NORTHWESTERN UNIVERSITY

The Interpretation of Prosodic Prominence Conveying Contrast and Intensity

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Linguistics

By Kate Sandberg

EVANSTON, ILLINOIS

June 2024
Abstract

This dissertation examines the associations between pragmatic meaning categories in English and specific realizations of prosodic prominence. It has been well-established that in Mainstream American English (MAE), prominence is often used to convey contrast. A more limited set of studies suggests that prosodic prominence may also be capable of conveying semantic intensity. Given the fact that prominence can be achieved via changes in multiple acoustic dimensions, the experiments discussed in this dissertation examine the extent to which listeners associate different realizations of prominence with either meaning category.

Previous studies have suggested that contrast and intensity are characterized by prominences featuring differences in pitch (high vs. low) and duration (long vs. extra-long). Therefore, this dissertation begins by asking listeners to make an inference about a speaker’s intended meaning after hearing a single instance of prominence produced on a gradable adjective. The results suggest that increased duration is highly associated with an intensified meaning, while differences in pitch height are less clearly associated with one vs. the other. The dissertation continues by investigating which combinations of pitch and duration serve as the strongest cues to contrast and intensity by asking listeners to select which of two prominences best conveys a single meaning category. The results from these experiments indicate that lower pitch and increased duration are preferred for both contrast and intensity, but that asymmetric lengthening of the onset is a particularly strong cue to semantic intensity only. Furthermore, in the final portion of this dissertation, I determine that asymmetric lengthening of the onset is highly preferred as a cue to intensity targeting the attitudinal dimension, as opposed to just the semantic one.
Overall, this dissertation provides additional insights listeners’ associations of different realizations of prominence with the relatively understudied meaning category of semantic intensity. Furthermore, this project contributes to the wider debate about the mapping between prominence and pragmatic meaning, suggesting that listeners do not simply pay attention to whether a word is categorically more prominent, but attend to differences in how that prominence is being achieved.
Acknowledgements

I could not have completed this thesis without the help of so many important individuals, each of whom deserves the greatest thanks. First and foremost, I must thank the members of my committee, including my advisor, Jennifer Cole, Gregory Ward, Eszter Ronai, and Ann Bradlow. I am eternally grateful for the feedback Ann provided in the early stages of my dissertation as a member of my prospectus committee. Her initial guidance surely helped avoid likely pitfalls later on in the research process. Gregory has supported my research from the early days of my academic career, serving as a second reader for my QP and allowing me to TA and guest lecture for his courses. As a sound researcher, I am so thankful for his insightful feedback and help ensuring that my understanding of the pragmatic literature was sound. I must also thank Eszter for filling in the semantic gaps in my project, and for being especially open and willing to meet with me in the final stages of my dissertation, even when she was on leave. And of course, many thanks go to Jennifer, who welcomed me into her lab, provided constant support, and challenged me to always make my project better. Without her, I would not have learned so much, nor believed that I could truly accomplish anything I set my mind to.

To the members of the Prosody and Speech Dynamics Lab, both current and former, including Dan Turner, Rebekah Stanhope, Thomas Sostarics, Shawn Foster, Züheyra Tokaç, Anna Robinson, and Jeremy Steffman: Thank you for providing me with a safe space to share my ideas, especially when they weren’t yet fully fleshed out. I am so grateful to have had the opportunity to work with such an incredible group of researchers. I must also extend my deepest gratitude to Chun Liang Chan, who helped me create my experiments, and who also dedicated countless hours to teaching me how to learn the basics of coding in JavaScript. I truly could not have done any of this without him. I must also thank all the other members of the Northwestern
Linguistics Department in general. Special thanks must be given to the members of my cohort: Chantal DeLeon, Thomas Sostarics, and Anna Robinson. Chantal, I am so happy to have met you, and I have so much respect and admiration for your positivity and the way that you support and care about all of those around you. To Anna and Thomas both: thank you for getting me through Covid. Your friendship and company were so important during a very dark time, and our game nights and social events were one of the few bright spots in that time of my life. To Thomas in particular: thank you for being willing to help with any and all of the technical parts of my project. You never made me feel embarrassed to ask questions, and you were always willing and ready to provide patient assistance. I am truly grateful to have had your company for all five years of graduate school. Special thanks must also be given to the other friends in the department that I have made throughout the years, including Maria Gavino, Wes Orth, and Cassie Davenport. I am so happy to have found friends in you all, and I will miss the E-crew forever!

On a personal note, I must thank my fiancé and partner in all things, Mark Agrios, for his unwavering support. From undergrad to graduate school, you are the one who kept my spirits up throughout the past five years, even when everything in me was telling me to give up. I truly could not have finished without your constant positivity, and I count myself lucky every day to have you by my side. Thanks should also be given to my mom and stepdad, Alison and Brian Shea, and to my dad, Scott Sandberg, for their constant encouragement. They were always there to answer a frustrated phone call, and they continue to support me in all things. Without them, I would not be the woman I am today.
Dedication

To my dad, who originally inspired the doctoral dream.
Contents

Abstract ...................................................................................................................................................... 2

Acknowledgements ...................................................................................................................................... 4

Dedication .................................................................................................................................................. 6

Contents .................................................................................................................................................... 7

List of Figures .......................................................................................................................................... 10

List of Tables ............................................................................................................................................ 12

General Introduction .............................................................................................................................. 13

Chapter 1: The Effect of Duration on Interpretations of Contrast and Semantic Intensity..... 21

Introduction ............................................................................................................................................. 21

Prosodic prominence ............................................................................................................................... 22

Prosodic Prominence and Information Structure ................................................................................. 24

Prosodic Prominence and Semantic Intensity ...................................................................................... 31

Goals of the Current Study ...................................................................................................................... 33

Experiment 1 ........................................................................................................................................... 34

Experiment 2 ........................................................................................................................................... 52

Overall Discussion ................................................................................................................................... 66

Conclusion ................................................................................................................................................. 73
Chapter 2: The Relative Contribution of Onset and Nucleus Duration to Interpretations of
Contrast and Semantic Intensity ............................................................................ 75

Introduction .................................................................................................................. 75

The Interpretation of Prosodic Prominence ................................................................. 76

The Acoustic Realization of Prominence ..................................................................... 78

The Current Study ......................................................................................................... 82

Experiment 1a – 1b ....................................................................................................... 83

Experiment 2 ................................................................................................................. 108

Overall Discussion ....................................................................................................... 115

Conclusion ..................................................................................................................... 119

Chapter 3: The Relative Contribution of Onset and Nucleus Duration to Interpretations of
Expressive Intensity ..................................................................................................... 121

Introduction ................................................................................................................. 121

Semantic/Expressive Intensity ..................................................................................... 122

Prosody and Intensity .................................................................................................. 125

Experiment 1 ................................................................................................................ 129

Experiment 2 ................................................................................................................. 146

Overall Discussion ....................................................................................................... 152

Conclusion ..................................................................................................................... 161
General Conclusion ........................................................................................................... 162

References ....................................................................................................................... 178

Appendix A ....................................................................................................................... 197
**List of Figures**

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Contrastive, Neutral and Intensifying Displays</td>
<td>36</td>
</tr>
<tr>
<td>2</td>
<td>Duration and Max Pitch of Contrastive, Intensifying and Neutral Utterances</td>
<td>37</td>
</tr>
<tr>
<td>3</td>
<td>Pitch Trajectories for Contrastive, Intensifying and Neutral Utterances</td>
<td>38</td>
</tr>
<tr>
<td>4</td>
<td>Original Contrastive, Intensifying and Neutral Productions and Their Final</td>
<td>42</td>
</tr>
<tr>
<td>5</td>
<td>Predicted Probability of a Contrastive and Intensifying Response for Contrastive, Intensifying and Neutral Prosodic Profiles</td>
<td>45</td>
</tr>
<tr>
<td>6</td>
<td>Predicted Probability of an Intensifying Response Across Duration and Prosodic Profile</td>
<td>49</td>
</tr>
<tr>
<td>7</td>
<td>Voice Quality an Preamble Speech Rate for Contrastive, Intensifying and Neutral Prosodic Profiles</td>
<td>54</td>
</tr>
<tr>
<td>8</td>
<td>Pitch Manipulation for the Adjective ‘Rainy’ in Experiment 2</td>
<td>57</td>
</tr>
<tr>
<td>9</td>
<td>Duration and Pitch Continuum Measures</td>
<td>58</td>
</tr>
<tr>
<td>10</td>
<td>Predicted Probability of a Match in Experiment 2</td>
<td>61</td>
</tr>
<tr>
<td>11</td>
<td>Acoustic Measures of the Unmanipulated Experimental Stimuli</td>
<td>84</td>
</tr>
<tr>
<td>12</td>
<td>Average F0 Across Contrastive and Intensifying Prosodic Profiles</td>
<td>86</td>
</tr>
<tr>
<td>13</td>
<td>Comparisons Tested in Experiment 1a-1b</td>
<td>89</td>
</tr>
<tr>
<td>14</td>
<td>Predicted Probabilities of Selecting an LL and SL Stimulus</td>
<td>94</td>
</tr>
<tr>
<td>15</td>
<td>Log Ability Scores for Interpretations of Semantic Intensity</td>
<td>99</td>
</tr>
<tr>
<td>16</td>
<td>Log Ability Scores for Interpretations of Contrast</td>
<td>102</td>
</tr>
<tr>
<td>17</td>
<td>Log Ability Scores for Judgments of Prosodic Prominence</td>
<td>111</td>
</tr>
</tbody>
</table>
Figure 18. Average F0 Across the Adjectives Produced with Intensifying and Contrastive Prosodic Profiles

Figure 19. Comparisons Tested in Experiment 1

Figure 20. Predicted Probability of Selection for Within-Pitch Trials

Figure 21. Log Ability Scores for an Interpretation of Expressive Intensity vs. Semantic Intensity

Figure 22. Mean Deviation from a Neutral Rating

Figure 23. Average Rating Per Adjective Across Duration and Prosodic Profile

Figure 24. Log Ability Scores for Interpretations of Semantic Intensity vs. Contrast

Figure 25. Predicted Probability of a Match for Neutral Prosodic Profiles
List of Tables

Table 1. Duration Manipulation Implemented in Experiment 1 ................................................................. 40

Table 2. Prosodic Profile x Duration Conditions in Experiment 1a-1b ................................................. 87

Table 3. Stimuli Utilized in Experiment 1 .................................................................................................. 132
General Introduction

A wealth of research focuses on the ability of prosodic prominence to convey aspects of information structure, such as **FOCUS** and **GIVENNESS**. For example, in the sentence, *I bought a new bike today*, prominence on *bike* can reflect the fact that the entity (bike) or the proposition (that a bike was bought) is being introduced into the discourse model for the first time, and thus constitutes new information rather than given (e.g., Halliday, 1967; Baumann & Riester, 2012). Prominence on *bike* might also convey focus by indicating that alternatives to the word *bike* are relevant to the interpretation of the expression, leading to inferences of exhaustivity (e.g., *I bought a new bike, and nothing else* (see Rooth, 1992, 1994 for a formal description of focus).

Prominence can be considered a gradient phenomenon, influenced by categorical changes in accentedness/accident type and by continuous changes in the acoustic dimensions of F0, duration, and intensity (e.g., Kochanski, Grabe, Coleman & Rosner, 2005; Cole, Mo, & Hasegawa-Johnson, 2010; Cole, Hualde, Smith, Eager, Mahrt & de Souzac, 2019; Im, Cole & Baumann, 2023; Bishop, Kuo & Kim, 2020; Fry, 1955; Sluijter & van Heuven, 1996b). For example, research suggests that dynamic pitch movements, often referred to as pitch accents, differ in their prominence, with prominence ratings increasing from L* to H* to L+H* (e.g., Bishop et al., 2020). Furthermore, increases in acoustic dimensions like duration or F0 can increase the likelihood that any categorical pitch accent is heard as more prominent (e.g., Ladd & Morton, 2007; Cole et al., 2010; Im et al., 2023). It is generally accepted that higher degrees of prominence map onto stronger types of meaning. For example, Baumann and Riester (2012) propose that higher positions on a **GIVENNESS HIERARCHY** are associated with higher degrees of prominence (i.e., more prominent pitch accents, such as H*). According to their model (which
uses the same term but different categories as the givenness hierarchy proposed by Gundel, Hedberg, & Zacharski, 1993), words differ in their informativity/salience in the discourse, ranging from GIVEN (explicitly mentioned or particularly salient) at the lowest end of the hierarchy to BRAND NEW (assumed to be unknown to the hearer) at the highest end. Im et al. (2023) provide support for this hypothesis by illustrating that, in production, higher pitch and a larger pitch range were more often associated with information that was new as opposed to given. Furthermore, in perception, participants did appear to rate new information as sounding more prominent.

A similar relationship as the one between prominence and givenness has been proposed for focus. For example, Cruschina (2021) argues that focus types vary in their degree of contrast based on the way in which the set of alternatives is pragmatically exploited (see also Molnár [2002] for the concept of a contrast hierarchy). Types of focus that increase the likelihood of, for example, a pragmatic inference of exhaustivity are more contrastive in the sense that they encourage a listener to make an inference about the exclusion of relevant alternatives. For example, at the lowest end of Cruschina’s hierarchy lies INFORMATION FOCUS, as is illustrated in (1).

(1) What would you like to eat?

I would like a [sandwich].

In this example, focusing the word sandwich gives rise to a contextually open set, and it does not necessarily encourage the listener to assume that the speaker would not also like, for example, chips. In this sense, the focused word (sandwich) is only contrastive in the sense that it is being singled out from a set of possible items to eat. Focus types that involve a higher degree
of contrast, on the other hand, such as **EXHAUSTIVE FOCUS**, encourage the listener to make an inference that the speaker **only** wants a sandwich, and not any other relevant food item. In English, this type of focus is typically associated with more prominent pitch accents such as L+H* (Watson, Tanenhaus & Gunlogson, 2008; Tomlinson, Gotzner & Bott, 2017).

Previous investigations suggest that higher positions on the contrast hierarchy are associated with even higher degrees of prominence than higher positions on the givenness hierarchy. In other words, in terms of their relation to prominence, the scales of givenness and contrast appear to be positioned on top of one another, although they overlap to some degree. For example, Katz and Selkirk (2011) argue that exhaustive focus as cued by a focus sensitive operator like **only** is realized with longer duration, greater intensity, and a more salient F0 movement compared to new information (which falls at the higher end of the givenness hierarchy). Furthermore, Krahmer and Swerts (2001) show that in perception, Dutch listeners perceive words with pitch accents marking contrastive focus as more prominent than words with a pitch accent marking its status as discourse-new. Thus, overall, higher positions on both the givenness and contrast hierarchies are realized with higher prominence, but higher degrees of contrast are realized with even higher degrees of prominence than the highest position on the givenness hierarchy (new).

In this dissertation, I discuss the fact that prominence has also been linked to another meaning dimension: that of **INTENSITY**. For example, producing the word **tall** with prosodic prominence in *the tree is tall* can convey that the tree being discussed is particularly tall. This
type of intensity\(^1\) might be referred to as SEMANTIC INTENSITY, as it appears to target the threshold on a measurement scale of degree values that is introduced in the denotation of an adjective itself (see Kennedy and McNally [2005] for a description of semantic intensity from the perspective of lexical intensifiers). Semantic intensity as conveyed by lexical intensifiers has also been modeled as an expressive phenomenon (Beltrama & Bochnak, 2015). For example, Beltrama and Bochnak argue that the Italian suffix -issimo conveys that “the speaker is in a state of maximal, emotional involvement about the fact that an individual x has property P…While a particular individual may not be the tallest in the context, taller individuals would not trigger any higher emotional involvement” (2015: 870). Since prosodic prominence has also been shown to convey expressive meaning related to emotions and attitudes (Ladd, 2008), I argue that prosody may also be capable of conveying two types of intensity: both semantic and expressive.

The mapping between prosodic prominence and the meaning category of intensity has not been well explored, and many existing investigations tend to conflate the notions of semantic and expressive intensity. For example, limited observations by Bolinger (2013) in English suggest that the intensification of degree words (i.e., semantic intensity) can be conveyed by prosodic prominence, and especially by exaggerated lengthening. Similarly, Jakobson discusses the notion of “emphatic prolongation,” although he argues that the effects of lengthening in this case are restricted to the expressive domain rather than the semantic one (1960: 354). Kohler (2006) and Niebuhr (2010) argue that, in German, prosodic prominence can strengthen the meaning of the word itself, suggesting that they are targeting semantic intensity. However, they also suggest that

\(^1\) The term ‘intensity’ can also be applied to acoustic phenomena. However, in the remainder of this dissertation, I use the term to refer to any type of meaning that involves some kind of underlying scale.
the realization of prosodic intensity differs depending on the valence of the attitude being simultaneously conveyed by the speaker (i.e., expressive intensity). Therefore, it is unclear from their analysis whether the realization of prominence they target is characteristic of semantic intensity vs. expressive intensity.

