrastive accent was present, the second NP was no longer accessed as a potential referent. Thus, pitch accent seems to work at an early stage of processing (such as Kintsch’s (1988) Construction phase).

Since Experiment 7 examined only the effects of the second NP, the next logical step was to uncover the priming patterns for the first NP. Given the immediacy of prosodic processing in Experiment 7, similar results were expected when probing the first NP.

---
9. The pre-pronoun effect could be the result of some residual priming of the second NP from the previous clause. The effect might not have occurred in the unstressed condition since the canonical prosodic contour emphases the new NP in the sentences (about 800 ms before the pronoun); whereas contrastive stress emphasizes the pronoun instead of the new NP in the stressed condition. Bringing attention to the new entity through emphasis might have suppressed any residual effects in the unstressed sentences. A pilot study with probe points 800 ms before the pronoun and at the onset of the pronoun (rather than at the offset) show a similar priming of the second NP 800 ms before the pronoun in the stressed condition, but no other priming effects. Thus, no residual priming is observed just before the pronoun.
Experiment 8: On-line Effects of 1st NP at Baseline and Pronoun

In Experiment 8, the same CMLP paradigm was used to investigate the reactivation of the first NP at the pronoun. As the results from Experiment 6 (off-line) indicated, the first NP is the preferred referent for sentences with a contrastive accent on the pronoun. Experiment 7 provided some evidence that prosodic cues constrain pronoun referent activation immediately. Therefore, a similar outcome was anticipated in Experiment 8. Instead of limiting access, pitch accent on the pronoun should license activation of the first NP. Therefore, it was hypothesized that the results would show a priming effect for the first NP in the stressed pronoun condition.

Method

Participants

As with Experiment 7, 48 subjects participated in the study. Of these, 12 were male and 36 were female. Forty-three participants were students from UC San Diego, UC Berkeley and Stanford. Five of the participants were working professionals. UC San Diego students were given course credit for experimental participation. Berkeley and Stanford students were paid $15 for the session, and the working professionals participated for free. All had learned English fluently by age six and had normal or corrected to normal hearing and vision. The average age was 21.8.
Materials

Discourses

The same discourses used in Experiment 7 were used in Experiment 8.

Probes

Instead of using related and control words associated with the second NP, the probes used in Experiment 8 were either related to the first NP, or were assigned control words matched to each of the related words. The same procedure for selecting probe words in Experiment 7 was used for Experiment 8. For example, one of the target discourses contained the following second sentence:

(61) “… An alien pinched an acrobat just behind the curtain and a ghost pinched her near the backdrop…”

The related probe word for the first NP, alien, was space. The control word for space, which matched in number of words, lexical decision reaction time, and frequency was union. The same non-words used in Experiment 7 were used in Experiment 8.

Design

As with the previous study, Experiment 8 was a within-subjects, 3-factors design (Prosody x Probe position x Probe type). Eight different scripts were prepared to counterbalance all conditions across discourses. The scripts were identical to those
used in Experiment 7 except the probe words flashed up on the screen were associated with the first NP instead of the second. The same order of presentation of the discourses and design from Experiment 7 were used in the current experiment.

Apparatus

The software and apparatus used in Experiment 7 were the same as those used in Experiment 8.

Procedure

The procedure used in Experiment 8 was identical to that of Experiment 7.

Results

For the off-line results, the average proportion of the time the second NP was chosen as the referent in discourses with an unstressed pronoun was .60. With contrastive stress, the proportion decreased to .35. The difference was statistically significant ($t_{47} = 3.77, p<.01$).

For each participant, eight means were calculated, one for each condition. Average mean reaction times across all subjects for each condition are shown in Figure 8 below.
Figure 8. Mean reaction times of probes related to the first NP and control probes when presented 800 ms before the pronoun and at the offset of the pronoun.

An ANOVA for a repeated measures design revealed a significant main effect for Probe type ($F_{(1,47)} = 6.817$, MSE = 33,825, $p = .01$). No other main effects or interactions for this ANOVA were significant.

Planned comparisons were carried out separately for each probe site and sentence type. Four analyses were performed, each comparing reaction times for related and control probes in four conditions (unstressed before, unstressed at, and stressed before and stressed at). The results showed no significant priming in the baseline position 800 ms before the pronoun for either condition. At the pronoun, there
was a significant effect only for discourses with contrastive stress \((F_{(1,47)} = 6.705, \text{MSE} = 32,561, p = .01)\).

Two additional ANOVAs were calculated to assess whether there were interactions between the baseline and pronoun positions. For these tests, the factors were Probe position and Probe type. Neither interaction was significant (Unstressed: \(F_{(1,47)} = 1.56, \text{MSE} = 5,941, p = .22\). Stressed: \(F_{(1,47)} = .58, \text{MSE} = 2,670, p = .45\).

**Discussion**

The off-line data validated that subjects’ interpretations shifted with contrastive stress. As hypothesized, the on-line results showed a priming effect at the pronoun in discourses with an accented pronoun. Again, the effect cannot be attributed to general facilitation of the related probes since no priming was observed before the pronoun. The fact that the interaction was not significant was most likely an issue of power.

There are two interesting consequences of the results: one related to prosody and the other to parallelism. The first is that prosodic information is used by the system in the early stages of processing. The fact that the first NP was activated only when stress appeared on the pronoun indicates that it was the pitch accent that licensed the activation. Consistent with Experiment 7, the on-line results reflect the off-line findings showing that the change in interpretation from prosody happened almost at the very moment the pronoun was heard.
The results also shed light on the real-time processing of parallelism. Note that priming was observed for the first NP only in the discourses with an accented pronoun. When there was no stress, the first NP was not accessed. Thus, it seems that the first NP is not licensed as a potential referent in parallel constructions. Work by Nicol (1988) had shown that all syntactically licit antecedents were activated when a pronoun was encountered in the sentence. Even though the first NP was syntactically licit, it was not reactivated in the parallel construction. It is possible that factors other than binding constraints that are still related to syntax, limit the suite of potential antecedents. Parallelism could function the same way a syntactic reflex does. Data from Frisch’s (2003) recent study corroborate parallelism’s immediate influence. Recall that Frisch observed a P600 effect when gender forced the referent to be an entity with an incongruent grammatical function in a parallel structure. P600 effects usually accompany syntactic anomalies, not semantic ones.

Taken together Experiments 7 and 8 answer the question of when prosodic information becomes available during language processing. Similar to other factors such as gender and syntactic binding relations, pitch accent constrains (and licenses) activation of potential referents at the moment the pronoun is heard. This finding is consistent with other on-line studies that have shown immediate effects of prosody.

The results from these experiments also provide insight into Venditti et al.’s (2001, 2002, 2003) findings. Although Venditti et al. observed activation of both referents in conditions with contrastive stress, the result was probably caused by an
ambiguous coherence relation. When the verb was presented before the pronoun as in the current materials, only the first NP was activated.

**Contrastive Stress and Processing Load**

Since Experiments 7 and 8 provided evidence of *when* prosodic processing takes place, the question of *how* prosody affects processing load was addressed. Understanding how contrastive stress changes normal processing contributes to a more accurate representation of prosody in models of pronoun referent resolution. If increases in processing are observed at the pronoun, then it is possible that contrastive stress is introducing an additional step during referent activation. This seems like a viable possibility since previous approaches to integrating prosody into theories of discourse anaphora have started with a solution without prosody and then showed how a rule changed the default with the presence of stress. However, if increases in reaction time are not observed until later, then the additional processing is happening in the Integration phase and possibly during a second-pass interpretation.

The first place to look for processing load differences is in the reaction time data from Experiments 7 and 8. Any increases in processing load caused by prosody would be reflected in longer reaction times, particularly to the control probes which are not related to anything within its associated discourse (as opposed to measuring facilitation differences). However, no main effects or significant differences between control probe reaction times were evident. Thus, the processing load for sentences with contrastive stress is no different from sentences without stress at the time the
pronoun is heard. Prosody does not seem to add another step or require additional processing resources during referent activation.

The next place to look for processing load differences is downstream from the pronoun. Even though prosody may be fast-acting in constraining referent access, there may still be some effects in a later Integration phase. Several researchers have reported a general increase in reaction times as a sentence progresses (Gernsbacher, 1989; Gernsbacher, Hargreaves & Beeman, 1989; McDonald & MacWhinney, 1995). Gernsbacher claims that this effect is the result of intense structure building at the beginning of the sentence followed by the less-strenuous process of placing remaining items within the existing structure. Given this, anomalies in processing load should be detected when there is a general slowing of reaction times over time. Slower reaction times would violate Gernsbacher’s observation of reaction times quickening as the sentence progresses.

In Experiments 9 and 10, two probe points after the pronoun were examined to test the hypothesis that contrastive stress adds to cognitive load in a later stage of processing. Experiments 9 and 10 were also designed to measure facilitation of the second NP and the first NP in both sentence structures. Understanding the priming patterns in addition to load demands helped divulge the type of processing taking place downstream from the pronoun.
Experiment 9: On-line Effects of 2nd NP at 400 ms and 800 ms

In Experiment 9, the same CMLP paradigm was used to investigate whether there were increases in reaction time and possible reactivation of the second NP in two positions downstream from the pronoun (400 ms and 800 ms). Since Kintsch set the timing of his Integration phase at about 400 ms downstream, it was hypothesized that an increase in processing load in sentences with contrastive stress would be detected 400 ms after the pronoun. If no reaction time increase was observed and the second NP was still activated for unstressed pronouns, the effect is probably due to residual priming at the pronoun. If the second NP was accessed for stressed pronouns, however, then a second-pass processing mechanism is most likely involved. Since the second NP was not originally activated as a potential referent of the stressed pronoun (as indicated by Experiments 7).

Method

Participants

Participants were 48 students from UC San Diego, UC Berkeley and Stanford and working professionals recruited from the community. UC San Diego students were given course credit for experimental participation. Berkeley and Stanford students were paid $15 for the session. Five professionals volunteered for free. All had learned English fluently by age six and had normal or corrected to normal hearing and vision. Of these, 15 were male and 33 were female. The average age was 21.3.
Materials

Discourses
The same discourses used in Experiments 7 and 8 were used in the current experiment.

Probes
The probe words used in Experiment 7 were used in Experiment 9. These probes were either semantic associates of the second NP, control probes matched to each of the related probes, or non-word filler probes.

Design and Apparatus
The same design from Experiment 7 and 8 were used here, except that the probe positions were different. Instead of presenting words 800 ms before the onset of the pronoun and at the offset of the pronoun, probe words appeared 400 ms after the offset of the pronoun or 800 ms after the pronoun. Scripts were counterbalanced so that half of the positions were 400 ms downstream and half 800 ms downstream. The same rotation of conditions in Experiments 7 and 8 were employed here as well.

Procedure
The procedure used in Experiments 7 and 8 was also used in Experiment 4.
Results

For the off-line questions, the average proportion of the time the second NP was chosen as the referent for unstressed pronouns was .67, compared to .34 for stressed pronouns. Again, the difference was statistically significant ($t_{47} > 10$, $p < .01$).

As with the previous experiments, eight means were calculated, one for each condition. Average mean reaction times across all subjects for each condition are shown in Figure 9 below.

![Mean Reaction Times of Control and Related Probes 400ms and 800ms After the Pronoun (2nd NP)](chart)

Figure 9. Mean reaction times of probes related to the second NP and control probes when presented 400 ms and 800 ms downstream.

Two ANOVAs for repeated measures were calculated: one two-factor ANOVA for control probes (Prosody x Probe position) and another, three-factor
ANOVA for both control and related (Prosody x Probe position x Probe type). The first ANOVA did not reveal any significant effects. The second ANOVA showed a significant main effect for Probe type ($F_{(1,47)} = 6.073$, $MSE = 18,398$, $p = .02$). No other main effects or interactions were significant.

Planned comparisons were motivated by the second ANOVA’s main effect of Probe type. The comparisons of related and control probes showed no significant priming of the second NP 400 ms downstream from the pronoun for either stressed or unstressed pronouns. However, 800 ms after the pronoun, a significant priming effect was observed in sentences with contrastive stress ($F_{(1,47)} = 4.503$, $MSE = 21,480$, $p = .04$).

The interactions between the positions 400 ms and 800 ms downstream did not show statistical significance.

**Discussion**

The off-line results showed that contrastive stress on the pronoun changed subjects’ interpretation of the referent.