Furthermore, while prominence conveying intensity is argued to be distinct from prominence conveying high degrees of contrast (e.g., Niebuhr, 2010), prior accounts do not directly compare the two realizations of prominence, making it unclear whether/how the two meanings can be distinguished from one another in an unspecified context. For example, if prominence conveys both high degrees of contrast and intensity, it is unclear how a listener would determine whether prominence on a word like *tall* is meant to convey that the tree is very tall (i.e., semantic intensity) or that the tree is tall but not also wide (i.e., exhaustive focus). One possibility is that the two meaning types differ in their prominence. For example, previous studies suggest that prominences conveying strong degrees of contrast are characterized by especially salient pitch movements (e.g., dynamic and heightened pitch courses) and minor lengthening (Pierrehumbert & Hirschberg, 1990; Watson, Tanenhaus & Gunlogson, 2008; Braun, Asano & Dehe, 2019; Ito & Speer, 2008), while prominences conveying intensity are characterized by exaggerated increases in duration and sometimes less salient pitch movements, such as lowered and narrowed pitch courses (Kohler, 2009; Niebuhr, 2010, Armstrong & Ward, 1926; Coustenoble & Armstrong, 1934). Based on prior theories of prominence, the lengthening present for both meaning types should increase their degree of accentual prominence. However, the less-salient pitch courses possible for intensity should lead to a lower level of prominence than that which is typical of focus types conveying strong degrees of contrast. These
generalizations would suggest that intensity is characterized by lower prominence than contrast, and that participants may distinguish between the two meanings based on overall prominence.

Alternatively, contrast and intensity might be conveyed by similar degrees of prominence but differ in the degree to which pitch vs. duration serve as the primary driver of their salience. This type of reasoning is consistent with Watson’s Multiple Source view of prominence, which argues that “different acoustic factors will influence prominence only in so far as they mark relevant information for the listener” (2010: 177). For example, Watson argues that the reason that F0 correlates with prominence is because it is used to mark information that is new, while duration primarily relates to speaker-related production processes, as it allows speakers to ensure comprehensibility while conserving articulatory effort. Thus, speakers tend to discount duration as a cue to information-structural meaning (Isaacs & Watson, 2008). However, it is possible that, in the context of intensity, duration plays a more important role in conveying prominence than F0, considering previous observations for its role in conveying intensity. Therefore, contrast and intensity would be characterized by similar degrees of prominence but differ in the acoustic cues by which that prominence was primarily achieved.

The goal of this dissertation is to probe the relationship between different realizations of prominence and semantic/expressive intensity, while also examining how these realizations differ from those that are typically associated with high degrees of contrast. In the chapters that follow, I report the findings from experiments that aim to tease apart the role played by duration in conveying these different meanings from the role played by pitch. In Chapter 1, I investigate whether listeners attend to differences in the acoustic realization of a prominence to arrive at distinct interpretations, and indeed whether those interpretations are consistent with those that
have been proposed in the literature for focus vs. intensity. To answer this question, I expose listeners to adjectives that were naturally produced with the intent to convey contrast vs. semantic intensity and manipulated to have long vs. short duration. I then ask them to make a judgment about which meaning (contrast vs. semantic intensity) they think the speaker was intending to convey when they uttered that sentence. The results of the experiment shed light on how differences in duration are used by listeners to make a judgment about an intended pragmatic meaning, and how this utilization differs depending on pitch.

In Chapter 2, I approach the same question from a different angle by testing which realizations of prominence serve as the strongest cues to an interpretation of contrast vs. intensity. For example, I play participants two stimuli differing in their pitch and duration and ask them to choose which stimulus conveys a higher degree of the quality being denoted by a critical adjective (Experiment 1a) or sounds more like a correction (Experiment 1b). In a follow-up experiment, I explore the degree to which judgments about prominence serve as a mediating factor for these results. Specifically, I ask participants to compare the same sets of stimuli in terms of their prominence, exploring whether stimuli that were rated as more prominent were also the ones that were more likely to convey contrast vs. semantic intensity. The results shed light on the degree to which changes in duration cue either interpretation when accompanied by intensifying (lower) vs. contrastive (higher) pitch, while also examining the relationship between the two meanings in terms of the level of prominence they are conveyed by.

Finally, in Chapter 3, I target questions about the ability of prosody to convey expressive intensity. Specifically, I use the same experimental paradigm as in Chapter 2, but I ask participants to make judgments about which stimulus sounds like it conveys a stronger opinion.
The results from this study shed light on the degree to which listeners associate different realizations of prominence with semantic vs. expressive intensity, while also providing insight into whether the two meanings are conveyed by distinct aspects of the acoustic signal. The chapter also reports the results of a study examining the attitudinal valence of each of the stimuli differing in pitch and duration (i.e., whether they are perceived as conveying a positive vs. negative opinion or attitude). This experiment probes how acoustic cues conveying stronger and more positive/negative attitudes might contribute to the association of certain realizations of prominence with interpretations of semantic vs. expressive intensity.

Overall, the experiments discussed in this dissertation shed light on the complex nature of prominence, suggesting that it can be conveyed by a multiplicity of cues, each of which may be conveying a different type of meaning. The key contribution of this project is to provide a clearer picture of the relationship between acoustic realizations of prominence and a meaning category that has received little attention in the literature: that of intensity, both semantic and expressive. The primary findings of the experiments reported on in the chapters that follow are that a) listeners associate increases in duration and intensifying (lower) pitch courses with interpretations of both semantic/expressive intensity and contrast, b) asymmetric lengthening in the onset while holding the duration of the nucleus constant serves as a stronger cue to interpretations of intensity than it does for contrast, and c) listeners associate slightly different realizations of prominence with expressive intensity than they do with semantic intensity, with onset lengthening being a particularly strong cue to attitudinal meaning.
Chapter 1: The Effect of Duration on Interpretations of Contrast and Semantic Intensity

Introduction

Prosodic prominence has been linked to a variety of pragmatic meanings. For example, producing the word *tall* in *the tree is tall* with prosodic prominence can encourage an inference about the exclusion of alternatives that match the word *tall* in semantic type, leading to perceptions of exhaustivity (e.g., *the tree is tall but not wide/green*) (as proposed by Horn, 1972; see Repp & Spalek, 2020 and Gotzner & Romoli, 2022 for recent overviews). A prominent realization of *tall* can also lead listeners to arrive at intensified interpretations (Bolinger, 2013; Armstrong & Ward, 1926; Kohler, 2006). For example, Armstrong and Ward (1926) suggest that increased prominence can add “intensity” to the meaning of a word. Similarly, Kohler argues that prominence can “amplify the meaning of words and…express a particularly great degree of what they imply” (2006: 2). In this sense, prosodic prominence appears to be able to convey intensity as is typically conveyed by lexical intensifiers. For example, producing the word *tall* in *the tree is tall* with prosodic prominence can presumably convey that the threshold of height above which an object qualifies as being tall is higher for this tree than it would be for an un-emphasized version of the word (i.e., that the tree is *very tall* as opposed to just *tall*) (see Kennedy and McNally [2005] for a description pertaining to lexical intensifiers). I refer to this type of intensity as SEMANTIC INTENSITY.

If prominence can convey both contrast and semantic intensity, it remains unclear how listeners distinguish between the two meanings when prominence is produced on an ambiguous word (i.e., a word that introduces the semantic primitives required for both contrast and intensity, such as alternatives that match the prominent word in semantic type), and when that word is
presented in the absence of context. Production studies investigating the realization of prominence conveying these two meanings suggest that listeners may be able to rely on differences in their acoustic manifestation to distinguish between them. For example, while prominence conveying contrast tends to be characterized by salient pitch movements and small or moderate increases in duration (Pierrehumbert & Hirschberg, 1990; further detail in Section 1.1), prominence conveying semantic intensity has been linked to exaggerated increases in duration and both more and less salient pitch movements (German: Kohler, 2006; Niebuhr, 2010; English: Armstrong & Ward, 1926; French: Coustenoble & Armstrong, 1934).

In this chapter, two experiments investigate how changes in duration affect listeners' interpretation of adjectives in utterances like *the tree is tall* as conveying contrast vs. semantic intensity, when those adjectives are characterized by prosodic realizations that were originally intended to convey contrast vs. semantic intensity. The results from this study show that prominence is not a unitary phenomenon. When presented with acoustically different types of prominence, listeners can arrive at different inferences about a speaker’s intended meaning. This section proceeds first with a review of previous work on the realization of prominence more generally, followed by a review of prominence intended to convey contrast vs. semantic intensity before introducing the experimental paradigms used in the current study.

**Prosodic prominence**

Before discussing the meanings conveyed by prosodic prominence and the specific acoustic manifestations that facilitate interpretations of contrast vs. semantic intensity, I must first introduce the notion of prominence as an acoustic and perceptual phenomenon. Prominence is typically defined as the degree to which a linguistic element stands out from the other elements
around it (Terken & Hermes, 2000). In English, it is primarily derived based on a word or syllable’s position in metrically strong positions within a hierarchical prosodic structure. For example, according to Autosegmental Metrical theory, the head of each level of the hierarchy is more prominent than the non-head elements (Terken & Hermes, 2000; Liberman, 1975; Pierrehumbert, 1980; Ladd, 2008). Syllables that are designated as the head of a foot are considered the most structurally prominent within the word (Cole et al., 2010). Similarly, words that are designated as the head of a prosodic phrase become phrasally prominent, standing out more than other words in the utterance. Listeners are biased to hear words or syllables appearing in these positions as more prominent, with the strongest likelihood of perceived prominence for words with lexical stress that occur in the default location for phrasal stress (also termed ‘nuclear’ stress), which in English is the rightmost position within a prosodic phrase (e.g., Cole et al., 2010).

According to Cole et al. (2010), the acoustic prominence of a word refers to the phonetic exponents of structural and phrasal prominence. Acoustic prominence is characterized by changes in one or more acoustic dimensions, including duration, intensity, and F0 (e.g., Kochanski, Grabe, Coleman & Rosner, 2005; Cole et al., 2010; Cole et al., 2019; Im, Cole & Baumann, 2023, inter alia). For example, in sentences without a marked focus, the word appearing in the nuclear position is typically associated with a phonological pitch accent, phonetically realized in a dynamic pitch movement that aligns with the stressed syllable of the word. Pitch accents are phonologically specified with tonal features, (high, low, or bitonal sequences) assigned to the stressed syllable. For example, Pierrehumbert (1980) distinguishes between an H* pitch accent, realized with a relatively high F0 target aligned with the stressed
syllable, and an L* pitch accent, realized with a relatively low F0 target. She also proposes three bitonal pitch accents, such as L+H*, which is realized with an initial low F0 target before rising quickly to a high target that is aligned with the stressed syllable. In their phonetic implementation, pitch accents differ in the scaling of F0 peaks and in the overall size of the F0 excursion, and there are corresponding differences in their perceived prominence. For example, Cole et al. (2010), Cole et al. (2019), and Im et al. (2023), reporting on findings from perceptual prominence ratings using Rapid Prosody Transcription, find that words with the greatest likelihood of being perceived as prominent were those bearing a L+H* pitch accent. Similarly, Bishop et al. (2020), using a similar task, illustrate that the likelihood of perceived prominence varied according to pitch accent in the order of L*< H*<L+H*. For words with the same pitch accent, variation in dimensions like F0 and intensity can also influence the relative perceived prominence of that accent. For example, Cole et al. (2010) and Cole et al. (2019) find that the likelihood that listeners perceive a word as prominent increases as that word’s duration and intensity also increases.

**Prosodic Prominence and Information Structure**

In English, as in other Germanic languages, prosodic prominence plays an important role in conveying pragmatic meaning related to information structure, or the packaging of information within an utterance (Chafe, 1976). For example, prominence can convey the distinction between **NEW vs. GIVEN** information (Halliday, 1967), sometimes also referred to as the **RHEME vs. THEME**, respectively (Steedman, 2000). The new/given distinction is based on the idea that, throughout a conversation, speakers are constantly acknowledging or making updates to information that is in the Common Ground (henceforth CG), or shared background knowledge of the two speakers.
(Stalnaker, 1978; 2002). For example, given information is typically defined as information that has already been established in the CG or is particularly salient in the context, while new information constitutes information the speaker intends to add to the CG.

Givenness is often considered to be a gradient phenomenon, moving from information that is evoked (previously mentioned), to inferable, to new (Prince, 1981; Gundel, et al., 1993; Chafe 1994; Lambrecht 1994; Baumann & Riester, 2012). This gradual nature of givenness is reflected in its acoustic manifestation, as prominence tends to increase as one moves up the givenness hierarchy proposed by Baumann and Riester (2012). For example, previously mentioned (evoked) referents tend to be deaccented, those denoting inferable information are marked with an L* pitch accent, and new referents are marked with an H* (Baumann, 2006; Baumann, Rohr & Grice, 2015; Rohr and Baumann, 2010). However, it is important to note that this relationship is not necessarily one-to-one. For example, Im et al. (2023) find that, in a corpus of speech from two TED talks, although more prominent pitch accents were more often associated with information ranked higher on the givenness hierarchy, this relationship was not statistically significant. In other words, the TED talk speakers showed only a weak preference for using more prominent pitch accents on more informative words, and the reason for this pattern may have nothing to do with a purposeful attempt to distinguish information status categories using pitch.

Prosodic prominence can also convey the distinction between BACKGROUND and FOCUS, sometimes also referred to as KONTRAST (Vallduví & Vilkuna, 1998). According to Rooth (1992), focus introduces a set of alternatives to the focused expression that are relevant for the interpretation of an utterance. For example, focusing the word Amy in (2) introduces alternatives that match the word Amy in semantic type, such as alternative individuals like Jim or Tom.
(2) [Amy] went shopping.

Crucial to this chapter is the notion of CONTRAST, a term which has been used somewhat inconsistently in the literature. According to Calhoun (2006), focus is synonymous with contrast (hence Vallduví & Vilkuna’s adoption of the orthographically distinct ‘kontrast’), as it singles out an element from among a set of alternatives and is therefore always contrastive in nature. However, she argues that different types of focus differ in their contrastiveness, or the speaker’s intent to make the alternatives salient, which influences the pragmatic effects of focus. Cruschina (2021) similarly argues that all focus is contrastive, but that focus types vary in their ‘degree’ of contrast based on the way in which the set of alternatives is pragmatically exploited (see also Molnár [2002] for the concept of a contrast hierarchy). For example, she proposes an increase in the degree of contrast each type of focus conveys that moves from INFORMATION FOCUS (which introduces an alternative set that is contextually open) to EXHAUSTIVE FOCUS (which excludes all relevant alternatives) to MIRATIVE FOCUS (which conveys unexpectedness with respect to the more likely alternatives) to CORRECTIVE FOCUS (which involves the direct negation of given alternatives). This hierarchy is based on the idea that the more active or salient the alternatives are in the context, the greater the degree of contrast. For example, information focus lies at the lowest end of the contrast hierarchy because it conveys contrast simply by being a member of the same set as the alternatives, as shown in (3).

(3). A: What did you buy at the store yesterday?
    B: I bought an apple.

In this example, the only interpretation that the speaker intends the listener to arrive at is that the speaker bought an apple, without saying anything about what other food might have been
bought. Exhaustive focus, on the other hand, which Cruschina (2021) describes as a sub-type of information focus, encourages the listener to make an inference about the exclusion of all alternatives inferred from the context. For example, in the case of (3), exhaustively focusing *apple* would convey that the speaker *only* bought apples, and nothing else. In many languages, this difference in contrastiveness is marked by syntactic strategies such as *FOCUS FRONTING*, when the focus constituent is moved to a left-peripheral position. For example, in Hungarian, focus fronting is only permitted for exhaustive focus and not for information focus (Cruschina, 2021). In English, this difference is more typically marked by prosodic prominence.

The final two types of focus that Cruschina (2021) discusses (and the most contrastive) are *MIRATIVE FOCUS* and *CORRECTIVE FOCUS*. Mirative focus refers to instances when attention is being drawn to the unlikeliness or unexpectedness of the asserted proposition with respect to the alternatives, as shown in (4).

(4) They told us there would only be chicken for dinner but instead we had [steak]_F.

In this instance, Cruschina (2021) describes the salient contrast as a contrast against expectations. Finally, corrective focus differs from these other three types of focus in that the alternatives must have been previously mentioned (or at least be especially salient in the context). Specifically, corrective focus conveys a direct rejection of an alternative already established in the CG (Gusenhoven, 2007). For example, in (5), corrective focus presupposes that an alternative to *Amy* is already salient in the context (as is the case in this example), and it conveys that the speaker is directly rejecting this alternative.

(5) A: Did John go shopping?
B: No, [Amy] went shopping.

In English, all types of focus are marked by pitch accents, regardless of their degree of contrastiveness. A word with contrastive focus is typically the most prominent element within a phrase. When contrastive focus is marked on a word that is non-final in the prosodic phrase, the greater prominence of the focused word is typically achieved by the deaccenting of following syllables (Rooth, 1992; Truckenbrodt, 1995; see Buring, 2016). Some researchers propose that the realization of focus differs from the realization of the new/given distinction, with focus being characterized by pitch accents that are more acoustically prominent relative to new information. For example, in a production experiment, Katz and Selkirk (2011) found that all-new sentences and sentences combining focused and discourse-new constituents differed in their patterns of prominence. Specifically, focused constituents (cued by a focus-sensitive operator like *only*) were more phonetically prominent than discourse-new constituents that lacked focus, with elements under focus showing greater duration and relative intensity in addition to a larger F0 movement. A slightly different pattern is reported by Breen et al. (2010), where it was found that participants did not always distinguish between new information and information that was focused in their productions, but when differences were produced, new information was realized with higher mean and maximum F0 on the new word than focused information, while focused information was realized with greater intensity on the focused word than new information.

Not only does the literature suggest that focus is realized with greater prominence relative to new information, but it also suggests that the degree of prominence increases as the degree of contrast a focus type conveys increases. For example, the H* and L+H* pitch accents proposed by Pierrehumbert and Hirschberg (1990) are argued to differ in their overall prominence by
virtue of their pitch spans, with L+H* having a larger pitch excursion, which makes it a more perceptually prominent pitch accent (Cole et al., 2010; Cole et al., 2019; Bishop et al., 2020). A related finding from a number of studies is that L+H* is more likely to activate alternatives (i.e., render them more salient) compared to H* (Watson, Tanenhaus & Gunlogson, 2008; Braun, Asano & Dehé, 2019, Ito & Speer, 2008). For example, Watson, Tanenhaus and Gunlogson (2008) exposed listeners to alternatives in dialogues like the one shown below:

(6) a. Click on the camel and the dog.

b. Move the dog to the right of the square.

c. Now, move the camel/candle below the triangle.

Using a visual world paradigm, they illustrate that listeners’ eye movements anticipated contrastive targets (candle) upon hearing L+H* in (3c), while they were equally likely to look at either given (camel) or contrastive targets (candle) when hearing H*. Ito and Speer (2008) illustrate that when listeners hear an adjective with a L+H* accent (‘BLUE ball’) they anticipate a referent to the immediately following noun that is an alternative to a discourse salient entity, and they exhibit disrupted processing when the noun is in fact unrelated to that entity. Finally, Braun and Tagliapietra (2010) illustrate that, even in a context that lacks explicit alternatives, contrastive accents (i.e., L+H*) facilitate the recall of contrastive associates. For example, hearing the word lagoon in an isolated sentence produced with the L+H* accent facilitated the recognition of a contrastive associate like river, but had no effect on the recognition of a non-contrastive semantic associate like water. Overall, the results of these studies suggest that interpretations involving alternatives are encouraged by increases in prominence, and that those increases in prominence are particularly associated with changes in pitch span and duration characteristic of an L+H* pitch accent as opposed to H*. 
Since increases in prominence increase the perceived contrastiveness of an accented word, they also increase the likelihood that a pragmatic implicature is derived. For example, Tomlinson, Gotzner and Bott (2017) illustrate that an L+H* accent facilitates exhaustive inferences relative to H*. In their experimental paradigm, which utilized mouse-tracking, participants read a sentence and then used their computer mouse to select which of two pictures (one representing an exhaustive interpretation and one not) best matched the meaning of the sentence. For example, if the sentence was, “Mark has a candle,” the non-exhaustive picture depicted a candle and a camel, while the exhaustive picture depicted only a candle. The researchers also had a control condition in which one picture was consistent with the sentence and one was not; for example, in the control condition, the sentence, “Mark has a candle,” was accompanied by a picture of a candle on the left and a camel on the right. The researchers measured ease of processing based on directness of the participant’s mouse path from initial position to the response. Overall, their results suggest that implicatures were derived earlier in sentence processing under L+H* than H*, as they saw more direct mouse paths in the inferential condition than in the control condition for L+H* than for H*. Based on the results of this study, the authors posit that L+H* may serve to elevate the salience of the relevant alternatives compared to H*, thereby increasing the degree of contrastiveness of a given focus type.

Overall, the literature on the realization of prominence conveying information structure suggests that increases in prominence increase the likelihood that an element is perceived as a) new, b) focused and c) more contrastive. Furthermore, the influence of prominence on the contrastiveness of an element increases the likelihood that a listener will arrive at an exhaustive interpretation when hearing a more prominent pitch movement. In the following section, I review
the literature on the realization of prominence conveying semantic intensity, comparing this realization of prominence to the realization for contrast discussed here.

**Prosodic Prominence and Semantic Intensity**

Prosodic prominence has also been linked to the meaning category of intensity. To date, intensity has primarily been modeled in relation to lexical intensifiers. For example, lexical intensifiers like *really* or *very* are said to target the literal meaning of gradable predicates like *tall*, which introduce measurement scales of ordered degree values in their denotations (Kennedy & McNally, 2005). Specifically, lexical intensifiers lead to an increase in the threshold ($\theta$) for a given adjective, which is typically defined as a value on the scale above which an entity is true of a given adjective. For example, qualifying as *tall* means “having a degree of height greater than $\theta$” (Bennett & Goodman, 2018: 148), and qualifying as *very tall* means having a degree of height that is higher than that of a tree that is only *tall*. This higher threshold is inferred by collecting the set of objects in the comparison class for which the bare adjective is true (for example, all types of trees that are above the threshold for *tall*), and then using that as the comparison class to infer a new threshold (Bennett & Goodman, 2018). In this way, a “*very tall tree* denotes something akin to *tall for a tall tree*” (Bennett & Goodman, 2018: 149). I refer to this type of intensity as **SEMANTIC INTENSITY**, as it targets a measurement scale provided by the denotation of the adjective itself.

Although prosodic prominence has not been specifically modeled in relation to semantic intensity, anecdotal observations suggest that it may perform a similar function as lexical intensifiers (Bolinger, 2013; Jakobson, 1960). For example, Bolinger (2013) discusses the fact that adjectives, nouns, and verbs can be intensified phonologically and may often be intensified
in other ways at the same time. Specifically, he links intensification by prosody to phonetic lengthening, and specifically, lengthening in the nucleus of the stressed syllable. According to prior investigations, prominence intended to convey intensity is not always realized in the same way as prominence intended to convey contrast. For example, early observations on English and French by Armstrong and Ward (1926) and Coustenoble and Armstrong (1934), respectively, suggest that prominence conveying semantic intensity can be expressed in one of two ways. The first way is somewhat consistent with prominence conveying contrast: It is characterized by a rising-falling F0 contour with shallow slopes and a high plateau, in addition to considerable lengthening of the emphasized vowel. However, the second means of expression differs considerably from prominence conveying contrast: It is described as involving a narrowed and lowered pitch course with a “short falling movement” and considerable lengthening of the syllable onset and preceding sound (Niebuhr, 2010: 3). These observations echo those made by other researchers noting the role that “emphatic prolongation” and “intensification by stretching” can play in conveying an intensified meaning in English (Bolinger, 2013: 281; Jakobson, 1960: 354). For example, Niebuhr (2010) finds in a production study conducted in German that the mean durations of either the onset or nucleus of words conveying intensity tended to be around 50% higher than the values typical of pitch-accented syllables in German.

Niebuhr (2010) also takes a deeper dive into the realization of prominence conveying semantic intensity more generally, although he does not directly compare intensity to contrast. For example, he finds that, in German, the two realizations identified by Armstrong and Ward (1926) and Coustenoble and Armstrong (1934) roughly map onto a distinction between semantic intensity conveying negative vs. positive attitudes. Furthermore, these specific manifestations differ in more ways than just those identified in these early studies. In his studies, negative
semantic intensity was indeed characterized by lowered and narrowed pitch courses, as it was most typically realized with a pointed F0 peak that fell sharply right after the vowel’s onset. Furthermore, words that were produced to convey negative semantic intensity featured lengthening of both the preceding sound and onset consonant along with quick and abrupt articulation of the vowel nucleus and pressed and irregular voice quality. Positive semantic intensity, rather than featuring a pointed F0 peak, tended to be characterized by an extended high F0 plateau featuring steep rises and shallow falls. Furthermore, positive semantic intensity tended to exhibit quickly articulated onset consonants and lengthened vowel nuclei in addition to sonorous, slightly breathy voice.

In perception, the characteristics of prominence conveying semantic intensity do appear to be perceived by listeners as sounding more prominent and more intensifying, at least for German speakers (Landgraf, 2014). Furthermore, participants are able to correctly identify these acoustic profiles as conveying positive vs. negative semantic intensity (Landgraf, 2014). However, to the best of my knowledge, no study to date has investigated listeners’ association of these realizations of prominence with semantic intensity in comparison to contrast. Thus, it remains unclear whether listeners could use these purported differences in acoustic implementation to determine a speaker’s intended meaning.

**Goals of the Current Study**

Taken together, these prior studies suggest that, although contrast and semantic intensity can both be conveyed by prosodic prominence, the two meanings might be characterized by unique realizations of prominence featuring differences in a variety of acoustic dimensions, including pitch, duration, and perhaps even voice quality. To the best of my knowledge, there
have thus far been no studies that directly test whether listeners can identify distinct realizations of prominence as conveying different types of meaning outside of information structure, nor are there any that compare contrast and semantic intensity in the same experiment. Therefore, in this study, I investigate whether listeners are sensitive to differences in the acoustic manifestation of prosodic prominence as cues to contrast vs. semantic intensity, asking whether their intuitions about a likely interpretation are consistent with the prosodic characteristics suggested in the studies above. Given prior associations of semantic intensity with large increases in duration, I specifically investigate how changes in duration affect the interpretation of words originally produced with pitch characteristics intended to convey contrast vs. semantic intensity, hypothesizing that sufficient lengthening on an adjective can encourage interpretations of semantic intensity over interpretations of contrast.

**Experiment 1**

*Materials*

The visual stimuli for this experiment were four image sets, each depicting five qualities described by gradable adjectives (for a total of 20 adjectives). Each image set was assigned 11 visual displays: one contrastive display, five intensifying displays, and five neutral displays (one for each of the five adjectives). The contrastive display featured five items that differed in qualities that could be described by alternative gradable adjectives. For example, one contrastive display consisted of five shirts differing in their pattern: one being *striped*, one being *spotted*, one being *bright*, one being *dark*, and one being *dirty*. Crucially, these qualities were depicted as being mutually exclusive. In other words, no shirt that was striped was also spotted. In this display, the relevant lexical item was written below each image (*striped, spotted, etc.*). The
intensifying displays featured the same items differing in the degree to which they instantiated the quality denoted by each of the adjectives also included in the contrastive display. For example, shirts varying in their number of stripes, shirts varying in their brightness, etc. Crucially, the image that instantiated the quality denoted by the adjective to the greatest degree was the same image featured in the contrastive visual display. Finally, neutral displays featured the critical item pictured among a collection of unrelated items. In these displays, the noun describing the object was written below each image (shirt, dog, candle, etc.). In this sense, the unrelated items were meant to focus as distractors rather than direct alternatives, which was the case in the intensifying and contrastive images.

For each display, an arrow pointed at the item being described by the audio stimulus. Examples of the visual stimuli are shown in Figure 1. The 20 critical adjectives were selected based on their abilities to be incorporated into groups of mutually exclusive gradable adjectives, to facilitate the creation of the visual stimuli. For example, the four image sets addressed emotions (happy, sad, mad, scared, tired), shirt patterns (striped, spotted, bright, dark, dirty), taste (spicy, sweet, sour, salty, bitter), and weather (sunny, rainy, windy, cloudy, snowy).
Figure 1.

*Contrastive, Neutral, and Intensifying Displays*

![Figure 1](image)

*Note.* This figure demonstrates (from left to right) examples of the visual stimuli for contrastive, neutral, and intensifying contexts.

The auditory stimuli for this experiment were sentences of the frame “The X is ADJ” naturally produced by the author of this study with prosody intended to convey semantic intensity, contrast, and a neutral, declarative meaning. As shown in Figure 2, these three prosodic realizations were acoustically distinct. The stimuli that were produced with the intent to convey semantic intensity were characterized by longer duration in both the nucleus and onset of the critical adjective compared to the neutral or contrastive stimuli and by lower pitch than the contrastive stimuli. Furthermore, consistent with prior observations for contrast, the stimuli that were produced with the intent to convey contrast featured higher pitch maxima on the critical adjective than the neutral stimuli.
Figure 2.
*Duration and Max Pitch of Contrastive, Intensifying and Neutral Utterances*

Note. Plot A illustrates the average duration (ms) of the nuclei and onsets of all 20 adjectives when produced with the intent to convey contrast, semantic intensity, and a neutral meaning, respectively. Plot B illustrates the maximum pitch value (Hz) realized across the entire adjective for each of the 20 adjectives produced with the intent to convey the same three meanings.

As shown in Figure 3, these naturally produced utterances also differed in the overall shape of their F0 trajectories. For example, the sentences produced with the intent to convey contrast featured a pitch trajectory that was higher across the entire utterance compared to sentences produced with the intent to convey semantic intensity. The contrastive stimuli also featured a lower and shallower F0 peak in the prenuclear material compared to the intensifying stimuli, which featured an F0 peak in the prenuclear material that was scaled slightly higher than the peak realized on the critical adjective.
Note. The raw pitch trajectories across the entire sentence for all 20 sentences produced with the intent to convey contrast, semantic intensity, or a neutral meaning. The portion of the contour outlined in red represents the F0 trajectory over the nuclear word (in this case, the critical adjective), while the portion of the contour outlined in blue represents the F0 trajectory over the material spanning from the beginning of the sentence up to the critical adjective (e.g., ‘the shirt is’). The bolded line illustrates an average F0 trajectory for each condition, calculated at 20 equally spaced points across the entire utterance.

To control for the differences in the F0 trajectories of the prenuclear material across the three conditions as shown in Figure 3, the adjectives were excised from the naturally produced utterances before manipulation. Although I controlled for the differences in F0 in the prenuclear material, I chose not to manipulate pitch on the critical adjective itself for three reasons: (1) because prior work identifies multiple pitch patterns that are characteristic of semantic intensity, and I lacked clear intuitions about which pattern would be more strongly associated with this
meaning type, (2) because it was not clear which dimension of pitch would have the greatest effect on interpretations of semantic intensity (i.e., excursion size, max F0, shape of pitch trajectory, etc.), and (3) because adjectives produced with the intent to convey contrast vs. semantic intensity were already distinct in terms of their pitch scaling in the original, natural productions. However, I did choose to directly manipulate the stressed vowel (nucleus) duration of the naturally produced contrastive and intensifying stimuli in Praat, in order to have controlled differences between items with longer vs. shorter duration. My decision to use resynthesis only for duration control does not allow me to draw conclusions about specific aspects of pitch on the critical adjective (e.g., F0 shape, excursion size) and their variable impact on the interpretation of semantic intensity vs. contrast, but it does allow me to investigate how duration affects the ability of a prosodic profile (the holistic pitch characteristics of an adjective produced with a specific meaning in mind) to convey its originally intended meaning (semantic intensity vs. contrast).
Table 1

*Duration Manipulation Implemented in Experiment 1*

<table>
<thead>
<tr>
<th>Word</th>
<th>Prosodic Profile</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Intensifying</strong></td>
<td></td>
</tr>
<tr>
<td>Bitter</td>
<td></td>
<td><strong>Long</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Onset = (\text{Dur(NeutralOnset)})</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Short</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Onset = (\text{Dur(NeutralOnset)})</td>
</tr>
<tr>
<td></td>
<td><strong>Contrastive</strong></td>
<td><strong>Long</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Onset = (\text{Dur(NeutralOnset)})</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Short</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Onset = (\text{Dur(NeutralOnset)})</td>
</tr>
</tbody>
</table>

*Note.* Visual depiction of the duration manipulation for the adjective ‘bitter’ with contrastive vs. intensifying prosodic profiles and long vs. short duration.

As shown in Table 1, the duration manipulation involved creating long and short duration conditions for the nucleus of each adjective produced with contrastive and intensifying prosodic profiles. The landmark for the short duration condition was based on the duration of the nucleus of that adjective when produced with a neutral prosodic profile. The landmark for the long condition was based on the duration of the nucleus of that adjective when produced with a contrastive prosodic profile multiplied by two. I chose this landmark because Niebuhr (2010) finds in his production experiment in German that the mean durations of consonants and vowels in intensified syllables were around 50% higher than the values previously reported to be typical of accented syllables in German. Therefore, doubling the duration of the nucleus in the
contrastive stimuli (a pitch-accented syllable) ensured that the duration was long enough to potentially facilitate an interpretation of semantic intensity.

The manipulation described above resulted in five unique intonational conditions for each adjective: neutral, contrastive-long, contrastive-short, intensified-long, and intensified-short. For all adjectives, the duration of the onset was also normalized by setting it equal to the duration of the onset of that adjective produced with a neutral prosodic profile. After the durations were manipulated, the intensity of all adjectives was normalized to match that of the neutral sentences. Finally, the adjectives were spliced back into their respective neutral sentences. An example of the original productions of the adjective mad intended to convey contrast, semantic intensity, and neutral, along with the word’s final realization in all five intonational conditions after manipulation, is shown in Figure 4.
Figure 4

Original Contrastive, Intensifying and Neutral Productions and Their Final Manipulations

Note. The F0 trajectory across the entire utterance for a) the original productions of the adjective *mad* when intended to convey contrast, semantic intensity, and a neutral meaning, and b) the final, long and short versions of the adjective *mad* when spliced back into a neutral sentence.
Participants

Forty-one students enrolled in introductory linguistics courses at Northwestern University received course credit for participating in the experiment. The data were filtered to include only responses from participants who self-reported as having normal hearing, corrected-to-normal vision, and no speech or reading impairments. Furthermore, all participants were native speakers of English who were born in the United States. Seven participants were excluded because they did not fit the criteria described above, resulting in a total of 34 participants.

Design and Procedure

Participants completed the experiment remotely via their own computers and were instructed to wear headphones. On each trial, they listened to an auditory stimulus while being presented with two side-by-side visual displays for that adjective. After listening to the stimulus, they were told to click on the display they thought the speaker had been looking at when they uttered the preceding sentence.

There were three types of trials in the experiment: trials pairing contrastive displays with intensifying displays, trials pairing neutral displays with contrastive displays, and trials comparing neutral displays with intensifying displays. Each participant was exposed to 20 adjectives but was only exposed to trials pairing contrastive with intensifying displays for 10 of those adjectives (striped, spotted, bright, dark, dirty, happy, sad, mad, tired, scared). The remaining 10 adjectives appeared with a balanced number of the two trial types that included neutral visual displays.

Each participant completed 100 trials (20 adjectives crossed with five intonational conditions). The trials were organized into five blocks, so that in each block, participants heard
all 20 adjectives realized with one of the five intonational conditions. The blocks then rotated through the stimuli so that each adjective occurred once in each intonational condition. The stimuli were randomized within blocks, and the block order was shuffled between participants. Finally, the presentation of the visual displays on either the left or the right was approximately balanced.