In the on-line portion of the experiment, no difference was observed between unstressed and stressed pronoun conditions 400 ms after the pronoun, against the hypothesized prediction. There did seem to be a different pattern after 800 ms as a result of prosody, where reaction times to control probes slowed, although the effect was not significant.
Of interest was the fact that priming occurred for the second NP in the stressed condition. This is the opposite result from the off-line interpretations, in which participants preferred the first NP with contrastive stress. The results suggest that a second-pass interpretation, possibly in the Integration phase given its late occurrence, is being considered for pronouns with stress. Thus, even though the correct pronoun is activated at the moment the pronoun is heard, the system continues to assess whether other referents might be more fitting.

Even though Experiment 9 provided an indication of how processing proceeds downstream, the data profile was not complete without probing the first NP.

Experiment 10: On-line Effects of 1st NP at 400 ms and 800 ms

Experiment 10 used identical materials and methodology to Experiment 8 (which focused on activation of the first NP). The only differences were the positions at which the probes appeared on the computer screen. As with the previous experiment, probe positions were 400 ms and 800 ms after the offset of the pronoun.

Results of Experiment 9 forced an adjustment to the hypothesized results. Instead of observing slowing effects 400 ms after the pronoun, processing load was anticipated 800 ms downstream. Since Experiment 9 provided evidence of some reactivation of the second NP downstream, the prediction was that there would be a second-pass interpretation – given that both NPs were potential referents, it was predicted that the first NP would also be considered and accessible 800 ms after the
pronoun. Priming effects would be observed, but only for discourses with a stressed pronoun.

Method

Participants

Participants were 48 students from UC San Diego, UC Berkeley and Stanford. UC San Diego students were given course credit for experimental participation. Berkeley and Stanford students were paid $15 for the session. There were 16 males and 32 females. The average age was 20.2. All had learned English fluently by age six and had normal or corrected to normal hearing and vision.

Materials

Discourses

The same discourses used in Experiments 7, 8 and 9 were used here.

Probes

The probe words used in Experiment 8 were used in Experiment 10. These probes were either semantic associates or control words for the first NP or non-word filler probes.

Design and Apparatus

Experiment 10 used the same design as Experiment 9.
Procedure

The procedure used in Experiments 7, 8, and 9 was also used in Experiment 10.

Results

In the off-line, participants selected the second NP as the referent in discourses with an unstressed pronoun at an average proportion of .60. With contrastive stress, the proportion decreased to .36. The difference was statistically significant ($t_{47} = 5.18$, $p < .01$).

As with the previous on-line experiments, eight means were calculated, one for each condition. Average mean reaction times across all subjects for each condition are shown in Figure 10 below.
Figure 10. Mean reaction times of probes related to the first NP and control probes when presented 400 ms and 800 ms downstream.

Again, two repeated measure ANOVAs were calculated. The first was for two-factors (Prosody x Probe position) and included only control probe data. The analysis revealed a main effect of Prosody ($F_{(1,47)} = 4.655, \text{MSE} = 20,708, p = .04$) and a significant interaction ($F_{(1,47)} = 5.191, \text{MSE} = 16,875, p = .03$). A mean paired comparison for the condition 800 ms downstream showed that reaction times to control probes were significantly slower in discourses with a stressed pronoun versus an unstressed pronoun.

A three-factor ANOVA for a repeated measures revealed a significant main effect for Prosody ($F_{(1,47)} = 6.637, \text{MSE} = 23,157, p = .01$) and a main effect of Probe
type \((F_{1,47} = 5.496, \text{MSE} = 21,301, p = .02)\). There was also a significant interaction of Prosody x Probe position x Probe type \((F_{1,47} = 4.079, \text{MSE} = 17,550, p = .049)\).

For analyzing reactivation of the first NP, planned comparisons between related and control probes were administered for each probe site and sentence type. The results showed no significant priming for either comparison 400 ms downstream. However, 800 ms after an accented pronoun, a significant priming effect was observed \((F_{1,47} = 7.372, \text{MSE} = 31,719, p = .01)\). There were no other significant priming effects.

Two additional ANOVAs were calculated to determine whether the interactions between the positions 400 ms and 800 ms downstream were significant in each condition. For stressed pronouns, the interaction showed a trend \((F_{1,47} = 2.92, \text{MSE} = 12,626, p = .09)\). The interaction for unstressed pronouns did not approach significance.

**Discussion**

The results of Experiment 10 revealed a significant slowing 800 ms downstream from the pronoun, but only in conditions with a stressed pronoun. This pattern was the exact opposite of reaction time data for discourses with an unstressed pronoun. When no contrastive stress was present, processing became faster over time, as observed in other studies (that have not investigated prosody). Thus, the experiment provides evidence that contrastive stress adds to the processing load downstream from the pronoun.
A clue as to the exact nature of the processing at this later stage comes from priming effects. Like the second NP in the previous experiment, the first NP was reactivated 800 ms after the pronoun. This effect lends further support to the hypothesis that a second-pass interpretation is being formulated in the Integration stage. Both potential referents are being considered during this second round of integration, but only in those discourses with contrastive stress. Thus, it is the prosodic information that is triggering a reinterpretation of the referent.

**General Discussion**

The results from the five experiments presented in this paper provided evidence of when prosody, namely pitch accent as contrastive emphasis, interacts with the processing of pronouns. Experiment 6 contributed baseline data from adult subjects on the effect of contrastive stress on interpretations of object pronouns in parallel constructions. The findings indicated that the strategy of parallel function helps the listener to identify the antecedent; and that contrastive stress serves to override this pattern, biasing the listener toward the other potential referent. Experiments 7 and 8 showed that the prosodic cue constrains and licenses activation of specific referents at the moment the pronoun is heard. In parallel structures, only the referent with congruent grammatical function is accessed (a result that corroborates recent research on the immediacy of parallelism processing by Frisch (2003)). When contrastive stress is present, the prosodic cue blocks access to the entity with the parallel grammatical role while activating the referent. This happens at a very early
stage in processing. Following the time course of reactivation further downstream, Experiments 9 and 10 revealed priming effects for both potential referents 800 ms after the pronoun, but only in discourses with contrastive stress. The stress seems to initiate a reanalysis of the pronoun’s referent as indicated by increased processing load.

Taken together the results have important implications for theories of discourse anaphora and on-line language processing. First, prosody is an important factor in pronoun referent resolution and should be included in future theories. Second, the current findings offer more evidence that prosody is considered in a very early stage of processing. Most other studies have used intonation and boundary pauses in temporarily ambiguous syntactic structures, while the work presented here focused on pitch accent. In the experiments that have looked at pitch accent and pronouns, the materials have employed subject pronouns, which have been shown to introduce temporary coherence relation ambiguities. By using object pronouns in the experiments reported above, the ambiguity has already been resolved by the time the pronoun is reached. The results indicate that prosody is processed immediately at the pronoun.

Finally, the results of the experiments reveal important information about how contrastive stress is processed during on-line language comprehension. Importantly, contrastive stress should not be viewed as an additional step in the process of activating potential referents. Instead of first figuring out what the preferred referent would be in a parallel structure without stress and then reversing this interpretation
when stress is detected, a more accurate description is one in which candidates are
generated based on the prosodic cue directly. There is no evidence that an additional
step is introduced or that increased processing resources are required at the moment
the pronoun is heard as a result of contrastive stress. The delay happens in a later
phase of processing when reanalysis occurs.

There are several possible reasons why contrastive stress might trigger
reinterpretation in this context. First, the reanalysis could be the by-product of a
reassessment that happens in all instances of contrastive stress. Whenever canonical
prosodic patterns are overridden by contrastive stress, the system triggers a
mechanism to confirm the interpretation. Another possibility is that contrastive stress
does just this: ‘contrasts’ two entities. The prosody signals that this contrast needs to
take place, and the system activates both entities so that it can make a comparison
between them, as a response to the prosodic cue. This latter proposal makes sense
given that contrastive stress is most effective in switching interpretations in discourses
with Resemblance coherence relations. Future research could determine which of
these scenarios is true.

With regard to the time course of processing prosodic cues, generally, future
research efforts could corroborate the current findings by using other methodologies.
For example, the semantic eye tracking technique could be used with the materials in
these experiments, alleviating any confounds from using subject pronouns. The
experimental sentences used here also have more material after the pronoun so that
end-of-sentence effects do not interfere with processing several hundred milliseconds downstream from the pronoun.

What is clear is that prosody is an important factor in pronoun referent resolution and deserves a more prominent role in future theories of the on-line processing of pronouns. In the next chapter, I propose a new model of discourse anaphora that accounts for the time course of pronoun processing while acknowledging the important role of prosody.
Many theories about discourse anaphora and pronoun referent assignment have been devised by researchers across the years. Table 8 lists some of these theories, models and frameworks with a brief description of each proposal and its scope (this list is not exhaustive):

Table 8.

Theories, Models, and Frameworks of Anaphora.

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Theory, Model or Framework</th>
<th>Gist</th>
<th>Scope</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grosz &amp; Sidner</td>
<td>Centering Theory</td>
<td>Defines discourse structure and describes how salience</td>
<td>Centering</td>
</tr>
<tr>
<td>(1986)</td>
<td></td>
<td>of centers influences coherence.</td>
<td></td>
</tr>
<tr>
<td>Cahn (1995),</td>
<td>Independent proposals to</td>
<td>Describes how prosody works within the Centering Framework.</td>
<td>Centering / Prosody</td>
</tr>
<tr>
<td>Nakatani (1997),</td>
<td>integrate prosody into</td>
<td></td>
<td></td>
</tr>
<tr>
<td>and Beaver (2002)</td>
<td>Centering Theory</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author</td>
<td>Title</td>
<td>Description</td>
<td>Frameworks</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>----------------------------</td>
</tr>
<tr>
<td>Hobbs (1979)</td>
<td>Coherence-Driven Theory</td>
<td>Describes how pragmatic information drives interpretations.</td>
<td>Pragmatics</td>
</tr>
<tr>
<td>Smyth (1994)</td>
<td>Extended Feature Match Hypothesis</td>
<td>Explains the trade-off between Centering and Parallelism preferences through structural feature matching.</td>
<td>Centering / Parallelism</td>
</tr>
<tr>
<td>Kehler (2002)</td>
<td>Coherence, Reference and the Theory of Grammar</td>
<td>Challenges theories based on competition and identifies the coherence relation as the driving force behind preferences.</td>
<td>Centering / Parallelism / Pragmatics / Coherence</td>
</tr>
<tr>
<td>Ehrlich (1980)</td>
<td>Comprehension of Pronouns</td>
<td>Establishes gender as initial activation mechanism followed by pragmatics.</td>
<td>Agreement / Pragmatics</td>
</tr>
<tr>
<td></td>
<td>Coreference</td>
<td>Shows how candidate set of referents is generated based on Binding Theory.</td>
<td>Agreement / Binding</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------------------------</td>
<td>------------------------------------------------------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Nicol &amp; Swinney</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(forthcoming)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Badecker &amp; Straub (1995)</td>
<td>The Binding Theory as a Weighted Constraint Model</td>
<td>Refutes Nicol’s (1988) approach and suggests that Binding is only a weighted constraint (along with semantics, pragmatics, and discourse factors).</td>
<td>Binding / Centering / Pragmatics</td>
</tr>
<tr>
<td>Kintsch (1988)</td>
<td>Construction-Integration Model</td>
<td>Presents a model that reflects the time course of activating a candidate set of referents and then shedding away those that are not contextually relevant.</td>
<td>Agreement / Pragmatics</td>
</tr>
<tr>
<td>Sanford &amp; Garrod (1981)</td>
<td>Memory Focus Model</td>
<td>Explains why there is a difference between processing pronouns versus repeated names in a framework with Explicit and Implicit Focus.</td>
<td>Pronouns vs. names</td>
</tr>
</tbody>
</table>
The twelve approaches presented above address a range of factors involved in pronoun referent resolution including agreement features, Binding, Centering, Parallelism, pragmatics, and prosody. Unfortunately, none of the approaches account for all these different phenomena in one comprehensive, holistic model.

In what follows, I will present a unified theory of pronoun referent resolution in English called the Discourse Anaphora Weighting Approach (DAWA). The DAWA is based on linguistic as well as psycholinguistic data and thus models language processing as it might happen in the human mind. This approach contrasts attentional and coherence-driven theories that cannot produce an interpretation until after an entire utterance has been processed. However, the approach borrows heavily from previous work, which has already solved many of specific problems (for example, syntactic constraints for binding and salience rankings). The motivation for creating such an approach is twofold: one, to account for the different factors involved in pronoun referent resolution in a unified way; and two, to promote prosody, an often neglected factor, as a key component in the resolution of anaphora.