Hypotheses

Based on prior observations for the prosodic realization of contrast and semantic intensity, I hypothesize that increases in duration will increase the likelihood of contrastive and intensifying interpretations relative to neutral ones, but that when the two interpretations are directly pitted against one another, listeners will be more likely to arrive at an interpretation of semantic intensity when hearing an adjective with increased duration than they will an interpretation of contrast. Furthermore, I predict that this effect of duration will be evident for stimuli featuring both intensifying and contrastive prosodic profiles. However, I also predict that there will be an interaction between prosodic profile and duration, such that stimuli featuring intensifying prosodic profiles and increased duration will be the most likely to convey semantic intensity, given that these profiles were characterized by lower pitch, and lower pitch profiles have been shown to be consistent only with interpretations of semantic intensity and not with contrast. Thus, overall, I predict that the likelihood of an intensifying interpretation will increase from intensified-long > intensified-short, contrastive-long, > contrastive-short and the likelihood of a contrastive interpretation will increase from contrastive-long > contrastive-short, intensified-long > intensified-short.
Analyses and Results

To first establish that both the contrastive and intensifying prosodic profiles were identified by listeners as conveying contrast vs. semantic intensity, respectively, compared to a neutral prosodic profile, I fit two logistic regression models to the data from trials comparing neutral displays to contrastive and intensifying displays. The models predicted response type (contrastive/intensifying, neutral) from the fixed effect of prosodic profile (intensifying, contrastive, neutral). The model also included random intercepts for participant and adjective. Prosodic profile was sum-coded with neutral pitch set to the reference level.

Figure 5
Predicted Probability of a Contrastive and Intensifying Response for Contrastive, Intensifying, and Neutral Prosodic Profiles
Note. The predicted probability of selecting a contrastive (top) or intensifying (bottom) display when hearing adjectives featuring neutral, intensifying, and contrastive prosodic profiles. The predicted probabilities are pictured along with the 95% confidence interval around the predicted probability.

Beginning with the trials comparing neutral and contrastive displays, as shown in Figure 5, the output of the model suggests that participants were significantly more likely to select a contrastive display when the prosodic profile was intensifying ($\beta_1=0.75$, $z=3.218$, $p<.001$) or contrastive ($\beta_1=0.84$, $z=3.329$, $p<.001$) compared to when it was neutral. Furthermore, a significant intercept of the model indicates that contrastive responses were overall more likely, as the log odds of selecting a contrastive display was 90% ($\beta_0 = 2.28$, $z=7.695$, $p<.001$). This
finding suggests that participants were biased to select contrastive displays compared to neutral displays across all three prosodic profiles.

Moving on to the trials comparing neutral and intensifying displays, the output of the second model suggests that participants were significantly more likely to select an intensifying display when the prosodic profile was intensifying ($\beta_1=0.6$, $z=1.885$, $p=0.05$) compared to when it was neutral. However, intensifying responses were not significantly more likely for contrastive prosodic profiles ($\beta_1=0.17$, $z=0.551$, $p=0.5$) compared to neutral. Finally, just as in the contrastive vs. neutral trials, intensifying responses were overwhelmingly more likely compared to neutral responses, as a significant intercept of the model suggests that the log odds of selecting an intensifying display while holding prosodic profile constant was 75% ($\beta_0 = 1.07$, $z=3.126$, $p<.001$).

The results from these two models first suggest that both the intensifying and contrastive prosodic profiles were distinct from the stimuli featuring neutral prosodic profiles. Specifically, the contrastive prosodic profiles had a higher likely of being interpreted as contrastive than the neutral prosodic profiles, while the intensifying prosodic profiles had a higher likelihood of being interpreted as both contrastive and intensified in comparison to the neutral prosodic profiles. The results from the contrastive model also show that contrastive and intensifying prosodic profiles were equally likely to be interpreted as contrastive (holding duration constant at its average). Finally, the models suggest that, overall, there was a bias for choosing a non-neutral display, and the bias for selecting the contrastive display was even stronger than the one for selecting an intensifying one, given that the contrastive displays were selected even more frequently than intensifying images on trials pairing neutral displays with contrastive/intensifying ones.
Having determined that the contrastive and intensifying prosodic profiles were indeed interpreted as sounding contrastive/intensifying compared to a neutral prosodic profile, I can now analyze the effect of duration on the ability of these acoustically distinct prosodic profiles to convey contrast and semantic intensity, respectively. To do so, I fit a mixed effects logistic regression model to the data from the trials comparing contrastive and intensifying displays and featuring only contrastive vs. intensifying prosodic profiles, as these were the only prosodic profiles with manipulated duration. The model predicted response type (intensifying, contrastive) from the fixed effects of prosodic profile (intensifying, contrastive) and duration (long, short), as well as the interaction between prosodic profile and duration. The model also included random intercepts for participant and adjective. All fixed effects were sum-coded. Contrastive prosodic profiles and short duration were mapped to -0.5 while intensifying prosodic profiles and long duration were mapped to 0.5.

As shown in Figure 6, the model results are consistent with my predictions, as they indicate a significant main effect of duration, such that the log odds of an intensifying response increased in the long duration condition relative to the short duration condition ($\beta = 2.38$, $z = 16.493, p < 0.001$). Furthermore, the model results support my hypothesis that intensifying prosodic profiles would increase the likelihood of an intensifying response relative to contrastive prosodic profiles, as they indicated a significant main effect of prosodic profile, such that the log odds of an intensifying response increased for intensifying prosodic profiles relative to contrastive prosodic profiles ($\beta = 0.545$, $z = 4.118, p < 0.001$). Finally, there was a significant interaction between prosodic profile and duration ($\beta = -0.64$, $z = -2.414, p < 0.05$). As shown in Figure 6, intensifying prosodic profiles led to a greater increase in the log odds of an intensifying response in the short duration condition than in the long duration condition. In other words,
duration was the primary cue to intensifying interpretations, since when long duration was present, prosodic profile had a smaller effect on responses. Furthermore, the condition with the greatest likelihood of being heard as intensifying was the stimulus featuring an intensifying prosodic profile and long duration, as I predicted.

**Figure 6**
*Predicted Probability of an Intensifying Response Across Duration and Prosodic Profiles*
Note. Predicted probability of an intensifying response plotted with 95% confidence intervals in both the long and short duration conditions and the contrastive and intensifying prosodic profiles. The dotted line represents a 50% chance of being selected.

Discussion

The results of this experiment illustrate that increased duration increases the likelihood of an intensifying response relative to a contrastive one, suggesting that increased duration is even more closely tied to the meaning category of semantic intensity than it is to contrast. Moreover, this cue seemed to distinguish interpretations of semantic intensity from interpretations of contrast to a stronger degree than differences in the prosodic profile characteristic of either interpretation. For example, while syllable lengthening had a large effect on the likelihood of an intensifying response relative to a contrastive one, regardless of prosodic profile, the effect of prosodic profile on participant responses was confined to the short duration condition, suggesting that listeners attended to changes in the prosodic profile as cues to semantic intensity only when increased duration was not already present.

Since contrastive responses were overall more likely on the trials comparing contrastive and intensifying interpretations, and because intensifying responses required a substantial increase in duration, it is possible that intensifying responses are simply associated with higher degrees of prominence than contrast. However, if this were true, I would also expect that the stimuli with the highest likelihood of conveying semantic intensity would be the stimuli with contrastive (i.e., higher) prosodic profiles and increased duration, as these two cues are reported to result in higher prominence (Gussenhoven, 2002, 2004; Cole et al., 2019; Im et al., 2023). Yet, the results indicate that the stimuli with contrastive vs. intensifying prosodic profiles and long
duration did not differ greatly in their likelihood of being perceived as contrastive vs. intensifying, suggesting that it is not a difference in overall prominence alone that conveys the distinction between semantic intensity and contrast.

Overall, the results of this experiment support the hypothesis that increases in duration increase the likelihood that participants arrive at an interpretation of semantic intensity as opposed to a contrastive one. However, they also suggest that increases in duration are more strongly tied to interpretations of semantic intensity than intensifying prosodic profiles are. For example, stimuli with both intensifying and contrastive prosodic profiles appeared to be compatible with intensifying responses. Nonetheless, longer duration seemed to be able to push both profiles toward an intensifying interpretation rather than a contrastive one. This finding is consistent with the idea that intensifying interpretations are available with a larger variety of prosodic profiles, as suggested by Niebuhr (2010).

However, I believe these results should be taken with caution, specifically in light of the resynthesis method used to lengthen the nucleus duration of the stimuli. First, since the duration manipulation was only applied to the adjective, which was then spliced back into a neutral utterance, in some cases this resulted in abrupt changes in speech rate between the preceding material and the emphasized adjective. A second consideration is that I did not re-align the pitch contour after duration manipulation, and I am not certain how in a natural production, increasing duration would impact the shape of a co-occurring pitch movement. A final consideration is that the durations chosen to represent the long and short duration conditions were based on limited data from the author of this study’s own productions. Therefore, it is possible that the long duration employed in this study was perceived by participants as unusually exaggerated, eliminating anything other than an intensifying interpretation.
In Experiment 2, I attempt to control for these inconsistencies by introducing variation around the duration values of the natural productions rather than manipulating their durations to match arbitrary and potentially exaggerated thresholds. Specifically, I ask how much variation in duration from a longer vs. shorter baseline is required to push a contrastive vs. intensifying prosodic profile from a contrastive interpretation to an intensifying one, and vice versa.

**Experiment 2**

In this experiment, I focus on the effect of increased duration on interpretations of semantic intensity as opposed to contrast when the adjective that features an increase in duration is realized with a prosodic profile originally intended to convey contrast vs. semantic intensity. Specifically, I test whether varying the duration of an adjective naturally produced with long or short duration, respectively, influences the way listeners associate prosodic profiles originally intended to convey contrast vs. semantic intensity with their respective interpretations. I predict that, for adjectives originally produced with a prosodic profile intended to convey semantic intensity (i.e., with lower pitch and increased duration), increasing the duration even further will increase participants’ associations of this prosodic profile with interpretations of semantic intensity, while decreasing it will reduce these associations. Conversely, I predict that increasing the duration of an adjective originally produced with prosody intended to convey contrast (i.e., with higher pitch and shorter duration) will decrease associations of this prosodic profile with interpretations of contrast, while decreasing duration will increase them.

In this experiment, I also investigate how variation in the overall scaling of the F0 trajectory across the entire utterance impacts the association of contrastive and intensifying prosodic profiles and changes in duration with interpretations of semantic intensity vs. contrast. For
example, I investigate whether decreasing the scaling of an F0 trajectory overall increases the likelihood of an intensifying interpretation relative to a contrastive one. This manipulation is largely exploratory, although I do hypothesize that interpretations of semantic intensity will be relatively unaffected by changes in overall F0 scaling, given Niebuhr’s (2010) observations that intensity can be characterized by both higher and lower pitch maxima, while interpretations of contrast may show greater sensitivity to the scaling of the entire utterance.

*Materials*

The auditory stimuli for this experiment were a subset of 10 adjectives used in Experiment 1 (*windy, rainy, snowy, happy, mad, salty, sour, spotted, dirty, bright*). Rather than utilizing splicing to control for the prosodic material included in the preamble of each production (the portion of the sentence preceding the adjective), in this experiment, each adjective was presented in the sentence of its original production, which preserved the pitch pattern over the entire sentence and resulted in three, unmanipulated prosodic profiles for each adjective: contrastive, intensifying, and neutral.
Figure 7

*Voice Quality and Preamble Speech Rate for Contrastive, Intensifying and Neutral Prosodic Profiles*

Note. Average H1-H2 values for the critical adjective (left) and average speech rate of the preamble (right) for all 10 adjectives used in Experiment 2.

Just as in Experiment 1, the naturally produced intensifying, contrastive, and neutral stimuli differed in their original pitch profiles and durations, with the contrastive stimuli featuring a higher mean value of max pitch and the intensifying stimuli featuring pitch movements with lower value pitch maxima. In terms of duration, the intensified stimuli once again featured higher mean onset and nucleus durations than the contrastive and neutral conditions.

Because the entire pitch contours of the stimuli were maintained in this experiment, I also measured a few other differences between the three prosodic profiles. For example, since
Niebuhr (2010) suggests that semantic intensity may be characterized by differences in voice quality, I measured the dB differences between the amplitudes of the first two harmonics (H1-H2) across the critical adjective in each condition using PraatSauce (Kirby, 2018b). Lower H1-H2 values (and especially negative ones) indicate pressed voice while higher values indicate breathy voice (Fischer-Jørgensen, 1967; Klatt and Klatt, 1990; Keating and Eposito, 2007). Modal voice is typically indicated by H1–H2 ±5 dB around 0 dB. As shown in Figure 7, the measurements suggest that the contrastive, intensifying and neutral conditions did differ slightly in voice quality, as the intensifying and neutral stimuli were characterized by lower H1-H2 values than the contrastive stimuli. However, all sets of stimuli were characterized by H1-H2 values that fell within the range of values for modal voice, suggesting that no condition was characterized by particularly irregular voice quality, at least for voice quality differences captured by H1-H2.

We also measured the speech rate of the preamble of each condition by dividing the number of syllables by the word duration in seconds. Higher values in this case indicate a faster speech rate, and lower values a slower speech rate. As shown in Figure 7, the intensifying stimuli were characterized by increased duration across the preamble, and coupled with the longer duration of the adjective, which indicates that the intensified stimuli are produced with a slower speech rate across the entire sentence, compared to the contrastive or neutral stimuli. Overall, the acoustic measures illustrated in Figure 7 indicate that the contrastive, intensified and neutral stimuli differed in a variety of ways across the entire utterance, and not just in the duration and pitch maxima of the critical adjective.

From each of these original productions, the adjectives were manipulated to create two continua for each prosodic profile: a duration continuum and a pitch continuum. The duration
continuum targeted the nucleus of the stressed syllable in each adjective and consisted of two steps down and two steps up in increments of 20% of the nucleus duration in the original production. As shown in Figure 8, the pitch continuum targeted the scaling of the pitch trajectory across the entire sentence and consisted of two steps down and two steps up in increments of 1 semitone from the original production. These manipulations resulted in a total of 5 steps for each continuum. The final duration (ms) and F0 (Hz) values at each step of the continuum for each adjective are shown in Figure 9. After manipulation, the intensity of the stimuli was normalized to be equal across the contrastive, intensified and neutral conditions to control for effects of intensity on prominence.
Figure 8

*Pitch Manipulation for the Adjective ‘Rainy’ in Experiment 2*

*Note.* Each line pictured in this figure represents one step of the pitch continuum for the adjective ‘rainy’ produced with a contrastive (top) and intensifying (bottom) prosodic profile.
**Figure 9**

*Duration and Pitch Continuum Measures*

![Graph showing duration and pitch continuum measures](image)

*Note.* The average pitch across the entire utterance (left) and the duration of the adjective’s nucleus (right) for each of the 10 adjectives at each step of the continuum when produced with contrastive and intensifying prosodic profiles in Experiment 2.

**Design and Procedure**

The procedure was the same as in Experiment 1. Participants completed the experiment remotely via their own computers and were instructed to wear headphones. On each trial, they listened to an auditory stimulus while being presented with two side-by-side visual displays for that adjective. After listening to the stimulus, they were told to click on the display they thought the speaker had been looking at when they uttered the preceding sentence.

The experiment was between-subjects, with half the participants hearing stimuli with manipulated durations and half hearing stimuli with manipulated pitch. On critical trials, participants were exposed to manipulations of the contrastive and intensifying stimuli and asked...
to make a comparison between a contrastive and intensifying image. On filler trials, participants were exposed to a manipulation of the neutral prosodic profile and asked to make a comparison between neutral and contrastive/intensifying images. The stimuli were organized into five blocks, with each adjective appearing three times in each block (once with each prosodic profile: contrastive, intensifying, and neutral). As a result, there were 10 contrastive, 10 intensifying, and 10 neutral prosodic profiles in each block. Half of the trials featuring neutral prosodic profiles compared neutral and contrastive images, and half compared neutral and intensifying images. The items were rotated through all five continuum steps throughout the five blocks. This design resulted in a total of 150 trials.

Participants

Sixty-three participants were recruited for this experiment through Prolific (32 in the duration manipulation and 31 in the pitch manipulation). The participants were all between the ages of 18 and 65 and reported normal-to-corrected hearing and vision in addition to no reading or learning impairments. All participants were also fluent speakers of American English who were born in the United States. Due to a failure to complete the entire experiment, three participants were excluded from the analysis, resulting in a total of 60 participants.

Analysis and Results

The data were coded according to whether the participants matched a prosodic profile to its original pragmatic intent. For example, given a trial with a contrastive prosodic profile, responses were coded with a 1 if participants selected the contrastive image and a 0 if they selected the intensifying image. Trials characterized by neutral prosodic profiles were excluded
from the analysis. However, results from these trials, accompanied by a brief discussion, can be found in Appendix A.

First, a logistic mixed effects regression predicted the probability of a match from the effects of continuum step (coded as continuous and centered at 0), prosodic profile (scaled-sum-coded with contrastive mapped to -0.5 and intensifying to 0.5), continuum type (scaled sum-coded with pitch mapped to -0.5 and duration to 0.5), and the three-way interaction between continuum step, prosodic profile, and continuum type. The model also included fixed effects for the continuous predictors of original onset duration and preamble duration (measured through the word ‘is, ‘in ‘The Noun is Adj’), which were z-scored using by-participant means and standard deviation. Finally, the model included random intercepts for adjective and participant.
**Figure 10**

*Predicted Probability of a Match in Experiment 2*

![Graph showing predicted probability of a match for duration and pitch conditions.](image)

*Note.* Predicted likelihood of matching a given prosodic profile with its originally intended interpretation at each step of the duration and pitch continuums.

As shown in Figure 10, matching on critical trials was overall very high, especially for the stimuli featuring contrastive prosodic profiles. In other words, participants consistently associated the contrastive and intensifying prosodic profiles with interpretations of contrast and semantic intensity, respectively. While the main effect of continuum step on matching was not significant ($\beta=-0.01, z=-0.64, p=0.5$), a significant interaction between continuum step and continuum type suggests that the effect of continuum step on the likelihood of a match was more prominent in the duration condition than in the pitch condition ($\beta=-0.12, z=-1.7, p<0.05$). A
significant interaction between continuum step and prosodic profile illustrates that continuum step had a larger effect on the likelihood of a match for the intensifying prosodic profiles than the contrastive prosodic profiles ($\beta=0.324$, $z=5.196$, $p<.001$). Finally, there was also a significant three-way interaction ($0.514$, $z=4.123$, $p<0.001$). This interaction supports the pattern evident in Figure 10: Namely, that continuum step had a greater effect on the likelihood of a match for intensifying prosodic profiles in the duration condition than it did in the pitch condition.

To better understand whether there was a main effect of continuum step in the duration and pitch conditions, respectively, mixed effects logistic regression models were applied to the duration and pitch conditions separately. The first model, applied to the duration condition, predicted the likelihood of a match from the fixed effects of continuum step and prosodic profile, in addition to the interaction between continuum step and prosodic profile. The model also included random intercepts for adjective and participant. The model indicated that the effect of continuum step on the likelihood of a match was not significant ($\beta=-0.07$, $z=-1.84$, $p=0.06$). However, the likelihood of a match was significantly lower for the intensifying prosodic profiles ($\beta=-0.86$, $z=-3.68$, $p<.001$). Furthermore, there was a significant interaction between continuum step and prosodic profile, with matches increasing to a greater degree with continuum step for an intensifying prosodic profile compared to a contrastive one ($\beta=0.57$, $z=7.38$, $p<.001$). In other words, increasing duration from an intensifying baseline increased listeners’ association of intensifying prosodic profiles (which featured lower pitch on the critical adjective in addition to differences in the speech rate of the preamble) with interpretations of semantic intensity. A visual inspection of the data indicates that the impact of duration continuum step on the likelihood of a match differed for the contrastive prosodic profiles compared to intensifying ones. For example,
the predicted probability of a match for intensifying prosodic profiles increased from 0.77 at step -2 of the continuum to 0.89 at step 5, which is equal to a positive change of 0.12. For the contrastive prosodic profiles, however, the predicted probability of a match decreased from 0.96 to 0.86, a change of equal magnitude. In other words, increasing duration increased listeners’ association of intensifying prosodic profiles with interpretations of semantic intensity, while increasing duration decreased the association of contrastive prosodic profiles with interpretations of contrast.

The second model, applied to the pitch condition, was exactly the same as the model applied to the duration condition. It predicted the likelihood of a match from the fixed effects of continuum step and prosodic profile in addition to the interaction between continuum step and prosodic profile. The model indicated that there was not a significant effect of continuum step on matches (β=0.03, z=0.7, p=0.483) nor was there a significant interaction between continuum step and prosodic profile (β=0.06, z=0.702, p=0.483). However, there was a significant effect of prosodic profile, with the likelihood of a match being overall lower for the intensifying profiles than for the contrastive ones (β=-2.59, z=-10.117, p<.001). Unlike in the duration condition, the effect of continuum step on the intensifying vs. contrastive prosodic profiles was similar. The predicted probability of a match for the contrastive prosodic profiles was about 0.98 at all five steps, while the predicted probability of a match for the intensifying prosodic profiles increased slightly by about 0.4. Overall, the results from this model suggest that manipulating the pitch scaling across an entire pitch contour did not influence listeners’ associations of variations in duration and intensifying and contrastive pitch profiles with interpretations of semantic intensity and contrast.
Discussion

The results from this experiment suggest that, when given more natural stimuli that introduce smaller variations in pitch and duration, participants are exceptionally good at distinguishing prosodic profiles intended to convey contrast from prosodic profiles intended to convey semantic intensity. In other words, contrastive and intensifying prosodic profiles always sounded contrastive/intensifying to participants, regardless of their duration or the scaling of their entire pitch contour. The fact that both the duration and pitch manipulations had only a minimal effect on the likelihood of a match suggests that, at least in this experiment, pitch differences related to the entire pitch contour, such as its shape or overall duration, played the primary role in distinguishing the two meanings.

Although the effect was small, manipulating duration did appear to influence listener’s associations of contrastive and intensifying prosodic profiles with interpretations of contrast and semantic intensity, respectively. For example, increasing duration increased listeners’ association of intensifying prosodic profiles with interpretations of semantic intensity, while increased duration decreased listeners’ association of contrastive prosodic profiles with interpretations of contrast. Because increased duration is assumed to increase prominence, this result suggests that it is not prominence alone that distinguishes the two meaning types. Otherwise, increasing duration for either prosodic profile, contrastive or intensifying, would have increased the likelihood of a match. Thus, the results suggest that the two types of meanings are associated with distinct acoustic realizations, rather than distinct overall levels of prominence.

The results from the pitch condition are particularly interesting, as they suggest that increasing the pitch scaling of the entire intensifying prosodic profile slightly increased listeners’
association of this profile with interpretations of semantic intensity, while it had no significant effect for contrastive prosodic profiles. This finding may be related to previous suggestions that overall pitch scaling is more closely related to changes in paralinguistic meaning, such as emotion (Niebuhr, 2010; Wickens & Perry, 2015). For example, Bänziger and Scherer (2005) find that the level of arousal underlying an emotion affects the global level and range of F0 contours. Similarly, Busso, Lee and Narayanan (2009) find that gross pitch contour statistics like F0 mean, maximum, minimum and range are more consistent with emotionally charged speech than changes in F0 shape. Intensity as a meaning category has been shown to be a pragmatic and expressive phenomenon in addition to a semantic one. For example, Beltrama (2018) argues that lexical intensifiers like *totally* can target lexical scales encoded in the denotation of gradable predicates but can also target pragmatic scales related to a speaker’s commitment to a proposition, primarily when the predicate being targeted by the intensifier does not lexicalise a bounded scale, or when the predicate is evaluative/subjective. Furthermore, intensifying suffixes like *-issimo* in Italian can sometimes function as mixed expressives (Gutzmann, 2011), whereby they indicate scalar information about a speaker’s attitude while also conveying information about semantic intensity (Beltrama & Bochnak, 2015). Considering the close connection between intensity, attitude/emotion and prosody, it is possible that the reason pitch scaling had a larger effect on stimuli featuring an intensifying prosodic pattern (compared to a contrastive one) was because this manipulation targeted the affective dimension of meaning, which is more relevant in an intensifying context than in a contrastive one. Since higher mean F0 values have been shown to be associated with higher degrees of emotional arousal (e.g., Bänziger & Scherer, 2005), raising the pitch scaling of the entire utterance in this experiment might have increased the emotional strength of the utterance as a whole, leading to a stronger connection with the
meaning category of intensity. Interpretations of contrast, on the other hand, are presumably influenced by changes at the phrasal level, when the adjective itself has higher pitch relative to the other word around it. This dimension was not explicitly manipulated in the experimental paradigm used here.

Alternatively, it may be difficult for participants to perceive differences in phrasal prominence when pitch scaling is manipulated over an entire utterance. Because the intensifying stimuli were scaled lower from the start (i.e., even before the manipulation), it’s possible that increasing the pitch scaling of the entire utterance increased matches because of initial perceptual difficulties at the lower end of the pitch manipulation continuum for these stimuli, when participants were less likely to hear the critical adjective as standing out from the other words around it. Increasing the pitch scaling, therefore, may have simply made it easier to hear the critical adjective as prominent for the stimuli featuring intensifying intonational tunes.

**Overall Discussion**

In this study, I investigated whether different acoustic manifestations of prosodic prominence are associated with interpretations of contrast vs. semantic intensity. In Experiment 1, I tested whether lengthening could increase the likelihood of an intensifying interpretation vs. a contrastive one, for both intensifying (low) and contrastive (high) prosodic profiles, and whether it does so to an even greater degree for a compatible pitch profile (intensifying). The results suggest that increased duration does indeed increase the likelihood of an intensifying response. However, the effect of duration was evident for both contrastive and intensifying prosodic profiles, suggesting that interpretations of semantic intensity are no more tied to
intensifying prosodic profiles they are to contrastive ones, especially when the duration is particularly long.

In Experiment 2, I investigated how more subtle variation in duration affects listeners’ association of contrastive and intensifying prosodic profiles with interpretations of contrast vs. semantic intensity. The first important finding from this experiment was that intensifying and contrastive prosodic profiles (featuring lower vs higher average pitch, respectively) were consistently associated with interpretations of semantic intensity vs. contrast, respectively. This finding contrasts with the finding from Experiment 1 by suggesting that intensifying prosodic profiles are a reliable cue to semantic intensity. The reason for this discrepancy in the results might stem from the fact that, in Experiment 2, I did not use splicing to control the material appearing in the preamble of each utterance, instead keeping each adjective with its naturally produced preamble, thereby leaving intact the entire pitch contour and any voice quality or duration characteristics associated with it. Thus, it is possible that the strong association of the intensifying and contrastive prosodic profiles with the correct (originally intended) interpretation in Experiment 2 is more-so due to one of these other dimensions rather than the simple effect of high vs. low pitch on the critical adjective.

In Experiment 2, I hypothesized that increasing the duration of an intensifying prosodic profile would increase listeners’ associations of this tune with interpretations of semantic intensity. I also manipulated the scaling of the entire pitch contour, to investigate how this affected listeners’ associations of contrastive and intensifying intonational tunes with interpretations of contrast and semantic intensity. The results provide support for the hypothesis that increasing duration increases participants’ associations of intensifying prosodic profiles with
interpretations of semantic intensity. Furthermore, they suggest that increasing duration decreases associations of contrastive prosodic profiles with interpretations of contrast. The results also do not suggest that the scaling of the entire pitch contour has much of an effect on the association of contrastive and intensifying prosodic profiles with interpretations of contrast and semantic intensity, as neither profile was significantly more or less likely to be associated with its respective interpretation when its pitch contour was lowered or raised.

The results of this study support findings from production studies in German indicating a connection between semantic intensity and (1) lowered pitch trajectories on the critical adjective and (2) increased duration. One possibility for why these two cues seem more closely connected to semantic intensity than they do to contrast is that both increased duration and lowered pitch convey semantic intensity via an iconic connection. For example, it is generally assumed that the relationship between form and meaning is arbitrary (Saussere, 1916). Yet, recent investigations into sound symbolism (e.g., Blasi, Wichmann, Hammarström, Stadler, & Christiansen, 2016; Perniss, Thompson, & Vigliocco, 2010; Winter & Perlman, 2021; Winter, Soskuthy, Perlman & Dingenmanse, 2022), co-speech gesture (Tieu, Pasternak, Schlenker & Chemla, 2017; Ebert, Pirillo & Walter, 2022), and ideophones (Barnes, Ebert, Hörnig, Stender & Barnes, 2022) suggest that iconic aspects of language play an important role in languages across the world. In fact, pragmatic theories of iconicity suggest that iconic elements contribute information that is not at-issue (Potts, 2005), and that they constitute part of what is implicated (Grice, 1975). Exactly what this contribution is remains a matter of debate, as some suggest it contributes something similar to expressives and appositives (Ebert & Ebert, 2014) while others argue that they function like presuppositions (Schlenker, 2018). Regardless of their specific contribution, the
fact that iconic components of language contribute information that is not at-issue suggests that reasoning about intensification based on prominence, despite being iconic, is still a type of pragmatic meaning calculation, and is thus relatively similar to the meaning computation that occurs for contrast.

It could be argued that lengthening a word to convey a higher degree of some adjectival quality constitutes an iconic contribution because it directly instantiates the degree of some quality in the length of the word. Indeed, Schlenker (2018) argues that lengthening the vowel in a scalar adjective has an iconic effect that is a direct mapping of the length of the talk. In other words, he argues that 'every sound unit maps onto a signified extra size unit' (Guerrini, 2020: 244). This argument is supported by findings that in a written, online social media corpus, lengthening is more acceptable for large-type adjectives (e.g., *big*) than for small-type adjectives (e.g., *tiny*), as an iconic contribution related to larger size would be incompatible with an adjective denoting small size (Fuchs, Savin, Solt & Ebert, 2019). Nygaard, Herold and Namy (2009) also find that when participants were asked to produce novel words (e.g., *blicked, seebow, daxen*) assuming they had a meaning associated with greater size (e.g., *big, tall*) their productions were longer than their antonyms (e.g., *small, short*), suggesting that speakers do often utilize duration as an iconic resource for cuing meaning related to size. Finally, lower pitch has also been shown to be associated with bigger size (Ohala, 1984), suggesting that the lower pitch maxima characteristic of the stimuli used here may also convey semantic intensity via an iconic connection. It is important to note that the adjectives utilized in this study did not all directly address size (e.g., *stripedness, rainyness*, etc.). However, I suggest that the iconic connection here may generalize to a relation between increased duration, lower pitch and any increase in degree, regardless of the specific dimension being referenced by a gradable adjective.
A second reason for the connection between increased duration and interpretations of intensification could relate to the notion of M-implicatures, which are based on Grice’s (1975) Manner Maxim and utterances that differ in their complexity. Under this model, more complex constructions are assumed to receive marked interpretations (Levinson, 2000). For example, Bennett and Goodman (2018) investigate the ability of lexical intensifiers varying in length (as measured by their number of letters) to shift the interpretation of a scalar adjective, arguing that intensifiers contribute nothing to the literal semantics of scalar adjectives, but that they encourage inferences about higher degree values by increasing the cost of an utterance. Their results support this hypothesis by indicating that intensifiers with more letters implied higher values on the measurement scales introduced by the scalar adjectives used in their study.

It could be argued that producing a word with increased duration to convey intensification is an especially marked choice for several reasons. For example, like contrast, prosody is an optional means for conveying semantic intensity. Just as contrast can be conveyed using focus-sensitive operators like *only*, semantic intensity can be more simply and clearly conveyed using a lexical intensifier like *really*. Thus, using increased duration to convey intensity may be viewed as a marked choice because it is unnecessary and less clear than its lexical alternatives. At an acoustic level, the pitch movements characteristic of intensity could also be seen as more marked. For example, Kugler and Genzel (2012) argue that perceptions of prominence might be better thought of as resulting from deviation from a neutral register rather than a specific directional change. In their study of the African language of Akan, they find that users often lower their pitch, rather than raise it, in order to draw the listener’s attention to a specific part of their utterance. The lowered pitch trajectories characteristic of prominence conveying intensity in the experiments discussed in this chapter, then, might be perceived by
listeners as a particularly marked choice relative to the higher pitch trajectories characteristic of contrast.

A third potential explanation for the connection between duration and semantic intensity lies in its further connection to emotional and attitudinal meaning. Prosody is often used to convey emotion/arousal in addition to referential meaning related to information structure. For example, Ladd writes that prosodic prominence “conveys some sort of special emotional involvement on the part of the speaker” (2008: 129). Furthermore, lexical intensifiers have been shown to be capable of modifying pragmatic and attitudinal scales in addition to semantic ones (Beltrama, 2018; Beltrama & Bochnak, 2015; Irwin, 2014), and are sometimes capable of expressing both simultaneously. Prior investigations of prosodic intensification suggest that, similar to lexical intensifiers, prosodic intensification may target semantic scales introduced by gradable adjectives while also signaling information about a speaker’s emotions or attitudes. For example, Niebuhr (2010) suggests that, in German, the pitch trajectory and location of segment lengthening is sensitive to the attitude a speaker intends to convey, with positive intensity featuring lengthening in the nucleus and negative intensity featuring lengthening in the onset. Landgraf (2014) confirms that German listeners perceive these different realizations as indicating positive vs. negative attitudes. Furthermore, she shows that pairing specific instantiations of prominence with words that have mismatching positive (e.g., wonderful) vs. negative (e.g., deplorable) semantics can result in the perception of irony. Overall, since the notion of intensification seems to be highly associated with attitudinal meaning, and because prosody is capable of conveying both propositional and paralinguistic meaning, it is an interesting question which type of degree scale the increases in duration implemented in this experiment were primarily targeting. On the one hand, it is possible that the pitch and duration manipulations
implemented here did indeed provide direct information to listeners about the location of a semantic threshold. However, it is also possible that these cues primarily signaled information about a speaker’s affective stance or emotional arousal, and that inferences about semantic intensity that listeners made were secondary inferences based on a listener’s knowledge of this information. Future research might consider this multi-faceted nature of intensification when investigating its prosodic realization. For example, since prosodic prominence can also convey expressive content related to attitudes and emotions, it is an interesting question whether listeners associate different forms of prominence with semantic vs. pragmatic/expressive intensity.