Discourse Anaphora Weighting Approach (DAWA)

The first assumption of the DAWA is that there is a verbal working memory buffer used by the language system to store incoming information. Several researchers have proposed architectures for this verbal memory store (Baddeley, 1986; Caplan & Waters, 1999; Just & Carpenter, 1992; Martin & Romani, 1994); however,
the architecture set forth here will combine several components of these proposals instead of settling on just one rendition. The functions that require memory and processing resources for anaphora resolution suggested here include the following:

Storage

- Temporarily store incoming acoustic information
- Store conceptual information extracted from the verbatim acoustic stream

Processing

- Process phonological information
- Process lexical-semantic information
- Process prosodic information
- Process syntax
- Process discourse
- Integrate propositions with contextual information and general world knowledge

Storage and Processing Resources

Based on these requirements, I propose two storage components (The components most likely do not reside in the same localized area of the brain, but the biological correlates of the architecture will not be explored here). The initial storage component, which I will call the Rich Verbatim Store (RVS), retains the exact form, syntax, prosodic information, and so forth of a phrase. This is most similar to
Baddeley’s phonological loop. The RVS itself retains information from the current clause, but this information decays rapidly. The other storage component, which I will refer to as the Conceptual Store (CS), keeps track of the entities within the discourse, their features, their relationships with each other, and pertinent information relevant to how these entities relate to the listener’s knowledge about the world. This component is comparable to Caplan & Water’s post-interpretive processing store. The rate of decay for this memory store is much slower. Also entities move in and out of focus through increases and decreases of activation levels. Outside the buffer is Long Term Memory (LTM) that represents an individual’s general knowledge. The LTM is conceptualized as an associative network as in Kintsch’s Construction-Integration Model.

Similar to Martin & Romani’s proposal, I suggest that there are several processing components as well: phonological (including segmenting words), lexical-semantic (looking up meanings of words and features associated with them such as gender or number), syntactic (building up syntactic structure), and propositional (integrating concepts introduced in the discourse with general world knowledge). In addition to those suggested by Martin & Romani, I include a processor for prosodic information, which can mark among other things pitch accent, and a processor for tracking discourse focus.

Equipped with the resources listed above, the act of resolving discourse anaphora begins. The general process of language comprehension can be represented in two phases: encoding information and integrating information. The specific process
of comprehending anaphora introduces two other steps: generating candidate antecedents and assessing whether the interpretations converge on a single candidate. The overview of the stages is illustrated in Figure 11. Each stage will be discussed in turn.

Figure 11. Stages of the Discourse Anaphora Weighting Approach.

Encoding

The early phases of processing involve extracting information from the RVS and encoding them into the CS. The subprocesses include phonological processing, lexical-semantic processing (including lexical tagging and LTM extraction), prosodic processing, and syntactic processing.

In phonological processing, individual words are identified from the sounds of the acoustic signal (word segmentation). In lexical-semantic processing, each word in
the RVS is tagged with feature information and is stored in a Lexical Feature Vector (LFV) in the CS, as shown in Figure 12.

Figure 12. Illustration of Encoding.

For nouns, features include such things as gender, number, animacy, and lexical form (pronominal vs. R-expression), as shown in Table 9.

Table 9.

Example of Lexical Feature Vector Information Tagged for Nouns.

<table>
<thead>
<tr>
<th>Incoming Word</th>
<th>Feature Label</th>
<th>Feature Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>John</td>
<td>Part of Speech</td>
<td>Noun</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td>Number</td>
<td>Singular</td>
</tr>
<tr>
<td></td>
<td>Animacy</td>
<td>Animate</td>
</tr>
<tr>
<td></td>
<td>Lexical Form</td>
<td>R-expression</td>
</tr>
</tbody>
</table>
For verbs, features include tense, argument structure and thematic structure (verb encoding will not be represented here). A word may bear multiple clusters of tags if it is ambiguous, regardless of the context of the utterance.

Concurrently, as tagging is taking place, each element is identified within the LTM store during the Encoding stage. The appropriate items are copied into the CS along with strongly associated neighbor nodes within the network, as in Kintsch’s Construction-Integration Model. This copying process is illustrated in Figure 13.

**Figure 13.** Illustration of mini-network in LTM copied into the CS for a given entity.

In copying the mini-networks into the CS, the connection weights between the nodes are also copied directly from the general knowledge network. The first-mentioned entity becomes the anchor for the discourse, with subsequent mini-networks from other entities and concepts linking to the anchor (in the spirit of Gernsbacher et al. (1989)’s structure-building framework).
Prosodic processing happens in parallel. Pitch accents over words or syllables are tracked such that elements are specifically marked for focus or contrast.

After a word has been tagged with its lexical and prosodic information, syntactic tagging takes place. The act of parsing the string of words and building up a syntactic structure occurs in the syntactic processor. The details of this mechanism go beyond the scope of this dissertation; however, many theories provide a detailed account of syntactic parsing (Frazier & Rayner, 1982; MacDonald, Pearlmutter & Seidenberg, 1994; Gibson, 1991). As an output of parsing, each entity is tagged with information about its grammatical function, and (as information about the verb becomes available), with its theta role. This syntactic information is stored in a Syntactic Feature Vector in the CS.

An entity is also tagged with two types of discourse information: (a) its salience value in the current utterance as described above, and (b) its history as the discourse topic. The topic of the previous sentence receives a T1 tag, the sentence before that T2, and the sentence before that T3. The history is quite short extending only a few sentences back. As new sentences are added to the discourse, the topic number is incremented. The same entity can retain several topic tags if it has been the topic for the past several sentences. Because the entity has already undergone syntactic tagging, the system can determine which entities occupy subject versus object position and which have been presented in the matrix clause versus embedded clauses. The salience of an entity is proportional to its rank as described in Brennan et al.’s (1987) algorithm for applying Centering theory to pronouns. This algorithm is consistent
with Matthews & Chodorow’s (1988) generalization of prominence of left-most positions compared to right-most, and shallow positions compared to deep. Of course, this information is not stored in the CS for the respective entity until it becomes available.

Once enough time has elapsed for the propositions in a clause to be tagged and copied from LTM to the CS, the acoustic information in the RVS starts to dissolve, freeing up resources for incoming phrases.

**Integrating**

For the propositional processing mechanism in which concepts are integrated with contextual information and general world knowledge, I will borrow from Kintsch’s (1988) Construction-Integration model. Here an inference mechanism generates additional connections to the nodes stored in the CS, based on a general understanding of the world and schemas for common activities such as flying in an airplane or swimming in a pool.

During the Integration process, activation spreads around the mini-network in the CS until the system stabilizes. Although rapid, the Integration process may be attempted several times if at first it fails to produce a coherent interpretation. The output of the Integration phase is a final interpretation based on activation values where the highest levels indicate those elements that are most important to the interpretation. Integration continues as new information is added to the CS through encoding.
Establishing the Coherence Relation

As a clausal boundary is reached (indicated by syntax and intonation), and as a new clause or sentence begins to populate RVS, discourse processing takes place in which the system establishes the coherence relation between the two clauses or sentences. The three classes of coherence relations that will be considered here are the three covered in Kehler’s (2002) book: Contiguity, Resemblance, and Cause-Effect. According to the DAWA, the default is Contiguity (Occasion) in which the speaker describes the events surrounding a central character or set of characters. The most salient entity is the subject, and future utterances that feature this entity as the topic result in the most coherent continuations. Resemblance incites a comparison between two entities and is often conveyed through parallel syntactic structure. Finally, Cause-Effect signifies a causal relation between the events in the discourse and its characters.

Although no rules exist which unconditionally dictate coherence relations, data from previous research points to consistent cues used by listeners to determine its relation. Again, Occasion is the default, so all cues must work to override it. In fact, even in the presence of strong cues that suggest otherwise, a decent percent (about 20 or so) of listeners will maintain the Occasion coherence relation irrespective of these other cues.

Resemblance is often conveyed by the conjunction “and” or the variations, “and then” or “and similarly” (Wolf et al., 2003). More importantly, the similarity of the syntactic structure to the previous clause indicates a Resemblance relation. As
Smyth (1994) showed, the more similar the two clauses (including parallel verbs, adjunct clauses and grammatical roles), the more listeners make interpretations based on parallelism. The syntactic processor provides for a mechanism to compare syntactic structure between clauses. The syntactic frame of the previous clause decays when syntactic processing is completed; however, the residue serves to facilitate processing if the structure of the subsequent clause is similar. The easier each item falls into place inside the residual syntactic frame of the previous clause, the more it denotes a Resemblance relation.

Since Occasion is the default and since Resemblance is most notably flagged via an utterance’s syntax, these two coherence relations are evident early in processing. In contrast, Cause-Effect requires more semantic and inferential integration.

For Cause-Effect relations, there are two prevalent cues. First is the connective *because*. The Cause-Effect implication inherent in semantics has received considerable support in empirical data. Several studies report different results in pronoun assignment in materials that do and do not use *because* (Ehrlich, 1980; McDonald & MacWhinney, 1995; Stevenson et al. 1994; Wolf et al. 2003). Another device is the semantics of the verb as is evident in the many studies investigating implicit causality (Caramazza et al., 1977; Caramazza & Gupta, 1979; Garnham et al., 1992; Garvey & Caramazza, 1974; McDonald & MacWhinney, 1995). Without these cues, the propositional processor must rely on general inference to extract causal relations between the entities and events conveyed in the discourse. Again, since
integration is necessary, interpretations based on Cause-Effect may not become evident immediately.

Although specific cues can constrain the set of coherence relation that can be inferred, these concepts do not have an explicit representation in the current model; rather they materialize through the organization of mechanisms and the final interpretation of the user.

As the new clause continues to unfold, the listener may encounter different types of discourse anaphora that must be resolved in order to construct meaning from the utterance. The following three sections address the different types of anaphora including reflexives, pronouns and anaphoric referring expressions.

**Reflexives**

As information is temporarily stored in the RVS, the phonological processor may identify reflexives such as *himself* in the acoustic stream. As with other nouns, the lexical-semantic processor tags the element with feature information including gender, number and animacy (in this example, male, singular, and animate, respectively). However, the processor cannot execute the additional step of isolating the element in the LTM store and copying it over into the CS. The reason for this is that the reflexive, in itself, is semantically empty and therefore does not possess its own semantic network. However, it does point back to an entity already in the discourse and therefore makes direct reference back to an already extracted mini-network inside the CS. Although LTM extraction does not take place directly, other
processing continues including prosodic and syntactic tagging. Next, the system must address the question of which entity is the reflexive’s antecedent.

The two steps in antecedent selection are (a) Generating potential candidate referents and (b) Assessing convergence.

**Generating Potential Candidate Referents for Reflexives**

The purpose of the Generating step is to identify entities in the current and previous clause that might be the antecedent of the reflexive. The first iteration involves entities only in the local discourse. This step takes about 200 ms and involves identification, activation and suppression mechanisms.

**Identification Mechanism 1: Feature Agreement**

When the Lexical Feature Vector of the anaphor has been populated, identify those entities that match the anaphor in all features (number, gender, animacy, etc.). If a feature is unknown, match entities with the features that are known. Also identify those entities whose feature values clash with the reflexive.

**Activation Mechanism 1a: Feature Agreement**

Activate the entities identified in Identification Mechanism 1 that have features that are congruent with the reflexive.
Suppression Mechanism 1b: Feature Agreement

Suppress the entities identified in Identification Mechanism 1 that have features that are incongruent with the reflexive.

Identification Mechanism 2: Binding

Identify the entities that bind the anaphor according to Chomsky’s definition of binding.

Activation Mechanism 2a: Principle A Adherence

Activate the entities that bind the anaphor.

Suppression Mechanism 2b: Principle A Adherence

Suppress the entities that do not bind the anaphor.

The Identification Mechanisms work in parallel and therefore the candidates generated from the syntactic search in Identification Mechanism 2 are not contingent on those found in Identification Mechanism 1. At first glance, then, it seems that Activation Mechanism 1a and Activation Mechanism 2a are redundant. Only one rule (the syntactic one) is necessary to settle on a single antecedent for the referent. The following is a justification for this parallel architecture.

Given claims by Badecker and Straub (1994) in which reading times are increased when a sentence contains two nouns of the same gender (even though only
one is syntactically licit as the antecedent of the reflexive), the data imply that both activation and suppression mechanisms are at work. One could argue that the candidates that match in gender are fed into Identification Mechanism 2. Passing two entities on to the next step might cause the increase in reading times. However, this architecture is not supported by the data: the same increase in reading times does not occur when subjects make correct interpretations of the sentence (Nicol, 1997) or when the entity is a possessive (Badecker & Straub, 1997 reported in Nicol & Swinney, (forthcoming)), as in the following:

(62) John thought that Beth’s cousin owed herself another opportunity to solve the problem.

If processing load is caused by passing two candidates to the next step instead of one, then gender ambiguity would cause a delay with possessives. Likewise, if binding constraints acted first, then gender ambiguity would never cause a delay because there would always be one and only one grammatically licit candidate passed to the next step. If, however, both processes are happening in parallel then the data can be justified. Two additional assumptions are required: (a) latencies result not only when passing two items to the next step but also when there are delays in mechanism completion times; (b) in the parallel process, when a single candidate emerges from either mechanism, that candidate is passed onto convergence immediately (as opposed to waiting for the other mechanism to complete). Although the mechanisms are
executed in parallel, agreement information becomes available sooner, in some cases, than syntactic activation/suppression. This phenomenon is most evident in sentence constructions such as the following:

(63) John thought that Ben owed himself another chance to solve the problem.

In this sentence, there are two subjects, one of the matrix clause and other of the relative clause. The mechanisms activating the correct subject and suppressing the incorrect one must decipher what the boundaries of the reflexive’s binding category are. As Nicol (1997) illustrated in her experiment in which incorrect responses tipped the scale of a statistically significant interaction, determining which is the sanctioned syntactic position can be tricky for participants, even without considering the on-line time course of the process. In these cases, where there are two subjects, the syntactic information takes more effort to resolve than agreement information, causing the increase in reading time observed in Badecker & Straub’s study with structures like (63) but not possessives. The possessive would be easier to resolve, so no extra processing time would be observed in completing the binding identification mechanism. The following figures and tables show three possible architectures and predictions of whether ambiguous gender should result in increased processing time. Only the model with parallel processing accurately represents the empirical data.
Figure 14. Illustration of Architecture 1. Candidates from Agreement Matching are fed into the mechanism that selects a candidate based on Binding Constraints.

Each step can add to the processing load. The amount of time is represented as three different variables: $\alpha$ for each candidate passed on from Mechanism 1 (agreement), $\beta$ for each candidate passed on from Mechanism 2 (binding), and $\chi$ for extra time needed to process structures with multiple subjects. To assess whether this model fits the data, each variable is assumed to equal one unit of time.
Table 10.

Units of Time Needed to Process Sentences for Architecture 1.

<table>
<thead>
<tr>
<th></th>
<th>Sentence with Two Subjects</th>
<th>Sentence with Possessive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ambiguous Gender</td>
<td>Unambiguous Gender</td>
</tr>
<tr>
<td>Gender Candidates</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>$\alpha$ * number of candidates</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Binding Candidates</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\chi$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Total Units of Time</td>
<td>4              $&gt;$ 3</td>
<td>3</td>
</tr>
<tr>
<td>Comment</td>
<td>Difference reflected in data.</td>
<td></td>
</tr>
</tbody>
</table>

In Table 10, the numbers represent blocks of time. A difference between two sentence types with and without gender ambiguity implies that a significant difference in
reading time will be observed in an experiment. With Architecture 1, the model does not accurately represent the data – no difference was observed for possessive sentences with and without gender ambiguity. The second architecture reverses the mechanisms such that only candidates generated from binding constraints are input into the agreement mechanism.

**Figure 15.** Illustration of Architecture 2. Candidates generated from Binding Constraints are fed into the mechanism that selects candidates based on Agreement Matching.
Table 11.
Units of Time Needed to Process Sentences for Architecture 2.

<table>
<thead>
<tr>
<th>Sentence with Two Subjects</th>
<th>Ambiguous Gender</th>
<th>Unambiguous Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binding Candidates</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(β * number of candidates)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Binding Computation (χ)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Gender Candidates</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(α * number of candidates)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Units of Time</td>
<td>3</td>
<td>= 3</td>
</tr>
<tr>
<td>Comment</td>
<td>Equal time NOT reflected in data.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sentence with Possessive</th>
<th>Ambiguous Gender</th>
<th>Unambiguous Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binding Candidates</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Binding Computation</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Gender Candidates</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Total Units of Time</td>
<td>2</td>
<td>= 2</td>
</tr>
<tr>
<td>Comment</td>
<td>Equal time reflected in data.</td>
<td></td>
</tr>
</tbody>
</table>

If the Binding Constraint mechanism occurs first and only generates one candidate, then gender ambiguity should never have an effect. However, this is not observed in the data. The final architecture shows how the two mechanism work in parallel.
In Architecture 3, as soon as a mechanism sufficiently activates a single candidate, then this candidate is selected as the referent. The system does not wait for a slow process to complete and does not pass more than one candidate onto the next step unless both mechanisms generate multiple candidates. Thus, the processing load is the lesser of the two amounts passed from each mechanism.
Table 12.

Units of Time Needed to Process Sentences for Architecture 3.

<table>
<thead>
<tr>
<th></th>
<th>Ambiguous Gender</th>
<th>Unambiguous Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sentence with Two Subjects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>Binding Candidates</td>
<td>Gender</td>
</tr>
<tr>
<td>Candidates</td>
<td>+</td>
<td>Candidates</td>
</tr>
<tr>
<td>Computation</td>
<td>2</td>
<td>1 + 1</td>
</tr>
<tr>
<td><strong>Total Units of Time</strong></td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Comment: Difference reflected in data.

<table>
<thead>
<tr>
<th></th>
<th>Ambiguous Gender</th>
<th>Unambiguous Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sentence with Possessive</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>Binding Candidates</td>
<td>Gender</td>
</tr>
<tr>
<td>Candidates</td>
<td>+</td>
<td>Candidates</td>
</tr>
<tr>
<td>Computation</td>
<td>2</td>
<td>1 + 0</td>
</tr>
<tr>
<td><strong>Total Units of Time</strong></td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Comment: No difference reflected in data.
The scenario in which parallel processing of feature agreement and binding constraints takes place complies better with the data, and therefore this approach has been incorporated into the DAWA. Feature agreement and binding are equally as important in resolving the pronoun in the model. If the candidate that matched in gender clashed with that for binding, the utterance would be registered as infelicitous.

The activation mechanisms embody a prime example of how the system exploits redundancy in order to validate its interpretation. Another instance of this is the complementary activation and suppression of the noun phrases in legal and illegal syntactic positions for coreference with the reflexive. Work by Nicol (1988), Gernsbacher (1989) and Badecker & Straub (1999) taken together suggest that both activation and suppression mechanisms alter the activation levels of the entities present in the CS.

Assessing Convergence of Reflexive Antecedents

In the Assessing Convergence step, the syntactic processor considers the candidates elicited in the Generating stage and decides which one is the antecedent of the reflexive (A similar Convergence stage will be used in examples with pronouns). The first step is to calculate the number of potential antecedent candidates that have reached a given threshold of activation after the applying the activation and suppression mechanisms.
Calculating the Number of Reflexive Antecedent Candidates

0, if the result is null, return to the Generating stage (adds to processing time).

1, if there is only one candidate, assign this candidate as the antecedent of the reflexive.

$X > 1$, if there is more than one candidate, keep the reflexive referent unresolved (adds to processing load). Go to Integrating phase.

Once an antecedent is chosen, the reflexive links directly to the conceptual information for its antecedent in the CS. In this sense the anaphor has a direct connection to the conceptual versus surface information as corroborated by results from Garrod et al. (1994) and Love & Swinney (1996).

If the system returned to the Generating stage because of a null result, it is possible that the subject of the binding category had not yet been reached (for example when the reflexive appears in a fronted embedded clause: “Himself included, John wanted the team to win the goal metal.”) Here, the reflexive would remain unassigned until more information became available in the acoustic stream.

Pronouns

When the phonological processor detects a pronoun such as her, the steps undertaken to uncover the antecedent are similar to that of reflexives, but they are more involved since, unlike reflexives, pronouns are free in their binding categories
and as such are much more flexible in terms of what they can refer to. The selection is not constrained to a special syntactic relation, as with reflexives.

When a pronoun is encountered, the lexical-semantic processor tags it with feature information including gender, number and animacy. Since pronouns are semantically impoverished and serve only as a variable for another entity, the step of searching through LTM to find the concept and copying it into the CS is bypassed and will be replaced with a direct link to the node of its antecedent. If the referent is a character outside the discourse, the pronoun points back to a mini-network created from an inference mechanism.

The pronoun is also tagged with prosodic and syntactic information. For prosodic tagging, two markings of note are focus and contrast.

After tagging, the system must take steps to resolve which entity is the pronoun’s antecedent and does so with the same stages identified for reflexives: (a) Generating potential candidate referents and (b) Assessing convergence. However, with pronouns, there is an additional sub-step between these stages that is dependent on coherence: (c1) Evaluating saliency or (c2) Evaluating parallelism.

**Generating Potential Candidate Referents for Pronouns**

First entities in the local discourse (current and previous clauses) become subject to activation and suppression mechanisms for generating referential candidates. This step takes about 200 ms as deduced by patterns in eye-tracking experiments that measure processing continuously (Arnold et al., 2000; Venditti et al.)
Again, the Identification Mechanisms for feature agreement and binding work in parallel.

Identification Mechanism 1: Feature Agreement

When the Lexical Feature Vector of the anaphor has been populated, identify those entities that match the anaphor in all features (number, gender, animacy, etc.). If a feature is unknown, match entities with the remaining features. Also identify those entities whose feature values clash with the pronoun.

Activation Mechanism 1a: Feature Agreement

Activate the entities identified in Identification Mechanism 1 that have features that are congruent with the pronoun.

Suppression Mechanism 1b: Feature Agreement

Suppress the entities identified in Identification Mechanism 1 that have features that are incongruent with the pronoun.

Identification Mechanism 2: Binding

Identify the entities that bind the anaphor according to Chomsky’s definition of binding.
Activation Mechanism 2c: Principle B Adherence
Activate the entities that do NOT bind the anaphor.

Suppression Mechanism 2d: Principle B Adherence
Suppress the entities that do bind the anaphor.

These two mechanisms are similar to those required for reflexive antecedent selection. The one difference, of course, is that the binding constraint must allow for the pronoun to be free in its binding category.

Assessing Convergence of Pronoun Antecedents
As with reflexives, the syntactic processor assesses whether the activation and suppression mechanisms of the previous Generating phase have converged onto a single referent. An illustration of the assessment process that determines whether an individual candidate moves onto the next stage (Assessing Convergence) is provided in Figure 17. In this illustration, $a$ is the activation level from Mechanism 1 (agreement), $b$ is the activation level for Mechanism 2 (binding), and $\theta$ is a threshold. Note, this process evaluates the activation and suppression levels from agreement and binding constrain mechanisms. This is different from the illustrations of processing load in the previous section. Since both agreement and binding mechanisms function in parallel, the processing undertaken in the assessment stage ensures that the proposed candidates do not violate agreement or binding constraints.
Since many entities could be viable options syntactically for the referent of a pronoun, other activation mechanisms must help to narrow down the selection (or strengthen the prospects of some of the candidates). For pronouns, the coherence relation in the discourse plays an important role in determining which other mechanisms will be utilized during processing.

**Calculating the Number of Pronoun Antecedent Candidates**

0, if the result is null, return to the Generating stage (adds to processing load).
If Resemblance (Weight value = 0.8, see below for explanation)

Go to Parallelism Mechanisms and then go to the Integrating phase.

If Occasion or Cause-Effect (Weight value = 1.0)

1, if there is only one candidate, assign this candidate as the antecedent of the pronoun. Check topic and reconcile shifts (see explanation below; adds to processing load). Go to Integrating phase.

X > 1, if there is more than one candidate, keep the pronoun referent unresolved (adds to processing load). Go through Salience Mechanisms and then go to the Integrating phase.

It is possible when dealing with pronouns (more so than reflexives) that the output of the activation mechanisms returns a null result. When this occurs, the search space must be expanded to include the global discourse and a second pass must be conducted. All the entities that were in the local discourse but failed to be activated are now suppressed. In addition, all entities that were in the CS but had been overshadowed by the more salient entities in the local discourse are brought into the foreground for the moment as the search for a referent continues. The same mechanisms function for the global discourse entities. If, after this second round, no referent is found, the system must search outside the discourse, which continues to add to the listener or reader’s processing load. When going outside the discourse, the
inference mechanism generates a mini-network in the CS, and the pronoun must point back to it and not an entity specifically mentioned in the discourse.