The results discussed in this chapter also have interesting implications for the mapping between prominence and meaning. For example, it is generally accepted that movement along information structural scales related to givenness (Baumann, 2006) and contrast corresponds with increases in acoustic prominence (Im et al., 2023). More prominent pitch accents are associated with the higher end of these scales (i.e., with newness and correction, respectively), while less prominent pitch accents are associated with the lower end (i.e., givenness and new information focus, respectively). Furthermore, the two scales do not seem to be directly parallel in terms of their relation to prominence. For example, while contrastive focus can be realized on both new and given information, pitch accents marking higher degrees of contrast (i.e., correction) are assumed to be more prominent than those marking higher degrees of givenness (i.e., newness). It is therefore interesting to consider where semantic intensity lies in relation to these two scales, in addition to whether or not it is itself as a meaning category gradient in any way. For example, if semantic intensity were to lie higher on the prominence scale than contrast, I would expect that the stimuli characterized by higher pitch and increased duration would be the
most likely to convey semantic intensity. However, my results suggest that this is not necessarily the case: rather, pitch movements that were scaled lower tended to be more associated with interpretations of intensity. This finding suggests that semantic intensity falls lower on the prominence hierarchy than contrast. However, future studies should investigate how interpretations of contrast and semantic intensity are mediated by variation in the overall degree of prominence, in an effort to better understand how intensity as a meaning category is distinct from contrast and givenness in terms of its form-function mapping.

Another interesting possible avenue of research involves the investigation of intensification accents realized in alternative sentence types. For example, Seeliger and Repp (2023) suggest that in polar interrogatives, information structure is marked differently than in assertions. Since polar questions are typically realized with rising pitch contours, some researchers propose that increasing prominence to convey contrast in polar interrogatives might involve lower pitch rather than higher, because prominence might be based on deviation from a pitch baseline for a given sentence type (e.g., Kügler & Genzel, 2012; Repp, 2020). For assertions, this baseline is low, but for questions, this baseline is high. Descriptions of the realization of prosodic intensification thus far remain limited to declaratives (as in the experimental materials used here). Therefore, future studies might examine the characteristics of intensifying prominence produced in alternative types of sentence types, such as polar interrogatives.

Conclusion

The results of this study indicate that listeners associate distinct realizations of prominence with interpretations of contrast and semantic intensity. In Experiment 1, I showed
that gradable adjectives featuring increased duration were more likely to be interpreted as conveying semantic intensity rather than contrast, regardless of whether they were produced with a pitch pattern over the adjective (prosodic profile) that was intended to convey intensity (lower pitch) or one intended to convey contrast (higher pitch). In Experiment 2, I also found that duration is more closely tied to interpretations of semantic intensity, and that pitch patterns across the entire sentence that were intended to convey intensity vs. contrast were also consistently associated with interpretations of semantic intensity. I also found that longer duration was more characteristic of an intensifying prosodic profile than a contrastive one. Specifically, increasing duration of the adjective decreased the likelihood that participants heard intensifying prosodic profiles as conveying semantic intensity rather than contrast. Conversely, increasing duration on adjectives produced with contrastive prosodic profiles decreased the likelihood that they were heard as contrastive. These differences suggest that prosody plays a role in conveying pragmatic meaning that is specific to a given meaning category, and that the realization of prominence can differ depending on the pragmatic meaning a speaker intends to convey.
Chapter 2: The Relative Contribution of Onset and Nucleus Duration to Interpretations of Contrast and Semantic Intensity

Introduction

While it is widely accepted that prominence has a pragmatic meaning function in Mainstream American English, work to date has primarily focused on its ability to convey aspects of information structure such as givenness and focus. Yet, anecdotal descriptions from early work on MAE and French intonation (Bolinger, 2013; Jakobson, 1960; Armstrong & Ward, 1926; Coustenoble & Armstrong, 1934) and recent experimental work in German (Kohler, 2006; Niebuhr, 2010) suggest an additional role for prominence in conveying semantic intensity for gradable adjectives. For example, producing the word tall in the tree is tall with prosodic prominence might convey that the tree is exceptionally tall. If this use of prominence in MAE is confirmed, it raises the question as to whether the distinct meaning functions of prominence (e.g., information structure, semantic intensity) correspond to distinctions in the phonological and/or phonetic correlates of prominence, and it also calls for a re-examination of the association between prominence (phonetic and phonological) and meaning.

In this chapter, a series of experiments tests the relative relation of prominence patterns over the phrase to interpretations of contrast and semantic intensity. The experiments investigate whether prominence expressed through F0 and duration patterns typically associated with information structure vs. semantic intensity are stronger cues to either interpretation. Furthermore, they investigate whether lengthening overall or asymmetric lengthening in the onset vs. nucleus serve is a stronger cue to either meaning type. The results from this study show
that different types of pragmatic meaning are variably associated with different acoustic realizations of prominence. In a follow-up experiment, I confirm that this effect is not mediated by prominence, as more prominent realizations were not always more likely to be associated with either pragmatic meaning type. These results suggest that the mapping between prominence and meaning cannot be reduced to a simple relation between increased perceptual salience and stronger meaning types. Rather, listeners appear to attend to different aspects of the acoustic realization of prominence when interpreting it as a cue to information structure vs. semantic intensity.

The Interpretation of Prosodic Prominence

In MAE, phrasal prominence often conveys focus, which, according to Rooth (1985, 1992) introduces alternatives to the focused expression. Specifically, focused expressions are said to have two meaning components: an ordinary value (i.e., a usual meaning according to compositional semantics) and a focus semantic value, which is a set of elements that match the focused expression in semantic type (Rooth, 1992). For example, focusing the word some in (6) is said to signal to the listener that alternatives contained in the focus semantic value for some (in this case, alternative quantifiers like all) are relevant to the interpretation of the utterance.

(6) Only \([\text{some}]_F\) of the students were late.

In the case of (6), the relevant alternatives are included in the denotation of the emphasized expression, as a quantifier like some can be associated with an entailment-based Horn scale of ordered alternatives, such as most or all (Horn, 1972). However, in the case of (7), the alternatives must be inferred from the context.

(7) Only \([\text{Mary}]_F\) was late.
For example, the listener might infer that in this case, the alternatives are other individuals who could possibly have been late, such as John or Amy.

Phrasal prominence has also been linked to the meaning category of semantic intensity. Within the semantic literature on gradable adjectives, intensification is typically described using threshold semantics (Kennedy & McNally, 2005). According to this theory, gradable adjectives like *tall* are associated with measurement scales of ordered degree values, as well as thresholds (θ) on these scales above which a given entity qualifies as true of an adjective. For example, to qualify as *tall*, a tree would need to be at least as tall as the threshold θ\textsubscript{tall}. Thresholds can be fixed by measure phrases, such as in a *five-foot tall tree*. In this case, the degree of height is explicitly specified by the measure phrase. However, modifiers like intensifiers can also alter thresholds, although they do not fix them to a specific degree value. For example, it is generally argued that a gradable adjective that is modified by an intensifier has the same semantics as an unmodified adjective, but with a new and increased threshold (Kennedy & McNally, 2005; Klein, 1980; Wheeler, 1972). In other words, a *very tall tree* is one that presumably has a height that is above the (higher) threshold for *tall* among a set of tall trees. Determining the exact value for all thresholds, whether intensified or not, involves pragmatic reasoning based on the context. For example, what counts as *tall* or *very tall* for a tree in a nursery may be very different from what constitutes *tall* or *very tall* for a tree in a forest.

It is important to note that, unlike focus, semantic intensity and the way in which it maps onto prosodic prominence has not been explicitly modeled or formalized. Rather, formal accounts of intensification remain restricted to lexical and syntactic strategies (e.g., Kennedy & McNally, 2005; Beltrama & Trotzke, 2019). However, various observations do suggest that
intensification can also be conveyed by prosody (Bolinger, 2013; Jakobson, 1960), and in fact, Beltrama and Trotzke (2019) mention in their paper discussing lexical and syntactic means for conveying intensification that phonetic strategies for conveying intensification, such as “emphatic prolongation” (Jakobson, 1960: 354), should be discussed and modeled in future studies. Thus, in this chapter, I assume that the semantic effects of prosody regarding intensity are similar to those that have been proposed for lexical intensification, although future studies should more clearly characterize this relationship.

If prosodic prominence can convey both focus and semantic intensity, it remains unclear how a listener can determine which meaning a speaker intends to convey when prominence is produced on a gradable adjective like tall. For example, in this situation, prominence might convey contrastive focus, leading a listener to conclude that the tree is tall as opposed to some other alternative (e.g., short, skinny, gigantic). On the other hand, prominence might convey semantic intensification, suggesting that the tree is particularly tall. In Chapter 1 of this dissertation, I present results suggesting that, in the absence of a disambiguating context, listeners can rely on differences in the acoustic realization of prominence to distinguish between these two meaning types. Therefore, in the following section, I review related findings from prior work on both the general realization of prosodic prominence and the specific manifestations that have been associated with contrast and semantic intensity, respectively.

The Acoustic Realization of Prominence

In production, prominence is implemented though various acoustic parameters, including F0, duration, and intensity (as a measure of acoustic energy) (e.g., Beckman, 1986; Kochanski, Grabe, Coleman & Rosner, 2005; Breen, Fedorenko, Wagner & Gibson, 2010; see Terken &
Hermes, 2000 for an overview). A variety of studies have confirmed that these SIGNAL-BASED factors lead a listener to perceive a word as prominent (e.g., Hermes & Rump, 1994; Fry, 1955, 1958; Van Katwijk, 1974; Cole et al., 2010; Mahrt, et al., 2012; Cole, et al., 2015; Cole et al., 2019; Baumann & Winter, 2018; see Ladd & Arvaniti, 2023 for a recent review). For example, from a phonological perspective, Autosegmental-Metrical theory claims that prominence results from the positioning of linguistic elements in metrically strong positions within a hierarchical prosodic structure (Liberman, 1975; Pierrehumbert, 1980; Ladd, 2008). Perception studies suggest that listeners are biased to hear a word with a nuclear accent (the rightmost accented word in the prosodic phrase) as prominent (e.g., Cole et al., 2010; Cole et al., 2019), independent of acoustic properties typically associated with prominence such as an F0 movement (Hermes & Rump, 1994).

Distinctions in the degree of phrasal prominence result from categorical distinctions between phonological pitch accents, or dynamic pitch movements that associate with elements appearing in metrically strong positions (Bolinger, 1958). Pierrehumbert (1980) proposes an inventory of pitch accents that includes three distinct types of rising accents in English (H*, L*+H, and L+H*), distinguished in the presence of a low target preceding the accentual peak and in the alignment of the accentual peak with respect to the stressed syllable, along with a low (L*) and falling accent (H+L*; later analyzed as ending in a downstepped high: H+!H*) (Beckman & Pierrehumbert, 1986; Beckman, Hirschberg, & Shattuck-Hufnagel, 2005). These pitch accents differ from one another in their acoustic prominence. For example, the greater pitch span of L+H* compared to H* results in corresponding differences in perceptual prominence ratings. Cole et al. (2010), Cole et al. (2019), and Im, Baumann and Cole (2023) all find evidence that, in
a Rapid Prosody Transcription task, the L+H* pitch accent was the most likely pitch accent type to be marked as prominent by participants. Similarly, Bishop et al. (2020), using a similar prominence rating task, illustrate that the likelihood of perceived prominence based on their participants’ responses varied according to pitch accent in the order of L*< H*<L+H*.

The specific realization of prominence conveying focus, and types of focus conveying stronger degrees of contrast in particular, has been well-investigated across languages. It is typically marked with a pitch accent, with acoustic correlates in measures of F0, duration, and intensity (Jackendoff, 1972; Truckenbrodt, 1995). Pierrehumbert and Hirschberg (1990) suggest a mapping between accent type and focus types: For example, they propose that L+H* marks contrastive focus while H* marks new information or broad focus (see Breen et al., 2010 as well). Corpus studies have failed to find a systematic correspondence between focus and choice of H* vs. L+H* (Hedberg & Sosa, 2006; Calhoun, 2004, 2006, 2012), but experimental evidence does suggest that in perception, the two accents have different implications for interpretation. For example, evidence from eye-tracking studies indicate that L+H* activates contextually provided contrastive alternatives relative to H* (Watson, Tanenhaus & Gunlogson, 2008; Braun, Asano & Dehé, 2019, Ito and Speer, 2008). Watson, Tanenhaus and Gunlogson (2008) show that, given a visual world paradigm, listeners’ eye movements anticipated contrastive targets upon hearing L+H*, while they were equally likely to look at either new or contrastive targets when hearing H*. These results suggest that, while the two accents may not be categorically distinct in terms of their meaning function, H* may be a more general prosodic marker for the discourse salience of a referent, while L+H* might bias the listener towards an interpretation relative to an alternative referent. Similarly, Ito and Speer (2008) report on an eye tracking study showing that L+H* can lead listeners on a garden path, where they anticipate an alternative referent before hearing
conflicting information in the following noun that identifies an unrelated referent. Finally, Tomlinson et al. (2017) illustrate that L+H* not only leads a listener to focus on contrastive alternatives but can encourage exhaustive inferences about the exclusion of those alternatives. For example, they show using a mouse-tracking paradigm that exhaustive implicatures were derived earlier in sentence processing with an L+H* pitch accent than with H*.

The specific realization of prominence conveying semantic intensity is relatively understudied. For example, beyond the observations made by Bolinger (2013) and Jakobson (1960) about semantic intensity’s connection to increased duration, early work by Armstrong and Ward (1926) and Coustenoble and Armstrong (1934) suggest that semantic intensity can be characterized by either a) rising-falling F0 peaks featuring high plateaus, in addition to considerable vowel lengthening or b) narrowed and lowered pitch trajectories featuring short, falling F0 movements. In a production study conducted in German, Niebuhr (2010) elicits semantic intensity conveying positive vs. negative attitudes and concludes that these two realizations can be mapped onto a distinction in attitudinal valence, with the latter being consistent with negative intensity and the former with positive. Furthermore, he investigates several other acoustic properties that might be characteristic of each type. For example, his analysis suggests that positive intensity is characterized by quickly articulated onset consonants and lengthened vowel nuclei, extended high F0 plateaus featuring steep rises and shallow falls, and sonorous, slightly breathy voice. Negative intensity, on the other hand, was characterized by lengthening of the preceding sound and onset consonant along with quick and abrupt articulation of the vowel nucleus, a pointed F0 peak that fell sharply right after the vowel’s onset and pressed and irregular voice quality. In a perception study conducted by Landgraf (2014), these characteristics were indeed interpreted as sounding more intensifying by German listeners.
Furthermore, they were correctly identified as conveying negative vs. positive attitudes (Landgraf, 2014). Overall, these initial investigations into the realization of prosodic prominence conveying semantic intensity suggest that its acoustic realization may differ significantly from that of contrastive focus.

In Chapter 1 of this dissertation, my results suggest that listeners do indeed associate these different realizations of prominence with semantic intensity vs. contrast. Yet, as previous perception studies for focus suggest, the mapping between prominence realizations (e.g., pitch accent types) and meaning (e.g., focus types) is not one-to-one. For example, Watson, Tanenhaus and Gunlogson (2008) illustrate that both H* and L+H* accents are compatible with a contrastive interpretation, although L+H* may be even more compatible than H*. In other words, certain realizations may bias an individual towards an interpretation slightly more than an alternative realization (Watson, Tanenhaus and Gunlogson, 2008). In the experiments conducted in Chapter 1, I forced listeners to make a binary distinction between contrast and semantic intensity for a single stimulus differing in its prosodic profile and duration, without allowing for the possibility that the stimuli could be variably associated with both interpretations. Therefore, in the experiments discussed here, I build on the results from Chapter 1 by asking participants to compare various realizations of prominence as best fitting an interpretation of contrast, and similarly for semantic intensity. By doing so, I hope to gain insight into the realization of prominence most likely to be associated with each meaning type.

**The Current Study**

In this study, I investigate the effect of variation in onset/nucleus duration on the interpretation of utterances with intensifying and contrastive prosodic profiles (i.e., utterances
naturally produced with the intent to convey semantic intensity and contrast). In two experiments, I ask participants to judge which of two prosodic realizations best represents an interpretation of contrast vs. semantic intensity (operationalized as sounding like it conveys a greater degree of the quality being denoted by the adjective or as sounding more like a correction). Stimuli were created from natural productions of sentences containing scalar adjectives in two prosodic conditions: with intonational properties (including pitch, duration, and voice quality) characteristic of a contrastive adjective or with properties characteristic of an adjective that conveys semantic intensity.

**Experiment 1a – 1b**

*Materials*

A total of six gradable adjectives were used as stimuli for this study: *bright, mad, sour, happy, salty, and windy*. These adjectives are a sub-set of adjectives used as stimuli in Chapter 1 and were selected from amongst this set based on their high level of perceived naturalness after the manipulations discussed below. Each adjective was produced in the sentence frame “The NOUN is ADJ” by the author of this study, in two pragmatic contexts: with the intention to convey either semantic intensity or contrast. As shown in Figure 11, this subset of adjectives exhibited the same pattern of acoustic differences as shown by the full set of adjectives in Chapter 1: Namely, the duration of the nucleus and onset was longer for the intensifying stimuli (Figure 11a), the pitch maximum on the adjective was higher for the contrastive stimuli (Figure 11b), the speech rate of the preamble for the intensifying stimuli was slower (Figure 11d), and the intensifying stimuli exhibited slightly less modal voice quality (Figure 11c) and a slightly higher maximum F0 value in the preamble (Figure 11e).
Figure 11

*Acoustic Measures of the Unmanipulated Experimental Stimuli*

*Note.* The graphs in this figure illustrate the average segment duration (a), max pitch values (b) and average H1-H2 values (c) for each of the six adjectives prior to manipulation, in addition to the speech rate (d) and max F0 (e) of the preamble material.
Overall, these initial productions are in line with the prosodic characteristics of semantic intensity discussed by Niebuhr (2010) for German. Furthermore, in Chapter 1 of this dissertation, I confirm that, in Experiment 2, at step 6 of the continuum (when the duration of the adjective was unmanipulated), these two realizations were correctly identified by participants as sounding contrastive vs. intensifying.