When the coherence relation is Resemblance, specific mechanisms for parallelism apply. Note the weight value is lower for Resemblance (0.8) than for Occasion (1.0). The weight value indicates the probability that a coherence relation will be selected given the appropriate cues. The weight values are based on data from previous experiments. For Resemblance, even though all the indicators of parallelism may be in place, the switch in referent interpretation from the default referent in subject position occurs less than 100 percent of the time, usually between 60 and 80 percent (Chambers & Smyth, 1998; also see data for Chapter 4). Given this observation, the weight value for the Resemblance path was set to 0.8. The probability is not whether the listener will select the Resemblance path versus the Occasion path, but rather the probability that the listener will go down a given path when all information suggests that that path is correct. If the coherence relation, is Occasion, listeners will go down this path 100% of the time; but when the coherence relation is Resemblance, listeners will only go down this path 80% of the time. As a side note, the weight values do not contribute to figuring the actual advantage of one candidate over the other. It simply indicates the probability that a given coherence relation will be chosen.

When the coherence relation is Occasion and the previous Generating stage produces only one candidate, this entity is assigned as the referent. If this referent has not been tagged with T1, as the previous topic of the discourse, then attention must be
shifted to this other entity. Eye-tracking reading time studies provide estimates of what this attention shift costs in terms of processing. For example, Garrod et al.’s (1994) data suggest that topic shifts present a 5 ms load per character slowing about 400 ms after the pronoun. Badecker & Straub (1994) also show longer latencies when the topic shifts compared to when it does not. Even though the system may have converged on a referent, the discourse processor still needs to maintain which entities are in the local focus and make shifts where necessary.

When more than one candidate is passed into the Assessing Convergence stage, this also increases the processing load. Garrod et al. (1994) shows a 3 or 4 ms latency increase per character when there are two candidates with the same gender. Extra processing for ambiguity has been shown in other studies as well (Boland, Acker and Wagner, 1998; Ehrlich, 1980; Hudson-D’Zmura & Tanenhaus, 1998; MacDonald & MacWhinney, 1990).

Evaluating Parallelism

Unlike other researchers who have considered multiple preference patterns and have devised theories based on competition (Kameyama, 1996; Stevenson et al. 1995), I adopt Kehler (2002)’s proposal that preferences result from the coherence relation of the pronoun’s clause with its preceding discourse. Thus, instead of an arbitrary hierarchy in which Centering preferences trump Parallelism but not general world
knowledge, the mechanisms in the DAWA only take effect when the relevant coherence relation has been established.\textsuperscript{10}

During syntactic processing, the system determines that syntactic structure for one clause is the same or similar to that of the previous clause. When enough cues lead to an interpretation based on Resemblance, specific mechanisms take effect, and prosodic tagging takes on an important role.

\textbf{When Coherence Relation is Resemblance}

Identification Mechanism 3: Parallel Grammatical Function

Identify the entity with the same grammatical function as the pronoun.

Likewise, identify the entity with the opposite grammatical function where subjects and objects are opposite.

Activation Mechanism 3a: Parallel Grammatical Function with Canonical Stress (Weight value = 0.8)

Activate the entity with the same grammatical function as the pronoun.

Suppression Mechanism 3b: Parallel Grammatical Function with Canonical Stress (Weight value = 0.8)

\textsuperscript{10} Although the DAWA adopts Kehler’s proposal that coherence relations are the underlying reason for differences in interpretations and not competition between preferences, Kehler takes the view that preferences are epiphenomenal side effects of the inference processes required to comprehend discourse. The DAWA cannot embody this claim since it is a model of anaphora and not general discourse processing.
Suppress the entity with the opposite grammatical function as the pronoun.

Activation Mechanism 3c: Parallel Grammatical Function with Contrast (Weight value = .5)
Activate the entity with the opposite grammatical function as the pronoun.

Suppression Mechanism 3d: Parallel Grammatical Function with Contrast (Weight value = .5)
Suppress the entity with the same grammatical function as the pronoun.

As mentioned previously, in any given experiment, the cues that indicate a different coherence relation than Occasion cannot bring participants to make that preference 100 percent of the time. Even less consistent are prosodic cues, and so these mechanisms tend to receive the lowest weights. Once the listener has decided that Resemblance is indeed the coherence relation, the mechanisms act like a syntactic reflex, with both activation and suppression at work. The resulting interpretation will either converge on one referent or result in an infelicitous interpretation.

The tendency to choose the element with identical grammatical function to the pronoun is so strong that it can trigger a P600 (Frisch, 2003) when an inconsistency
occurs with regard to gender. An example of this is the sentence, “Homer worships Marge and Lisa worships him too.” However, in some cases, parallelism might not be as strict, and the listener may be able to salvage the interpretation by switching back to an Occasion coherence relation in order to accommodate the gender cue (at the cost of additional processing).

Evaluating Saliency

The mechanisms for the default Occasion coherence relation work a bit differently than those for Resemblance because there is no suppression.

When Coherence Relation is Occasion

Identification Mechanism 4: Salience

Rank the salience of each according to grammatical function and depth of embedding according to Matthews and Chodorow (1988).

Activation Mechanism 4a: Salience with Canonical Stress

(Weight value = 1.0)

Activate each entity in the local discourse in proportion to its salience.

Activation Mechanism 4b: Salience with Contrast (Weight value = .2)

When the pronoun is tagged with contrast, promote the second-most salient entity over the first.
Activation Mechanism 4c: Salience with Focus Stress

(Weight value = .1)

When an entity is focus marked with a pitch accent, increase the activation of the focus-marked NP.

The following is a possible algorithm for assigning relative activation levels to candidates based on salience.

(64)  a. Assume that the lowest ranked entity is assigned x and the activation level for each rank above this increases by y. This is the candidate’s rank value.

b. Multiply the candidate’s rank value by 1 over the total number of candidates.

The algorithm in (64) assumes that activation increases monotonically with rank.

Mechanism 4b changes the rank of an entity based on prosodic information.

Note that Activation Mechanism 4b has a very low weight value. Although contrastive stress does sway participants to choose the “other” referent (Venditti et al. 2002), it does so only a small fraction of the time compared to the overwhelming choice of the subject in most cases. Note that this contrasts with Kameyama’s (1997) claim that stress always induces the selection of the “other” referent. Although this is
true most of the time when the coherence relation is Resemblance, this is not the case within Occasion discourses, where the effect influences about 20% of responses.

Activation Mechanism 4c accounts for the increased selection of potential referents marked with focus from pitch accent. Here again, the weight of the mechanism is quite small since the effect only increases the probability that participants will choose a particular referent. In Mechanism 4c, the rank order of the candidates is not changed, but the amount of activation for the focus-marked candidate is increased.

Note, that there are no suppression mechanisms when dealing with narratives. Thus, regardless of the candidate selected as the referent, none should be considered infelicitous. The output of the Evaluating Saliency mechanisms are the candidate referents with specific activation levels based on saliency. These are incorporated in the nodes in the CS and passed into the Integrating phase, which will attempt to disambiguate the pronoun.

**Integrating and Pronoun Resolution**

If a referent has already been selected for the pronoun, the Integrating phase validates that the entity’s action in the discourse makes sense. If Integrating uncovers a flaw in logic or implausible situation, then the processor may initiate several more cycles to try to resolve the confusion. For example, in Garrod et al.’s (1994) study, an incongruent verb caused an increase in reading times at the verb when the gender of the pronoun was unambiguous and the same referent was the expected topic of the
discourse. The latency happened almost immediately upon encountering the verb because no additional processing was spent on shifts of attention or gender ambiguities. The referent had already been selected and it was just a matter of integrating the referent into the context of the discourse at the verb.

The situation is a bit different when the referent has been left unresolved in the Assessing Convergence stage. Now, it is up to the Integrating phase to offer some insight into which candidate is the actual referent. The activation levels from the Evaluating Saliency step are important in defining preferences during the first few cycles of the Integrating phase. Thus, we see in Arnold et al.’s (2000) experiment, once the participants heard the verb after the ambiguous pronoun and validated that the most salient entity made sense with picture, the probability of eye fixations on the correct referent immediately increased. However, in the condition in which the picture did not confirm the expected referent based on saliency, eye fixation remained at the same level for both candidates, presumably as Integrating cycles continued. Consistent with this interaction between saliency and its confirmation through pragmatics, Matthews and Chodorow (1988) observed slower reading times when pragmatics forced the referent of the pronoun to be the entity that was not the most salient in the local discourse. It seems that saliency seeds the beginning cycles of the Integrating process, and when the expected referent is not validated from the context, more cycles are needed to resolve the ambiguity.

Thus, theta roles and pragmatics become important drivers of interpretation at this later point in processing. Taken from Kintsch’s model, an inference mechanism
generates links to the entities in the CS based on the context of the discourse and a general understanding of the world. For example, it knows that X might amaze Y because X holds certain character traits and generates links accordingly. Then spreading activation settles on which candidate is the best for the pronoun given the context.

**Anaphoric Referential Expressions**

Often a name previously mentioned in the discourse will appear again. The encoding and integrating functions work just the same: the entity is tagged, a copy from LTM is made, and the nodes in the CS are integrated into the discourse. This process often takes longer than extracting meaning from pronouns and reflexives because of the act of searching through LTM and copying nodes into the CS. Pronouns and reflexives simply link back to nodes already copied into the CS thereby skipping an expensive processing step. However, the system detects when it has made a duplicate copy during the Integrating phase because it keeps track of the characters important to the dialog and registers that two share the same characteristics. The nodes are assimilated into one, but this merger is expensive and costs extra processing time. The impact is only less noticeable when the entity’s representation in the discourse has decayed as a result of not being mentioned or referred to over time. In these cases the copied nodes do not have to be merged as much as reinstated into the CS memory buffer. Thus, the steps of copying and merging result in the repeated name penalty observed in many experiments.
Together, the phases of Encoding and Integrating, Generating Candidates and Assessing Convergence for reflexives, pronouns and anaphoric referring expressions constitute the DAWA. The test of the DAWA, of course, is its ability to account for the data reported in the literature. What follows are a series of examples from experiments mentioned in previous chapters illustrating how the DAWA works and accounts for previously reported findings.

Examples

Taking some examples from the psycholinguistic literature, here is how the DAWA handles pronoun resolution with gender cues, binding constraints, centering, parallelism, pragmatic preferences, and prosody. Once a potential antecedent has secured an advantage over another candidate, such advantage can be reflected in a psychological measurement.

Gender and Centering Preferences

In Arnold et al.’s (2000) article, gender is used to force pronoun interpretations in various discourses that explore differences in Centering preferences. Their materials consist of the following types of discourses:

(65) Donald is bringing some mail to Minnie while a violent storm is beginning. He’s carrying an umbrella …
When readers first encounter *Donald*, Encoding begins.

<table>
<thead>
<tr>
<th>Incoming Word</th>
<th>Feature Label</th>
<th>Feature Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Donald</td>
<td>Lexical Feature Vector</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Part of Speech</td>
<td>Noun</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td>Number</td>
<td>Singular</td>
</tr>
<tr>
<td></td>
<td>Animacy</td>
<td>Animate</td>
</tr>
<tr>
<td></td>
<td>Lexical Form</td>
<td>R-expression</td>
</tr>
<tr>
<td></td>
<td>Syntactic Feature Vector</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Grammatical Function</td>
<td>Subject</td>
</tr>
<tr>
<td></td>
<td>Clause</td>
<td>Matrix</td>
</tr>
<tr>
<td></td>
<td>Discourse Feature Vector</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Focus Role</td>
<td>T1</td>
</tr>
<tr>
<td></td>
<td>Salience</td>
<td>Subject</td>
</tr>
</tbody>
</table>

A node for *Donald* is extracted from LTM and is stored in the CS. Encoding continues with *Minnie*.

<table>
<thead>
<tr>
<th>Incoming Word</th>
<th>Feature Label</th>
<th>Feature Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minnie</td>
<td>Lexical Feature Vector</td>
<td></td>
</tr>
</tbody>
</table>
Another network is copied into the CS. With regard to the coherence relation, there is nothing to bring the reader from departing from the default of Occasion at this point. Processing continues. Then, the reader encounters the pronoun *He*. The pronoun also undergoes Encoding.

<table>
<thead>
<tr>
<th>Incoming Word</th>
<th>Feature Label</th>
<th>Feature Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>He</td>
<td>Lexical Feature Vector</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Part of Speech</td>
<td>Pronoun</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td>Number</td>
<td>Singular</td>
</tr>
</tbody>
</table>
The focus role of *Donald* increments from T1 to T2 now that a new sentence has begun.

After the pronoun is tagged, the system must figure out which referent it is referring to. The first step is Generating Potential Candidate Referents for Pronouns. Each candidate’s activation level is assessed as illustrated in Figure 17. Note that the numerical values are simply a tool to show the relative activation of the candidate antecedents. For simplicity, activation values equal 1.0 and suppression values equal –1.0. The threshold ($\theta$) for moving on to the Assessing Convergence stage is 0. The numbers used below in no way reflect the actual reaction times or selection probabilities.

Identification Mechanism 1: Feature Agreement

*He* = Male; *Donald* = Male; *Minnie* = Female
\[ He = Donald \]

\[ He \neq Minnie \]

Activation Mechanism 1a: Feature Agreement

\[ Donald = 1.0 \]

Suppression Mechanism 1b: Feature Agreement

\[ Minnie = -1.0 \]

Identification Mechanism 2: Binding

\[ Donald \text{ does not bind } He. \]

\[ Minnie \text{ does not bind } He. \]

Activation Mechanism 2c: Principle B Adherence

\[ Donald = 1.0 \]

\[ Minnie = 1.0 \]

Suppression Mechanism 2d: Principle B Adherence

N/A

**Total Activation Value** \((a + b)\) (After about 200 ms)

\[ Donald = 1.0 \text{ (agreement)} + 1.0 \text{ (binding)} = 2.0 \]

\[ Minnie = -1.0 \text{ (agreement)} + 1.0 \text{ (binding)} = 0 \]

According to the DAWA, after about 200 ms, there is a clear advantage for *Donald* as the referent of the pronoun *He*. The DAWA predicts that eye fixations would begin to
increase in probability to Donald at this stage in processing, which is exactly what is observed in the empirical data.

Since only one candidate has a value above 0, this candidate is assigned as the antecedent of the pronoun. Since Donald has been assigned T2 and He T1, and since Donald = He, then there is no shift in attention. The Integrating phase confirms that Donald is indeed the one carrying an umbrella and so no second pass analyses need take place.

When the materials present both Donald and Minnie again, but use She as the pronoun, the totals activation values are flipped:

\[
\text{Donald} = -1.0 \text{ (agreement)} + 1.0 \text{ (binding)} = 0
\]

\[
\text{Minnie} = 1.0 \text{ (agreement)} + 1.0 \text{ (binding)} = 2.0
\]

After 200 ms, it is clear who the referent of the pronoun is here as well, and, as predicted, eye fixations gravitate toward the target referent, Minnie. One difference between this scenario and the previous one is that Minnie has not been assigned a focus role. Thus, a shift in focus must take place between Donald and Minnie. Although the eye fixation measurement is not sensitive to any difference between the two conditions aside from the target referent, the DAWA would predict a slowing a few hundred milliseconds after the pronoun because of the shift in attention.
Arnold et al. also explored materials with ambiguous pronouns. Instead of the female character, *Minnie*, the discourse presents *Mickey* along with *Donald Duck*. The following are the mechanisms triggered in the DAWA for these discourses.

Identification Mechanism 1: Feature Agreement

\[ He = \text{Male}; \ Donald = \text{Male}; \ Mickey = \text{Male} \]

\[ He = Donald \]

\[ He = Mickey \]

Activation Mechanism 1a: Feature Agreement

\[ Donald = 1.0 \]

\[ Mickey = 1.0 \]

Suppression Mechanism 1b: Feature Agreement

N/A

Identification Mechanism 2: Binding

*Donald* does not bind *He*.

*Mickey* does not bind *He*.

Activation Mechanism 2c: Principle B Adherence

\[ Donald = 1.0 \]

\[ Mickey = 1.0 \]

Suppression Mechanism 2d: Principle B Adherence

N/A
**Total Activation Value** (After about 200 ms)

*Donald* = 1.0 (agreement) + 1.0 (binding) = 2.0

*Mickey* = 1.0 (agreement) + 1.0 (binding) = 2.0

Both candidates are active after about 200 ms. Since the coherence relation is Occasion, processing continues in Evaluating Saliency. Since activation values of the two candidates are the same going into Evaluating Saliency, the values resulting from this step will be the relative difference in activation between the two. As described in Mechanism 4, the values are based on the salience of each candidate. Using the algorithm in (64), with $x = 1.0$ and $y = 1.0$, the following relative activation values result:

Identification Mechanism 4: Salience

*Donald* = Subject

*Mickey* = Indirect Object

Activation Mechanism 4a: Salience with Canonical Stress

*Donald* = 1.0

*Mickey* = 0.5
Total Activation Value going into the Integrating Phase

Donald = 1.0 (salience) = 1.0
Mickey = 0.5 (salience) = 0.5

Here, there is an advantage for the first-mentioned entity, but there has been no commitment of which is the referent as was the case with differentiating gender cues. After 200 ms, the DAWA predicts that there will be a preference for Donald. This slight preference for the first-mentioned character is apparent in MacDonald and MacWhinney’s (1990) data. Since both referents are activated, both are candidates in the Integrating phase.

In Arnold et al.’s study, visual context offers a clear tool for disambiguating the pronoun. The story describes how “He’s carrying …” something. Connections between each candidate’s node in the CS and a new node for carrying are created. Then the Integrating cycles begin. In one condition, the picture shows that Donald is carrying two things (and umbrella and a mailbag), while Mickey is not carrying anything at all. For participants viewing this picture, the Integrating cycle quickly determines that it must be Donald who is the referent of the pronoun. Since Donald is the most salient of the two candidates, spreading activation stabilizes quickly, since Donald already has a head start from the discourse context.

In another condition, however, each character is carrying something. Donald is carrying a mailbag, and Mickey is carrying an umbrella. In this case, the Integrating cycles cannot converge on a single referent quickly and the pronoun remains
unresolved until further information becomes available. The data reflect this split between the two candidates by showing equal eye fixations to each character. Although, Mickey receives more eye fixations about 1200 ms after the pronoun, the difference between the two is still not large.

The extent to which the system makes an assumption about the most salient candidate depends on the strength of the salience. This is demonstrated clearly in Arnold et al.’s second experiment in which the topic of the previous two sentences \((Donald \text{ and } He = Donald)\) changes in the third \((He = Mickey)\), unbeknownst to the listener. In this condition, the referent was assumed to be the same as the topic in the previous sentences. Evidence for this comes from eye fixations to the wrong referent even as far downstream as 1200 ms. According to the DAWA, both Donald and Mickey are passed to the Assessing Convergence stage. However, in the Evaluating Saliency step, only Donald receives an activation score since only Donald appeared in the previous sentence:

Identification Mechanism 4: Salience

\(He = Donald = \text{Subject}\)

Activation Mechanism 4a: Salience with Canonical Stress

\(Donald = 1.0\)
Total Activation Value going into the Integrating Phase

Donald = 1.0 (salience) = 1.0

Mickey = 0 (salience) = 0

As a result, Donald is much more prominent in the first cycles of the Integrating phase (as is reflected in the high number of eye fixations); however, after a while, inconsistent information in the picture brings the Integrating phase to suppress Donald as the correct referent choice.

Prosody and Parallelism Preferences

When the coherence relation is Resemblance, other activation and suppression mechanisms take effect. The examples showing prosody and parallelism using the DAWA will be from Chapter 4.¹¹

The following is a target sentence within an example discourse:

(66) A cowboy pushed a robber into the bar, and a waiter pushed him into the poker table.

Taking this discourse through the DAWA, the listener first starts to hear the story and begins the Encoding process with the word *cowboy.*

¹¹ Venditti’s experiment (the only other on-line study dealing with prosody and parallelism) confounds the effects of Narrative and Resemblance at the point of the pronoun.
A node for *cowboy* is identified in LTM and copied into the CS. Then, the phonological processor encounters the word *robber*, which is encoded as follows:

<table>
<thead>
<tr>
<th>Incoming Word</th>
<th>Feature Label</th>
<th>Feature Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>robber</td>
<td>Lexical Feature Vector</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Part of Speech</td>
<td>Noun</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td>Number</td>
<td>Plural</td>
</tr>
<tr>
<td></td>
<td>Animacy</td>
<td>Animate</td>
</tr>
<tr>
<td></td>
<td>Lexical Form</td>
<td>R-expression</td>
</tr>
<tr>
<td></td>
<td>Syntactic Feature Vector</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Grammatical Function</td>
<td>Subject</td>
</tr>
<tr>
<td></td>
<td>Clause</td>
<td>Matrix</td>
</tr>
<tr>
<td></td>
<td>Discourse Feature Vector</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Focus Role</td>
<td>T1</td>
</tr>
<tr>
<td></td>
<td>Salience</td>
<td>Subject</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Incoming Word</th>
<th>Feature Label</th>
<th>Feature Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cowboy</td>
<td>Lexical Feature Vector</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Part of Speech</td>
<td>Noun</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td>Number</td>
<td>Plural</td>
</tr>
<tr>
<td></td>
<td>Animacy</td>
<td>Animate</td>
</tr>
</tbody>
</table>

Table: Lexical Feature Vector

<table>
<thead>
<tr>
<th>Incoming Word</th>
<th>Feature Label</th>
<th>Feature Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cowboy</td>
<td>Lexical Feature Vector</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Part of Speech</td>
<td>Noun</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td>Number</td>
<td>Plural</td>
</tr>
<tr>
<td></td>
<td>Animacy</td>
<td>Animate</td>
</tr>
<tr>
<td></td>
<td>Lexical Form</td>
<td>R-expression</td>
</tr>
<tr>
<td></td>
<td>Syntactic Feature Vector</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Grammatical Function</td>
<td>Subject</td>
</tr>
<tr>
<td></td>
<td>Clause</td>
<td>Matrix</td>
</tr>
<tr>
<td></td>
<td>Discourse Feature Vector</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Focus Role</td>
<td>T1</td>
</tr>
<tr>
<td></td>
<td>Salience</td>
<td>Subject</td>
</tr>
</tbody>
</table>
Another node is copied into the CS. Next, the system encounters *waiter*, which is also tagged:

<table>
<thead>
<tr>
<th>Incoming Word</th>
<th>Feature Label</th>
<th>Feature Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>waiter</td>
<td>Lexical Feature Vector</td>
<td></td>
</tr>
<tr>
<td>Part of Speech</td>
<td>Noun</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>Male</td>
<td></td>
</tr>
<tr>
<td>Number</td>
<td>Singular</td>
<td></td>
</tr>
<tr>
<td>Animacy</td>
<td>Animate</td>
<td></td>
</tr>
<tr>
<td>Lexical Form</td>
<td>R-expression</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Syntactic Feature Vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammatical Function</td>
</tr>
<tr>
<td>Clause</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Discourse Feature Vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focus Role</td>
</tr>
<tr>
<td>Salience</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Syntactic Feature Vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammatical Function</td>
</tr>
</tbody>
</table>
The verb as well as the syntactic structure is exactly the same as in the first clause.
Thus, the listener should assume a Resemblance coherence relation (80 percent of the
listeners will do so). And finally, the listener hears *him*:

<table>
<thead>
<tr>
<th>Incoming Word</th>
<th>Feature Label</th>
<th>Feature Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>him</td>
<td>Lexical Feature Vector</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Part of Speech</td>
<td>Pronoun</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td>Number</td>
<td>Singular</td>
</tr>
<tr>
<td></td>
<td>Animacy</td>
<td>Animate</td>
</tr>
<tr>
<td></td>
<td>Lexical Form</td>
<td>Pronoun</td>
</tr>
<tr>
<td></td>
<td>Syntactic Feature Vector</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Grammatical Function</td>
<td>Object</td>
</tr>
<tr>
<td></td>
<td>Clause</td>
<td>Matrix (conjoined)</td>
</tr>
<tr>
<td></td>
<td>Discourse Feature Vector</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Focus Role</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>Salience</td>
<td>Object</td>
</tr>
</tbody>
</table>
Now, the system begins identifying the referent by generating potential candidates.