In this experiment, as in Chapter 1, the decision was made to leave the natural productions intact, without manipulating their pitch, in order to preserve all properties of the original productions other than those that were explicitly manipulated (duration), allowing for the possibility that cues for intensity or contrast are distributed before or after the adjective, and also to allow a more stringent test of durational variation on the adjective itself. For example, in addition to the differences in pitch scaling, duration, and voice quality as described above, the original productions also featured differences in the shape of the F0 movement across the entire sentence and on the adjective itself (Figure 12). These differences might also serve as cues to contrast and intensity, and thus were preserved in the signal. Thus, throughout the remainder of this chapter, I refer to these original pitch, duration, and voice quality characteristics as contrastive and intensifying PROSODIC PROFILES.
Figure 12

Average F0 Across Contrastive and Intensifying Prosodic Profiles

*Note.* The raw F0 trajectories for all six adjectives produced with the intent to convey contrast vs. semantic intensity. The portion of the contour outlined in blue represents the F0 trajectory over the material spanning from the beginning of the sentence through the word immediately preceding the critical adjective (e.g., ‘the shirt is’). The portion of the contour outlined in red represents the F0 trajectory over the critical adjective itself. The solid black line illustrates the average F0 trajectory for the contrastive and intensifying conditions.

For each adjective, four types of stimuli were created based on manipulating the onset and nucleus of each adjective featuring each prosodic profile to create a longer and shorter version of each. The short duration condition was set equal to a decrease in duration of 40% from the original contrastive duration, while the long duration condition was equal to an increase of 40% from the original intensifying duration. These durations were equal to the shortest and longest durations from the duration continuum implemented in Experiment 2 of Chapter 1. Crossing these two duration manipulations in two segments (onset and nucleus) with the two naturally
produced prosodic profiles led to the creation of eight unique stimuli for each adjective, as shown in Table 2.

**Table 2**

*Prosodic Profile x Duration Conditions in Experiment 1a-1b*

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Prosodic Profile</th>
<th>Onset Duration</th>
<th>Nucleus Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLL</td>
<td>Contrast</td>
<td>Long</td>
<td>Long</td>
</tr>
<tr>
<td>CLS</td>
<td>Contrast</td>
<td>Long</td>
<td>Short</td>
</tr>
<tr>
<td>CSL</td>
<td>Contrast</td>
<td>Short</td>
<td>Long</td>
</tr>
<tr>
<td>CSS</td>
<td>Contrast</td>
<td>Short</td>
<td>Short</td>
</tr>
<tr>
<td>ILL</td>
<td>Intensity</td>
<td>Long</td>
<td>Long</td>
</tr>
<tr>
<td>ILS</td>
<td>Intensity</td>
<td>Long</td>
<td>Short</td>
</tr>
<tr>
<td>ISL</td>
<td>Intensity</td>
<td>Short</td>
<td>Long</td>
</tr>
<tr>
<td>ISS</td>
<td>Intensity</td>
<td>Short</td>
<td>Short</td>
</tr>
</tbody>
</table>

*Note.* All eight unique stimuli for each adjective. The stimulus column indicates the abbreviation used to refer to each stimulus type throughout the remainder of this chapter.

**Design and Procedure**

Participants completed a 2 Alternative Forced Choice (2AFC) task that required them to select which of two stimuli sounded either a) like it conveyed a higher degree of the quality being denoted by a given adjective (Experiment 1a) or b) more like a correction (Experiment 1b).
Four of the comparisons were across-pitch: They involved one stimulus featuring each prosodic profile (e.g., one with a contrastive prosodic profile [higher pitch] and one with an intensifying prosodic profile [lower pitch]). The two other comparisons were within-pitch: They isolated the effects of duration while holding the prosodic profile constant, comparing the SS and LL conditions or the SL and LS conditions for stimuli featuring the same prosodic profile (e.g., ILL vs. ISS, ILS vs. ISL, CLL vs. CSS, CLS vs. CSL). These additional trials allowed us to test for a difference between lengthening overall or between lengthening of the onset vs. nucleus, respectively, for each prosodic profile.

Both Experiment 1a and Experiment 1b were between-subjects, with half of the participants hearing trials that compared all the durational variants of stimuli with intensifying prosodic profiles (e.g., ISS, ILL, ISL, and ILS) with the shortest and longest variants of stimuli with a contrastive prosodic profile (e.g., CSS and CLL). The other half of the participants heard trials that compared all durational variants of the contrastive stimuli (e.g., CSS, CLL, CSL, CLS) with the shortest and longest variants of the intensifying stimuli (ISS and ILL). This design yielded the critical comparisons indicated in Figure 13.

While a corrective meaning might be considered a specific sub-type of focus/contrast, and thus might exclude interpretations of contrast more generally, I chose to adopt this wording for the sake of clarity. Participants seemed much more likely to understand what it would mean for an utterance to sound like a correction than what it would mean for an utterance to sound ‘contrastive.’

I chose not to adopt the fully-crossed design because this would have yielded 56 unique combinations of different stimuli, which would result in an extremely long experiment. Furthermore, I did not have specific hypotheses about the likelihood of, for example, an ILS stimulus compared to a CSL stimulus.
Figure 13

Comparisons Tested in Experiment 1a-1b

Intensifying Half

<table>
<thead>
<tr>
<th>Across Pitch</th>
<th>ILL</th>
<th>ILS</th>
<th>ISL</th>
<th>ISS</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLL</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>CSS</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Within Pitch</th>
<th>ISS</th>
<th>ISL</th>
</tr>
</thead>
<tbody>
<tr>
<td>ILL</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>ILS</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

Contrastive Half

<table>
<thead>
<tr>
<th>Across Pitch</th>
<th>CLL</th>
<th>CLS</th>
<th>CSL</th>
<th>CSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ILL</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>ISS</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Within Pitch</th>
<th>CSS</th>
<th>CSL</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLL</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>CLS</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

Note. The comparisons that were tested in each half of the between-subjects experiment. The comparisons tested are represented with an X.

The eight comparisons were divided into two blocks such that all the trials in a single block involved comparisons of four durational variants of stimuli featuring one prosodic profile (intensifying or contrastive) with only one durational variant from the other prosodic profile. For example, for participants hearing all durational variants of the stimuli with an intensifying prosodic profile, Block 1 compared all four intensifying durations (ILL, ISS, ISL and ILS) to the CLL condition along with two comparisons within the intensifying stimuli (ILL vs. ISS and ILS...
vs. ISL), while Block 2 compared all four intensifying durations to the CSS condition, in addition to the same two within-intonation comparisons. This resulted in a total of six comparisons for each adjective in each block: four across pitch (e.g., ILL v. CLL, ISS v. CLL, ISL v. CLL, ILS v. CLL) and two within (e.g., ILS v. ISL, ISS v. ILL). Whether the contrastive or intensifying stimulus appeared on the left or the right was counterbalanced across blocks. Participants heard each block twice: once with each of the presentation orders.

Within each block, the stimuli were organized into six sub-blocks based on adjective, so that each adjective occurred once in each block, in a different trial condition. The adjectives were then rotated through all conditions over the course of the six blocks. Six adjectives produced in six sub-blocks, for each of 4 blocks, resulted in a total of 144 trials.

Participants completed the experiment remotely using their own laptop or desktop computer and headphones. Before beginning the experiment, they were required to test their headphones for sound quality. In Experiment 1a, testing interpretations of intensity, on each trial, participants were presented with two audio files and told to select the visual icon representing the utterance in which they thought the object being described exhibited a greater degree of the adjectival quality describing it (e.g., for the adjective “bright,” the participant would be asked which shirt was brighter). In the experiment testing interpretations of contrast (Experiment 1b), conducted with a separate set of participants, participants were asked to select the icon representing the utterance that they thought sounded more like a correction.

The four, complementary hypotheses for this study are as outlined below:

**Hypothesis 1:** Interpretations of contrast and semantic intensity are conditioned by different prosodic profiles. Prediction: Prosodic profiles featuring lower pitch will be a strong cue to
semantic intensity while prosodic profiles featuring higher pitch will be a strong cue to interpretations of contrast.

**Hypothesis 2**: An increase in word duration increases the likelihood of the word being interpreted as contrastive or intensified. Prediction: A word with longer duration is more likely to be identified as conveying contrast or semantic intensity, compared to a word with shorter duration.

**Hypothesis 3**: An asymmetric increase in duration restricted to the stressed syllable onset, while holding nucleus duration steady, increases the likelihood of an interpretation of semantic intensity, while a contrastive interpretation is not similarly influenced by asymmetric lengthening. Prediction: A word with only a lengthened onset or only a lengthened nucleus will be more likely to be identified as conveying semantic intensity than contrast, compared to a word with lengthening in both the onset and nucleus.

**Hypothesis 4**: The prosodic profile of an adjective has a stronger effect on the interpretation of semantic intensity or contrast than does durational lengthening (onset, nucleus, or both). Prediction: An interpretation of semantic intensity will be overall more likely for stimuli with intensifying (low-pitch) prosodic profiles, and similarly, an interpretation of contrast will be more likely for stimuli with contrastive (high-pitch) prosodic profile, regardless of whether the stimulus has longer or shorter duration.

**Participants**

A total of 180 participants were recruited from Prolific for this experiment, 90 for the experiment testing interpretations of semantic intensity (Experiment 1a) and 90 for the experiment testing interpretations of contrast (Experiment 1b). Each experiment was between subjects, with 45
participants participating in the contrast manipulation condition (with stimuli CLL, CSS, CSL, CLS, ILL, ISS) and 45 in the intensity manipulation (with stimuli ILL, ISS, ISL, ILS, CLL, CSS). The participants were all between the ages of 18-65 and self-reported as having normal-to-corrected hearing and vision and no reading or learning impairments. All participants were fluent speakers of American English who had been born in the United States. Due to a failure to complete the entire experiment, four participants were excluded from Experiment 1a, leading to a total of 86 participants. Furthermore, ten participants were excluded from Experiment 1b, resulting in a total of 80 participants.

Analysis and Results

For each experiment, I conducted separate analyses of the within-pitch and the across-pitch trials. The within-pitch analysis (e.g., CLL vs. CSS, ILL vs. ISS) targets Hypothesis 2: It tests for an effect of lengthened duration on the interpretation of both semantic intensity and contrast, for both intensifying and contrastive prosodic profiles. The across-pitch analysis (e.g., CLL vs. ILL, CLL vs. ISS, CSL vs. ISS) targets Hypothesis 1: It tests the effect of the prosodic profile (contrastive vs. intensifying) on the interpretation of contrast and semantic intensity for adjectives in all conditions of duration. Both analyses also target Hypotheses 3 and 4 by probing the interaction between duration and prosodic profile (i.e., whether increased duration is more/less compatible with an interpretation of semantic intensity when it is paired with high vs. low pitch). Each of these two analyses and their respective results are discussed in turn below.

---

4 The decision to analyze these trials separately was made because of the difference in the type of comparisons applied to each stimulus. For example, across both experiments, the CLL stimulus was compared to ILL, ILS, ISL, ISS and CSS stimuli, while the CLS stimulus was compared only to ILL, ILS, ISL, ISS and CSL stimuli. As shown in Fig. 3, the within-pitch trials compared only the LL vs. SS conditions and LS vs. SL conditions for a single prosodic profile (e.g., CLL vs. CSS) and did not feature across-pitch comparisons.
Within-Pitch Analyses. Data from Experiments 1a and 1b were pooled to test for an effect of duration, within each condition of prosodic profile, on the selection of items as conveying contrast or semantic intensity. Separate mixed effects logistic regression models were fit to both types of within-pitch trials (LL vs. SS and LS vs. SL). These models predicted the likelihood of selecting the longer stimulus (LL) or the stimulus featuring a longer nucleus (SL) from the fixed effects of prosodic profile (contrastive vs. intensifying), experiment (1a [semantic intensity] vs. 1b [contrast]), and order of presentation (1 vs. 2) in addition to the interaction between prosodic profile and experiment. The model also included by-participant and by-adjective random intercepts. The predictors were all sum-coded, with reference levels set to intensifying prosodic profiles and Experiment 1b (contrast), respectively.

Beginning with the LL vs. SS trials, as shown in Figure 14, the model suggests that the longer stimulus (LL) was much more likely to be selected than the SS stimulus, regardless of whether the stimulus featured an intensifying or contrastive prosodic profile or whether it was being interpreted as contrast vs. semantic intensity. For example, a significant intercept in the model suggests that, holding all factors constant, the LL stimulus had an 88% chance of being selected ($\beta_0=1.98$, $z=6.58$, $p<.001$). In fact, the model also suggests that there was not a significant difference between the likelihood of selecting the LL stimulus when produced with an intensifying (low) vs. contrastive (high) prosodic profile ($\beta_1=0.7$, $z=1.595$, $p=0.11$) or in Experiment 1a vs. Experiment 1b ($\beta_1=-0.08$, $z=-0.22$, $p=0.82$). There was also not a significant interaction between prosodic profile and duration ($\beta_1=0.16$, $z=0.215$, $p=0.830$). These support Hypothesis 1 by indicating that longer duration is more compatible with both interpretations of
contrast and semantic intensity. However, they do not support Hypothesis 2, as they suggest that this pattern is true regardless of prosodic profile, rather than vice versa.

**Figure 14**

*Predicted Probabilities of Selecting an LL and SL Stimulus*

![Predicted Probabilities of Selecting an LL and SL Stimulus](image)

*Note.* Predicted probabilities of selecting the LL (15a) or SL (15b) stimulus in both the intensifying and contrastive experiments along with the 95% CI around the predicted probability.

Moving on to the trials comparing LS and SL, I see that the difference in the likelihood of selecting the LS vs. SL stimulus was much smaller than the difference between LL and SS. The intercept of the model was not significant ($\beta_0=-0.28$, $z=-1.27$, $p=0.2$) and suggests that, holding all factors constant, the likelihood of selecting the SL stimulus was about 43%. Unlike for the LL
vs. SS trials, the model results also suggest that the likelihood of selection differed slightly depending on the experiment and on the prosodic profile the adjective was realized with. For example, while there was not a significant difference in its likelihood of being selected in Experiment 1a vs. 1b, nor when paired with contrastive vs. intensifying prosodic profiles, there was a significant interaction between prosodic profile and experiment ($\beta_1=1.22$, $z=2.988$, $p<.01$).

As shown in Figure 14, the likelihood that the SL stimulus was selected was significantly higher when realized with contrastive prosodic profiles in the intensifying experiment only. Conversely, the likelihood that LS stimulus was selected was higher in the intensifying prosodic profile condition than in the contrastive prosodic profile condition. In the contrastive experiment, there was not a large distinction between the likelihood of selecting either stimulus based on prosodic profile. Finally, within the stimuli featuring contrastive prosodic profiles (CLS, CSL), the likelihood of selecting SL decreased from Experiment 1a to Experiment 1b, while for the intensifying prosodic profiles, it increased. Taken together, these results suggest that the distinction between the LS and SL stimuli was more apparent for a) the stimuli featuring Intensifying prosodic profiles (lower pitch) and b) when participants were making judgments of semantic intensity. Furthermore, there was a difference in directionality: in Experiment 1a, SL was only more likely with contrastive prosodic profiles, while in Experiment 1b, it was more likely with intensifying prosodic profiles.

It is important to note that both models indicated a significant effect of order on the likelihood of selection, with the likelihood that LL ($\beta=0.46$, $z=4.844$, $p<.001$) or LS ($\beta=-0.45$, $z=-6.045$, $p<.001$) was selected being higher when it appeared on the right than on the left. This result appears to stem from a slight preference of the participants for selecting the stimulus on
the right (e.g., for the LS vs. SL trials, 54% of the responses involved selection on the right).
Assuming that participants initially played the two stimuli from left to right, it possible that, after hearing a baseline stimulus for comparison on the left, participants felt more comfortable selecting the stimulus on the right or were at least less likely to return their cursor back to the left to make their final selection.

Overall, the important takeaways from this analysis are: a) that stimuli characterized by long duration are more likely to be selected as sounding more intensifying and more like a correction, regardless of the prosodic profile they are realized with and b) stimuli featuring lengthening only in the onset are more likely to be selected as sounding intensifying, and especially when accompanied by intensifying prosodic profiles, while stimuli featuring lengthening only in the nucleus were more likely to be selected as sounding more like a correction, regardless of their prosodic profile. These findings suggest that differences in duration are more variably associated with interpretations of semantic intensity, but that both meaning types are more compatible with increased duration more generally.
**Across-Pitch Comparisons.** To test for an effect of prosodic profile on the likelihood of an interpretation of contrast vs. semantic intensity and to probe the interaction between prosodic profile and duration, I fit a Bradley-Terry model (Bradley & Terry, 1952) to the across-pitch trials from each half of each experiment using the BradleyTerry2 package in R (Firth & Turner, 2012). A Bradley-Terry model predicts the outcome of a paired comparison by assuming that, in a contest between two players (in this case, one stimulus featuring an intensifying prosodic profile and one featuring a contrastive prosodic profile), the odds that one player beats another depends on a latent variable that may be thought of as ‘ability.’ This model allows us to understand which players in my experiment (operationalized as different combinations of prosodic profile and duration) are most likely to cue a given interpretation regardless of the player they are paired up against.

In these experiments, ‘ability’ can be considered a measure of representativeness of a given meaning type. Specifically, it indicates a) how associated a stimulus featuring an intensifying vs. contrastive prosodic profile is with a given meaning type and b) how much this association depends on the duration configuration of its competitor. For example, the ‘ability,’ score for ILL represents its likelihood of being chosen as sounding more intensifying than all duration configurations of contrastive prosodic profiles (CLL, CLS, CSL and CSS). The differences in ability for the stimuli featuring intensifying prosodic profiles therefore represent differences in how likely each intensifying duration configuration was to be selected as sounding like it conveyed a stronger degree of intensity compared to all the contrastive stimuli it was paired up against.