The identification, activation, and suppression mechanisms output the following:

Identification Mechanism 1: Feature Agreement

\[
\begin{align*}
\text{him} & = \text{Male}; \text{cowboy} = \text{Male}; \text{robber} = \text{?(male)}; \text{waiter} = \text{Male} \\
\text{him} & = \text{cowboy} \\
\text{him} & = \text{robber} \\
\text{him} & = \text{waiter}
\end{align*}
\]

Activation Mechanism 1a: Feature Agreement

\[
\begin{align*}
\text{cowboy} & = 1.0 \\
\text{robber} & = 1.0 \\
\text{waiter} & = 1.0
\end{align*}
\]

Suppression Mechanism 1b: Feature Agreement

N/A

Identification Mechanism 2: Binding

\[
\begin{align*}
\text{cowboy} & \text{ does not bind } \text{him}. \\
\text{robber} & \text{ does not bind } \text{him}. \\
\text{waiter} & \text{ does bind } \text{him}.
\end{align*}
\]
Activation Mechanism 2c: Principle B Adherence

cowboy = 1.0

robber = 1.0

Suppression Mechanism 2d: Principle B Adherence

waiter = -1.0

Total Activation Value

cowboy = 1.0 (agreement) + 1.0 (binding) = 2.0

robber = 1.0 (agreement) + 1.0 (binding) = 2.0

waiter = 1.0 (agreement) + -1.0 (binding) = 0

Both cowboy and robber are passed into the Assessing Convergence phase. Now since the coherence relation is Resemblance, the mechanisms in Evaluate Parallelism are executed:

Identification Mechanism 3: Parallel Grammatical Function

cowboy = opposite grammatical function

robber = parallel grammatical function

Activation Mechanism 3a: Parallel Grammatical Function with Canonical Stress

robber = 1.0
Suppression Mechanism 3b: Parallel Grammatical Function with Canonical Stress

\[ cowboy = -1.0 \]

**Total Activation Value going into Integration Phase**

\[ cowboy = -1.0 \text{ (grammatical function) } = -1.0 \]

\[ robber = 1.0 \text{ (grammatical function) } = 1.0 \]

Given the Resemblance coherence relation, the second NP, \textit{robber}, should be activated as the referent, very soon after the pronoun is heard. Parallelism acts quickly, like a syntactic reflex, even before the Integrating phase. And as predicted, the data show a priming effect for the second NP after the pronoun, but not for the first NP.

Now, when contrastive stress is realized on the pronoun, different mechanisms are administered:

Identification Mechanism 1: Feature Agreement

Same as above.

Identification Mechanism 2: Binding

Same as above.
Identification Mechanism 3: Parallel Grammatical Function

cowboy = opposite grammatical function

robber = parallel grammatical function

 Activation Mechanism 3c: Parallel Grammatical Function with

Contrast

cowboy = 1.0

Suppression Mechanism 3b: Parallel Grammatical Function with

Contrast

robber = -1.0

Total Activation Value going into Integrating Phase

cowboy = 1.0 (grammatical function) = 1.0

robber = -1.0 (grammatical function) = -1.0

With contrastive stress, the first NP is more prominent than the second, as the on-line and off-line results confirm. The effect is observed almost immediately since it happens before the Integrating phase.

Prosody and Pragmatics

With interpretations that depend on pragmatics, the activation levels of the potential referents soon after the pronoun should be similar (with a slight preference for the more salient candidate). Then, several hundred milliseconds later, after
Integrating, the contextually relevant referent should take prominence. This is exactly what Gernsbacher (1989) reported in her fourth experiment, which used materials such as the following:

(67) Bill lost a tennis match to John. Accepting the defeat, he walked quickly toward the showers.

(68) Bill lost a tennis match to John. Enjoying the victory, he walked quickly toward the showers.

About 200 ms after the pronoun, both candidates should be accessible with the first-mentioned enjoying a slight advantage.

**Total Activation Value**

Bill = 1.0 (salience) = 1.0

John = 0.5 (salience) = 0.5

Only at the end of the sentence does the referent that is consistent with the pragmatic content of the story show a significant lead over of the nonreferent. This is true even though the relevant information occurs before the pronoun. This shows that the identification, activation, and suppression mechanisms are executed before the Integrating phase begins.
When focus stress is added to the mix, other mechanisms are used. The following is an example discourse from Chapter 3:

(69) Wendy was taking a cruise to Alaska during her spring break. As she was going to her room to grab a book, she noticed a paramedic calling the captain down the hall. Later Wendy bumped into him and asked if everything was okay.

The DAWA mechanisms result in the following relative scores when listeners come upon the pronoun *him*:

**Total Activation Value After Generating Phase**

- **paramedic** = 1.0 (agreement) + 1.0 (binding) = 2.0
- **captain** = 1.0 (agreement) + 1.0 (binding) = 2.0
- **Wendy** = -1.0 (agreement) + -1.0 (binding) = -2.0

Wendy does not move into the Assessing Convergence phase because the activation values for this referent are not above 0. After the Evaluating Saliency phase, the entities would have the following relative activation values (again numerical values are only for illustration purposes, where the incresae in activation for focus stress equals 0.3).
Total Activation Value with Focus Stress on 1\textsuperscript{st} NP

\begin{align*}
\text{paramedic} &= 1.0 \text{ (salience)} + 0.3 \text{ (focus stress)} = 1.3 \\
\text{captain} &= 0.5 \text{ (salience)} = 0.5
\end{align*}

Although both \textit{paramedic} and \textit{captain} are accessible, there is a preference for \textit{paramedic} since its grammatical function elevates its salience in the discourse and since focus stress augments it. When focus stress is placed on the second NP, \textit{captain}, the salience for this NP is amplified:

Total Activation Value with Focus Stress on 2\textsuperscript{nd} NP

\begin{align*}
\text{paramedic} &= 1.0 \text{ (salience)} = 1.0 \\
\text{captain} &= 0.5 \text{ (salience)} + 0.3 \text{ (focus stress)} = 0.8
\end{align*}

At this point the Assessing Convergence stage would establish that both possibilities should remain open. In the Integrating phase where the Patient of the action is promoted as a potential referent of the pronoun, the high activation value of the second NP pushes it over the first NP in some cases, resulting in a higher probability of second NP being chosen when delivered with focus stress.

In discourses that present the potential referents in a passive construction, the first NP is the Patient, and therefore, the small increase in salience from focus stress would not make a difference in the interpretation – without the theta role helping to
overcome the first NP’s slightly higher activation level, the first NP remains the referent of choice in these sentence constructions.

Conclusion

Since the psychological data often present contradictions with regard to when different effects take place, it is difficult to make claims about precisely when certain effects happen. For example, some articles report a delay in the availability of gender information, while others show a clear influence immediately after the pronoun. However, there are patterns across many studies, and these suggest that feature agreement information is available immediately as are binding constraints (when grammatical function differs between participants) and parallelism preferences based on syntax. Prosodic cues also seem to influence activation immediately. These trends are all embodied in the mechanisms of the DAWA.

There are several important differentiators between the DAWA and other theories. First, the approach does not rely on artificial hierarchies, and there are no rules of competition between Centering preferences, parallelism, and pragmatics because the mechanisms change with the coherence relations. Looking at all the examples presented above, the DAWA accounts for the data through identification, activation, and suppression mechanisms, which are all sensitive to coherence relations.

The DAWA also provides an explanation of how prosody interacts with discourses with different coherence relations. Pitch accent, for example, has a very
different pattern of influence in Occasion settings versus those with Resemblance, and pitch accent that conveys contrastive stress affects interpretations in very different ways than focus stress.

Finally, the different aspects of a referent’s suitability including gender agreement, binding constraints, and discourse cues such as first-mention preferences and parallelism all become available around the same time. That said, there is probably a cascading effect in the completion of the mechanisms – agreement feature matching and binding constraints probably become available first, with discourse factors occurring slightly later. Of course pragmatic information follows all of these in the Integrating phase of the DAWA. Although such subtle cascading before Integrating is not evident in something like a picture eye fixation task (which usually depicts low but equal fixation probabilities of both candidates at the pronoun, and then increasing glances at the preferred referent about 200 ms after the pronoun), there is indication of this cascading from processing load experiments, in particular reading time studies. Here, gender ambiguities introduce latencies immediately at the pronoun (in some studies) and shifts in discourse focus cause slower reaction times at the following verb.

Now that the influences of prosody on pronoun referent resolution have been incorporated into an overall framework of the real-time representation of discourse anaphora, concrete predictions can be tested. For example, other methodologies such as eye tracking and ERPs could be used to show prosody’s immediate effect on pronoun referent activation. For eye tracking, similar materials could be used as in
Chapter 4. For ERPs, different discourses in which gender forces coreference with anomalous referents based on the pitch accent could be developed with the expectation that P600s would result. With regard to focus stress, more examples of pitch accents on referents could be explored, especially outside the Occasion coherence relation. Additionally, the subtle timing differences between agreement features, binding, discourse effects, and prosody put forth in the DAWA could be tested further. For instance, processing load could be measured for parallel structures with ambiguous gender. The DAWA predicts that processing load would increase because agreement information is available slightly sooner than parallelism preferences. Clearly more experimentation is needed to uncover nuances in timing, and the theory will either account for such findings or require revisions.

Even so, the DAWA provides a framework for real-time discourse anaphora resolution with sensitivity towards coherence relations and an acknowledgement of prosody’s important role. The hope is that it will continue to offer a comprehensive representation of how referents are resolved in discourse and drive future research on the topic.
## APPENDIX A: MATERIALS FOR EXPERIMENTS IN CHAPTER 3

<table>
<thead>
<tr>
<th>Experiment 1 (active)</th>
<th>Experiment 2 (passive)</th>
<th>Experiment 3 (conjunction)</th>
<th>Experiment 4 (source/goal)</th>
<th>Experiment 5 (goal/source)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a Karen knew some of her friends were going to this year's Halloween party.</td>
<td>1b When she arrived, she scanned the room and saw a sheriff serving a pirate near the punch bowl.</td>
<td>1c Right away Karen recognized him and went over to say &quot;hi&quot;.</td>
<td>2a On Margaret's first day on the job she was introduced to everyone on her floor.</td>
<td>2b During lunch she noticed an executive questioning a scientist in the kitchen.</td>
</tr>
<tr>
<td>2c Later in a meeting, Margaret addressed him by name and impressed the others.</td>
<td>3a Dr. Sullivan was traveling to Europe on a very busy flight from New York.</td>
<td>3b During the flight, he noticed a stewardess counseling a nurse at the front of the cabin.</td>
<td>3b During the flight, he noticed a stewardess counseling a nurse at the front of the cabin.</td>
<td>3b During the flight, he noticed a stewardess counseling a nurse at the front of the cabin.</td>
</tr>
</tbody>
</table>
3c Dr. Sullivan called her over to see if there was an emergency.

4a When her parents were visiting, Candace took them to the old village where people dressed up like characters from centuries ago.

4b Just as everyone was getting hungry, she saw a physician approaching a saint near a stone cottage.

4c Sarah quickly went over to him and asked where the nearest café was.

5a Rebecca realized that some of the recipients of the prestigious awards ceremony were from her home town.

5b During the ceremony, she saw an orthodontist congratulating an instructor near the front stage.

5c After the presentation Rebecca approached him and mentioned that she had gone to the same school.

6a Rita was at the space center doing some research for her final project.

6b She overheard a father quizzing an astronaut near the anti-gravity room.
6c Rita decided to go over to him and catch the tail end of the conversation.

7a Linda made sure she attended the career fair at school.

7b During the fair, she saw a pilot swatting a dentist near a booth wall.

7c Linda decided to approach him and ask about career opportunities.

8a Helen loved all sports so she was very excited when she got tickets to see the Indy 500.

8b Before the race, she noticed a photographer encouraging a racer in one of the buildings.

8c Helen apologized to him for interrupting but then asked to stay and chat.

9a Ann was put in charge of coordinating the several million-dollar wedding.

9b Everything was going smoothly, when she came across a caterer scolding the groom near the roasted duck.

9c Everything was going smoothly, when she came across a caterer scolding the groom near the roasted duck.
9c Ann thought something was wrong so she went over to him and asked if everything was okay.

| 10a | Steven was in town visiting and he decided to take his brother's kids with him to the bank. |
| 10b | As he was standing in line, he saw his niece entertaining a teller near the waiting area. As he was standing in line, he saw his niece being entertained by a teller near the waiting area. As he was standing in line, he saw his niece teaching a magic trick to a teller near the waiting area. As he was standing in line, he saw his niece learning a magic trick from a teller near the waiting area. |
| 10c | When Steven was done, he went over to her and smiled. |

| 11a | At the race track, Ellen never knew who to bet on and always picked horses based on their names. |
| 11b | Before the next race, she overheard a programmer advising a jockey near the snack bar. Right before the next race, she overheard a programmer being advised by a jockey near the snack bar. Right before the next race, she noticed a programmer and a jockey near the snack bar. Right before the next race, she overheard a programmer offering a calculator to a jockey near the snack bar. Right before the next race, she overheard a programmer taking a calculator from a jockey near the snack bar. |
| 11c | Ellen was tempted to go over to him and ask about betting statistics. |

| 12a | After graduating, Kristina decided to go to the annual children's book trade show in New York to make some contacts. |
| 12b | In the reception area, she overheard a critic recommending a publisher during the coffee break. In the reception area, she overheard a critic being recommended by a publisher during the coffee break. In the reception area, she saw a critic hand an autographed book to a publisher. In the reception area, she saw a critic receive an autographed book from a publisher. |
At the next break, she asked him for advice on how to get a foot in the door.

Every year Alisha went to the golf tournament at the resort in the foothills.

During the game, Alisha noticed her brother being questioned by a caddie in the sand pit.

Later, at the reception, she saw him and asked whether there were any extra tickets to the final tournament.

At the car repair shop Deborah was waiting in line behind a couple who looked quite irritated.

She couldn't help but notice the husband belittling the mechanic near the counter.

Deborah started getting concerned, so before it was her turn, she asked him if there was anything she should know before turning her car in.

Greg always grew impatient with the lines in the ski lodge snack bar.

One day as he was waiting in line, he saw a cashier lecturing a skier in front of a spilled beer.

Greg gave her a napkin because he wanted to get out on the trails before nightfall.
16a  Wendy was taking a cruise to Alaska during her spring break.

16b  As she was going to her room to grab a book, she noticed a paramedic calling the captain down the hall.  
     As she was going to her room to grab a book, she noticed a paramedic being called by the captain down the hall.  
     As she was going to her room to grab a book, she noticed a paramedic and the captain down the hall.  
     As she was going to her room to grab a book, she noticed a paramedic throwing a life vest to the captain down the hall.  
     As she was going to her room to grab a book, she noticed a paramedic catching a life vest from the captain down the hall.

16c  Later Wendy bumped into him and asked if everything was okay.
### APPENDIX B: MATERIALS FOR EXPERIMENTS IN CHAPTER 4

<table>
<thead>
<tr>
<th>Discourses presented aurally</th>
<th>Probes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1.</strong> In the television series, a family is brought to different planets in search of places to live. In one episode, an astronaut rescued a father from quicksand and a Martian rescued him from volcanic lava. The stars of the show always survived.</td>
<td>Space Union      Mother Rabbit</td>
</tr>
<tr>
<td><strong>2.</strong> Nominations for the institution’s awards always initiated discussions about favoritism. This year, a chancellor nominated a coach for the prestigious award and an administrator nominated him for a position on the board. The next day, rumors were already flying.</td>
<td>School Sprint     Team Text</td>
</tr>
<tr>
<td><strong>3.</strong> No cameras were allowed at the performance of Peter Pan. But secretly, a gymnast video taped a mermaid in the act before intermission and a nanny video taped her in the routine just before the finale. Luckily no one got caught.</td>
<td>Athlete Puppets   Ocean Wagon</td>
</tr>
<tr>
<td><strong>4.</strong> The guests at the royal ball often exchanged lavish gifts. This year, a girl surprised the queen with a diamond pendant and a baron surprised her with a pearl necklace. Everyone took note so they could outdo each other next year.</td>
<td>Boy Mob           King Poet</td>
</tr>
<tr>
<td><strong>5.</strong> The most violent day at the saloon started when during a civil game of poker, someone's mother was insulted. A cowboy pushed a robber into the bar and a waiter pushed him into the poker table. Luckily no one was killed.</td>
<td>Horse March       Bank Rice</td>
</tr>
<tr>
<td><strong>6.</strong> At the reunion, everyone poured over a booklet describing people's accomplishments. A pilot intimidated a dentist because of a degree from Yale and a lawyer intimidated him because of a degree from Harvard. But by the end of the night, no one cared.</td>
<td>Plane Field       Teeth Socks</td>
</tr>
</tbody>
</table>
7. In medieval times, it was illegal to accommodate anyone who was banished. But one day, a monk hid a traveler in the stables, and a servant hid him in the kitchen. Had word gotten out, they all would have been sentenced to death.

8. In the fairy tale, evildoers were sent out to spy on the castle. A witch watched a lady inside the tallest tower and a fortuneteller watched her within the crystal ball. The good characters seemed doomed, but the story had a happy ending.

9. In the Las Vegas hotel, rumors spread that something had happened in the theater. A gambler talked to a reporter about a fire and a vacationer talked to him about a loose tiger. They were all disappointed to learn that it was nothing.

10. In the slapstick movie, the main characters were on a cruise trying to catch a pesky fly. In the dining area, a sailor accidentally swatted a baker on the nose and a bartender swatted him on the back of the head. Even after this, the fly escaped.

11. The community center offered a support group to help reduce stress. The teacher listened to the widow about struggles with the children and the model listened to her about self-esteem problems. Almost all of the members signed up for the following session.

12. The country club was an excellent place to network. During the tournament, someone’s brother invited a caddie to a dinner party and a tennis player invited him to lunch downtown. By the end of the game, almost everyone had made new contacts.

13. No one was hurt in the accident near the YMCA, but several people got into an argument. A biker interrupted a lifeguard every couple of sentences and a yoga instructor interrupted him every other word. Finally, the police had to keep the peace.

14. In the Jeopardy game, the contestants even started making rude comments to one another. The critic
insulted the librarian after the Daily Double and the engineer insulted her just before the commercial break. They were all warned that poor sportsmanship was not encouraged.

15. At the beach, someone's wallet was discovered missing and a fight broke out. A surfer punched a crook in the face and a volleyball player punched him in the arm. So many people got involved that the police had to break it up.

16. During the holidays, a small group was stationed at the military base. The soldier teased the chef about the burnt pancakes and the chaplain teased him about believing in Santa Claus. After the holidays, they all tried to keep in touch.

17. The new chat room became very popular on the Internet. In one room, a psychic advised a caregiver about a new relationship and a housewife advised her about cleaning carpets. All the advice was taken at face value since no one had official credentials.

18. The regulars of the resort often teased each other in the lodge. A cashier pestered a skier for not having any ones and a teenager pestered her for wearing a coat from the 70s. The newcomers never knew what the think.

19. On the commuter rail into campus, several professors observed others for the purpose of research. A sociologist watched the conductor as passengers boarded and a psychologist watched him as tickets were collected. Eventually, there was enough material for a few publications.

20. Each year, the company’s internal groups were asked to give a talk. An executive invited a scientist to present research findings and a consultant asked him to present future plans. After the presentations, the company decided how to allocate funds.

21. At the career fair, professionals shared their experiences with students and with each other. This year, a decorator inspired a minor to take an internship and a
recruiter inspired him to enroll in classes. Because of its success the fair soon doubled in size.

22. Thursdays at the college pub were particularly fun because it was game night. One evening, a graduate befriended a sculptor during a chess game and an illustrator befriended her during a Pictionary game. There was a terrible uproar when game night was canceled.

23. One night, there were a few stragglers in the building after-hours. On the way out, a chiropractor noticed a criminal sneaking into the building and a custodian noticed him hiding in the shadows. When a car alarm went off, they all froze in fear.

24. Excited by their costumes for the Halloween play, some of the third graders started rough-housing back stage. An alien pinched an acrobat just behind the curtain and a ghost pinched her near the backdrop. Soon the whole audience heard the giggling back stage.

25. The opera company was getting ready for opening night, when a small earthquake hit the city. The beautician called the seamstress about any injuries and the singer called her about the performance. Soon, there was a notice that the show would go on.

26. In the middle of the flight to San Diego, there was a request for medical staff. A stewardess heard a nurse talk about labor pains and a native heard her talk about a hospital. Within minutes, there was an emergency landing.

27. When some new people were hired at the theater, there was tension between workers. A performer accused a supervisor just before the show and a lighting crewman accused him during intermission. But after the standing ovation, all the accusations were forgotten.

28. Many high-profile people were scheduled to appear at the town’s centennial celebration. During the speeches, a tourist videotaped a bishop for a family keepsake and a guard videotaped him as a security measure. The program ran smoothly but it was a bit boring.
29. B movies always did very well at the town's theater. In this week's show, a **warrior** attacked a **hiker** in a forest and a monster attacked him near a swamp. It was such a hit that the first week was sold out.

30. Things started changing in the wholesome neighborhood when casinos were legalized. In one day, someone's **spouse** saw a **pickpocket** at a buffet dinner and a neighbor saw him at a seedy club. Many families started to move away.

31. After the tornado hit the small town, everyone banded together. A **plumber** helped an **usher** fix a leak at the theater and a mechanic helped him fix a flat tire. For many, it was like living in a new town.

32. On the day of the big wedding, many people didn't know where to go. The **photographer** asked the **mother** where the church was and an in-law asked her where the reception was. Eventually everyone got to where they needed to be.

33. When war broke out, life drastically changed, unexpected attacks happened regularly. One day, a **hunter** just missed an **enemy** hiding in the fields and a commander just missed him along the riverbank. It was shocking to many how the country had been infiltrated.

34. The city put on a free dinner with live music for Independence Day. Before dinner, a **veteran** nudged a **cook** during the raising of the flag, and a vocalist nudged him during the playing of the national anthem. Many went away feeling inspired.

35. Each year there was an awards ceremony to recognize outstanding citizens. This year, the **mayor** honored a **director** with the gold key and a ranger honored him with a kindness to nature award. Many were upset that the best-kept-lawn trophy wasn’t given out.

36. One area of the park’s walkway was consistently causing accidents. In one hour, a **swimmer** bumped into a **wanderer** on the way to practice and a jogger bumped
into him on the way to the track. Soon a street light was installed for pedestrians.

37. A fight broke out during the homecoming game; it was so violent that many were arrested. An officer threatened a kicker on the field and a teammate threatened him in the parking lot. All in all, 14 people were taken downtown.

38. Competition was fierce at the pie-eating contest this year. A partner harassed a wrestler for complaining of a stomach ache and a farmer harassed him for dropping out after one bite. The contest was a disaster and was canceled the following year.

39. Soon after the cruise had left, a terrible storm hit and caused some damage. A paramedic assisted the captain with the minor injury and an electrician assisted him with the failed lighting. A rescue helicopter was eventually called in for help.

40. During the new branch's grand opening, a huge household came in and started a scene. A niece argued with a teller about the negative balance and a roommate argued with her about a bounced check. By 5 o’clock, several had asked to be transferred.

41. The neighborhood was very well kept, but a few properties in the outskirts were in shambles. A painter confronted the landlord about a leaky faucet and an inspector confronted him about hazardous fumes. Before long, the housing committee had to step in.

42. At the community center, there was a party to celebrate birthdays every month. This month, a prankster tricked a toddler with candles that didn’t blow out and a guardian tricked him with an exploding jar of jellybeans. It was definitely the best party yet.

43. Even though there was a scandal in town, everyone pretended that nothing had happened. One day a professor greeted a parent near the bookstore and a friend greeted him outside the cafe. No one breathed a word of the event until behind closed doors.
44. When the location of the celebrities’ wedding was changed on a whim, everything had to be coordinated quickly. The **publisher** contacted the **caterer** before the ceremony and the florist contacted her before the reception. Even with all the frustration, everyone remained calm and professional.

45. This year’s Thanksgiving parade was so crowded there was hardly any room to stand. A **sheriff** bumped into a **pirate** near the Treasure Island float and a juggler bumped into him by the food stand. Everyone came away with a few extra bruises.

46. On the news, there was a great deal of coverage of the summer Olympics. One **student** interviewed a **diver** about necessary training and a sports writer interviewed him about role models. The University gave many sports teams grants to buy recordings of the stories.

47. A terrible hurricane affected everyone in the beach town. A **patient** shocked a **journalist** with a story about an overturned car and a scuba diver shocked him with a story about a flying boat. Although there was a huge amount of damage, everyone survived.

48. Many first-hand experts were filmed in the new program about dangerous sports. A **trainer** interviewed a **skydiver** about necessary gear and an anchorman interviewed him about mental preparation. The program was one of the most popular of the year.
REFERENCES


Poster presented at the Fifteenth Annual CUNY Conference on Human Sentence Processing.