As discussed above, the experiments conducted here were between-subjects, with half of the participants hearing all durational variants of the intensifying prosodic profiles compared to
two levels of the contrastive prosodic profiles (CLL, CSS) and half of the participants hearing all
durational variants of the contrastive prosodic profiles compared to two levels of the intensifying
prosodic profiles (ILL, ISS). Because of these differences in comparison, separate Bradley-Terry
models were fit to the across-pitch comparisons for each half of each experiment. The model
applied to the contrastive half of the experiment (the one testing all durational variants of
contrastive stimuli) tests for the effect of duration on the likelihood of selection for contrastive
prosodic profiles. The model applied to the intensifying half of the experiment (the one testing
all durational variants of intensifying stimuli) tests for the effect of duration on the likelihood of
selection for intensifying prosodic profiles.

The output of the models provides maximum likelihood estimates and standard errors for
each stimulus’s log ability score. For both models, stimulus type (the ‘players’) was entered into
the model as a six-level factor (Contrastive Half: CLL, CLS, CSL, CSS, ILL, ISS; Intensifying
Half: ILL, ILS, ISL, ISS, CLL, CSS). For the contrastive half of the experiment, CSS was set to
the reference level, and for the intensifying half of the experiment, ISS was set to the reference
level. The models also included a contest-specific predictor for order of appearance (1 vs. 2),
which accounts for the fact that a specific position might have biased a player to “win” a given
matchup, where “winning” means being selected as the stimulus that best represents the meaning
(contrast or semantic intensity). The model did not include a random factor for participant,
meaning that it analyzes all response data without considering the participant who produced a
certain response.
Figure 15

Log Ability Scores for Interpretations of Semantic Intensity

Note. Log ability scores for each unique stimulus in Experiment 1a in both the contrastive (left) and intensity (right) halves of the experiment.

Beginning with Experiment 1a, as shown in Figure 15, the results from the Bradley-Terry model applied to the contrastive half of the experiment support the hypothesis that intensifying prosodic profiles are a stronger cue to interpretations of semantic intensity than contrastive ones (Hypothesis 1), as intensifying stimuli were more likely to be selected than contrastive ones, regardless of whether they were short or long. The results also support the hypothesis that increased duration serves as a cue to interpretations of semantic intensity (Hypothesis 2). For example, within the stimuli featuring contrastive prosodic profiles, stimuli with any amount of lengthening had higher odds of winning a given matchup compared to the shortest contrastive stimulus (CSS). Similarly, within the intensifying stimuli, the longer intensifying stimulus (ILL) was also more likely to be selected than the shorter intensifying stimulus (ISS) ($\beta=1.58$, $z=16.2$, ...
Regarding Hypothesis 3, the model results do not support the idea that asymmetric lengthening of the onset serves as a strong cue to interpretations of semantic intensity. For example, the CLL stimulus (which featured lengthening in the onset and nucleus) was significantly more likely to be selected than the stimulus featuring lengthening in only the onset ($\beta = 0.5, z = 5.05, p < .001$) and only in the nucleus ($\beta = 0.39, z = 3.58, p < .01$). Furthermore, the CSL and CLS stimuli did not have significantly different odds of being selected ($\beta = 0.17, z = 1.5, p = 0.13$).

Moving on to the intensifying half of the experiment, I see that the model results provide further support for the hypothesis that duration serves as a cue to interpretations of semantic intensity (Hypothesis 2). For example, once again, all three of the intensifying stimuli featuring some lengthening were significantly more likely to be selected than the shortest intensifying stimulus (ISS). Furthermore, the likelihood of selection increased from ISS to ISL to ILL/ILS, although ILL did not have significantly higher odds of being selected than ILS ($\beta = -0.06, z = -0.49, p = 0.6$). The fact that the ILS and ILL stimuli did not differ in their likelihood of selection supports Hypothesis 3 by suggesting that, at least for the stimuli featuring intensifying prosodic profiles, asymmetric lengthening of the onset was as strong of a cue to interpretations of semantic intensity as lengthening overall. Within the contrastive stimuli, the effect of duration was also present, as CLL had significantly higher odds of being selected than CSS ($\beta = 1.187, z = 13.75, p < .001$). Finally, once again, the intensifying prosodic profiles all had a higher likelihood of being selected than the contrastive prosodic profiles, providing further support for Hypothesis 1.
Considering the results of these two models in tandem, it appears that prosodic profile has a stronger effect on the interpretation of semantic intensity than duration (thus supporting Hypothesis 4), but smaller effects of duration within each prosodic profile can also boost an intensified meaning. A further finding is that for the intensifying stimuli only, an equal boosting effect of duration is observed for stimuli with a long onset (ILL, ILS), compared to stimuli with a long nucleus, suggesting that asymmetric lengthening is a particularly strong cue to semantic intensity, though only for intensifying prosodic profiles. For contrastive stimuli, the effect of duration appeared to be sensitive to the length of both the onset and nucleus, with additive effects that didn’t favor onset length over nucleus length.

Moving on to Experiment 1b, which tested interpretations of contrast (Figure 16), I find that the Bradley-Terry model applied to the contrastive half of the experiment illustrates a similar effect of duration on interpretations of contrast for contrastive stimuli as observed for interpretations of semantic intensity in Experiment 1a. For example, within the intensifying stimuli, the longer intensifying stimulus (ILL) was more likely to be selected than ISS ($\beta=0.79$, $z=11.68$, $p<.001$). Similarly, all of the contrastive stimuli featuring lengthening had significantly higher odds of being selected than the CSS stimulus, providing support for Hypothesis 2. The CSL and CLS stimuli were also not significantly different ($\beta=0.07$, $z=0.8$, $p=0.4$), in accordance with my hypothesis that asymmetric lengthening of the onset is only a productive cue to semantic intensity, and not for contrast (Hypothesis 3). Finally, the results do not support the hypothesis that contrastive prosodic profiles would encourage intensifying responses relative to intensifying ones (Hypothesis 1), as in this experiment, intensifying prosodic profiles sometimes had a higher likelihood of winning a given matchup than contrastive prosodic profiles. For
example, the ILL stimulus had a higher likelihood of being selected than even the longest contrastive stimulus (CLL). However, overall, differences in prosodic profile appeared to be a weaker cue to interpretations of contrast than they were to semantic intensity, as the intensifying stimuli and contrastive stimuli did not differ in the likelihoods of winning a given matchup to the same degree as they did for interpretations of semantic intensity. For example, the CLS stimulus was not significantly different from the ISS stimulus ($\beta=-0.07$, $z=-0.958$, $p=0.33$).

**Figure 16**

*Log Ability Scores for Interpretations of Contrast*

![Diagram](attachment:image.png)

*Note.* Log ability scores for each unique stimulus in the contrastive (left) and intensifying (right) halves of Experiment 1b.

Moving on to the model applied to the intensifying half of the experiment, I see a similar effect of duration for the intensifying stimuli. For example, all of the intensifying stimuli
featuring lengthening were more likely to be selected than the shortest intensifying stimulus (ISS). Similarly, the longest contrastive stimulus (CLL) was more likely to win a given matchup than the shorter contrastive stimulus (CSS) \( (\beta = 1.2, z=15.65, p<.001) \). The results for the intensifying stimuli provide support for the hypothesis that asymmetric lengthening of the onset is only a productive cue to semantic intensity, and not to contrast, by illustrating that, when participants were asked about contrast, the ILS and ISL stimuli did not have significantly different odds of being selected \( (\beta=0.02, z=0.26, p=0.79) \). Finally, once again, in contrast to Hypothesis 1, the intensifying stimuli were overall more likely to be chosen than the contrastive stimuli, although just as in the contrastive half of the experiment, the shortest intensifying stimulus (ISS) did not have significantly higher odds of being selected than the longer contrastive stimulus \( (\beta=-0.02, z=-0.382, p=0.7) \).

Taking the results of these two models together, I see that, in accordance with Hypothesis 4, regarding the superiority of prosodic profile as a cue to interpretations of contrast and semantic intensity, prosodic profile had a larger effect on interpretations of contrast than duration. However, once again, duration can further boost a contrastive meaning within a given prosodic profile, thus supporting Hypothesis 2. Of particular note from the outcome of Experiment 1b is that contrastive prosodic profiles did not encourage contrastive interpretations relative to intensifying ones. In fact, the opposite pattern was true, and intensifying prosodic profiles appeared to cue interpretations of contrast in a similar manner as they did for interpretations of semantic intensity.

Considering the output of these four models, I can now qualitatively compare the results from Experiment 1a (regarding semantic intensity) and Experiment 1b (regarding contrast) to
further examine how these findings provide support for Hypotheses 3 and 4. First, regarding the hypothesis about asymmetric lengthening in either the nucleus or onset serving as a strong cue to interpretations of semantic intensity (Hypothesis 3), it is interesting to note that the distinction between lengthening in the onset and lengthening in the nucleus that was evident for interpretations of semantic intensity in Experiment 1a was not apparent in the contrastive/corrective context instantiated in Experiment 1b. For example, both CLS and CSL and ILS and ISL did not have significantly different chances of winning a matchup against an intensifying stimulus in Experiment 1b, unlike in Experiment 1a, when ILS had significantly higher odds of winning a given matchup than ISL. This finding suggests that listeners show greater sensitivity to differences in where in the syllable lengthening occurs when interpreting prominence as a cue to semantic intensity, although only when an adjective is produced with an intensifying prosodic profile. In these cases specifically, participants appear to hear lengthening in the onset as a stronger cue to semantic intensity than lengthening in the nucleus.

Regarding Hypothesis 4, about the superiority of prosodic profile as a cue to semantic intensity and contrast as opposed to duration, the results indicate that, across both experiments, the stimuli featuring intensifying prosodic profiles were stronger cues to both meaning types compared to stimuli with contrastive prosodic profiles. If duration, on the other hand, were more important, I would have expected to see variation in the likelihood of selection based on duration, without a clear separation between the stimuli featuring different prosodic profiles.

Discussion

Overall, the results from this experiment provide robust support for two of the hypotheses introduced earlier, and partial support for a third. For example, regarding the ability of
intensifying prosodic profiles to cue interpretations of semantic intensity and contrastive prosodic profiles to cue interpretations of contrast (Hypothesis 1), in the across-pitch analysis, intensifying prosodic profiles did seem to be the preferred instantiation of prominence conveying semantic intensity, as listeners were significantly more likely to select a stimulus featuring an intensifying prosodic profile as sounding more intensifying compared to a contrastive one. However, contrary to my original hypothesis, interpretations of contrast were also cued by contrastive prosodic profiles, although the stimuli featuring contrastive profiles did have overall higher odds of being selected as sounding more like a correction in Experiment 1b (when participants were asked about contrast) than in Experiment 1a, suggesting that the two interpretations do differ slightly in terms of their ability to be cued by contrastive vs. intensifying prosodic profiles.

Regarding Hypothesis 2, which suggested that interpretations of contrast and semantic intensity would both be cued by increased duration, the results from the within-pitch analysis suggest that, indeed, for each prosodic profile, LL stimuli were significantly more likely to be selected as sounding more intensifying or more contrastive than SS stimuli, regardless of the prosodic profile of the stimulus. The results from the across-pitch analysis similarly show that increasing the duration of the contrastive or intensifying adjectives increased their likelihood of being selected as contrastive or intensified when the competing stimulus featured a different prosodic profile, providing further support for this hypothesis.

Moving on to Hypotheses 3 and 4, both of which address the interaction between prosodic profile and duration, I observe that the results from both the within- and across-pitch analyses provide support for the idea that a) prosodic profile serves as a stronger cue to
interpretations of intensity and contrast than duration and b) asymmetric lengthening of the onset is a strong cue to interpretations of semantic intensity, but not contrast. For example, in the across-pitch analysis, in Experiment 1a, lengthening in the onset was more likely to be selected as sounding more intensifying than lengthening in the nucleus. However, in Experiment 1b, there were no significant differences between the ILS and ISL stimuli, suggesting that where in the syllable the lengthening occurs has little effect on the compatibility of a given prominence realization for interpretations of contrast. Furthermore, in the within-pitch analysis, there was a significant difference between the LS and SL stimuli only in Experiment 1a and only when the stimulus was produced with an Intensifying prosodic profile. These findings are in line with Niebuhr’s (2010) suggestion that onset vs. nucleus lengthening plays an important role in conveying positive vs. negative stances in the realm of intensification, and it suggests that this contribution is not as important when conveying contrast, which, intuitively, might seem less intertwined with attitude from the start.

Finally, as suggested by Hypothesis 4, the effect of duration on the likelihood of selection was not as great as the effect of prosodic profile, as duration effects on the likelihood of selection were primarily observed within-pitch as opposed to across. For example, in the across-pitch analysis, the largest separation between stimuli in terms of their likelihood of being selected was between the stimuli featuring intensifying and contrastive prosodic profiles, as each intensifying stimulus was more likely to be selected than its opposing contrastive stimulus. In other words, the variation seemed to be primarily within-pitch rather than across-pitch. Both the within-pitch and across-pitch analyses suggest a significant effect of order of appearance on the likelihood that a given stimulus won a given matchup, with the likelihood increasing when that stimulus
appeared on the right rather than the left. As previously suggested, this result is due to a slight preference of participants to select the stimulus on the right, although this bias was not so great as to negate the effects elaborated on here.

Overall, the results from this experiment suggest that, contrary to my initial hypothesis, the stimuli characterized by intensifying prosodic profiles were able to cue interpretations of both semantic intensity and contrast. One possible explanation for this result is that the interpretation of contrast/semantic intensity is mediated by a word’s degree of prominence. For example, as mentioned in the introduction, pitch accents can be ranked relative to one another in terms of their prominence (Cole et al., 2010; Cole et al., 2019; Im et al., 2023; Bishop et al., 2020), and there can be additional variation within pitch accent category in the degree of perceived prominence depending on the acoustic realization of the accented syllable or word (e.g., via gradient changes in pitch range, duration or intensity) (e.g., Ladd & Morton, 1997). It has also been suggested that stronger (i.e., more informative) types of meaning map onto higher levels of prominence (e.g., H* vs. L+H* for new vs. contrastive focus) (e.g., Fery, 2017). Therefore, it is possible that the effects shown here are due to differences in prominence between the stimuli based on the acoustic differences between the two prosodic profiles, with the lowered pitch profiles and slower speech rate characteristic of the intensifying prosodic profiles being stronger cues to prominence (or cues to a stronger degree of prominence), and thus more likely for any marked interpretation, e.g., semantic intensity or contrast.

However, this hypothesis contradicts previous theories of prominence, which argue that the most prominent pitch movement is the high-rising one typically associated with contrastive focus, due to its higher pitch maximum (Gussenhoven, 2002; Pierrehumbert & Hirschberg,
Yet, other proposals suggest that perceived prominence may have less to do with a simple distinction between high vs. low pitch and may instead be derived based on unexpectedness/markedness. For example, Kugler and Genzel suggest that prominence is characterized by ‘a deviation from an unmarked prosodic structure,’ and do not discriminate between the directionality of this deviation (2011:1). Therefore, it is possible that the stimuli featuring intensifying prosodic profiles in this study were perceived as overwhelmingly more prominent than the stimuli featuring contrastive ones, perhaps because of the markedness of the pitch profile or because of other differences in the stimuli. Because of these conflicting results, I conducted a follow-up study asking participants to make metalinguistic judgments about prominence, in order to explore the degree to which prominence might serve as a mediating factor for the interpretation of contrast and semantic intensity.

Experiment 2

Introduction and Hypotheses

The perception of prosodic prominence is a complex phenomenon, and it has been shown to be influenced by both signal-based factors in addition to expectation-driven ones (e.g., Cole et al., 2010). For example, Cole et al., (2019) show that aspects of the acoustic signal such as a slower speech rate (as measured in longer duration) can increase the likelihood that an element is perceived as prominent, in addition to elevated F0 or higher intensity. From an expectation standpoint, a word is more likely to be produced with acoustic prominence (larger F0 excursions resulting in enhanced F0 extrema; longer duration; greater acoustic energy; among other properties) depending on its information structure properties (lexical and referential givenness, focus), its surprisal value (measured in terms of uni- or n-gram frequency in a large speech
corpus), and its status as the phrase-final (nuclear) vs. non-final accented word in its prosodic phrasal domain. Accordingly, listeners expect words in prominence-lending contexts to be more prominent, and these expectations result in a contextual prominence bias that is reflected in explicit prominence judgments (Cole et al. 2010, 2019).

In Experiment 1a and b, I tested the ability of a given acoustic manifestation of prominence to cue an interpretation of contrast vs. semantic intensity. The results suggest that intensifying prosodic profiles are more associated with interpretations of both contrast and intensity, and I hypothesize that this might be due to a higher degree of perceived prominence for this specific manifestation of pitch relative to the stimuli featuring contrastive prosodic profiles. Therefore, in this experiment, I test whether the pattern of results seen in Experiments 1a and b can be explained by metalinguistic judgments about prominence, hypothesizing that both interpretations are encouraged by pitch and duration patterns that confer greater prominence. Considering the results from Experiments 1a and b, in the context of this experiment, I predict that stimuli characterized by intensifying prosodic profiles will be more likely to sound prominent than stimuli characterized by contrastive ones (ILL, ILS, ISL, ISS > CLL, CLS, CSL, CSS). Furthermore, stimuli characterized by longer duration will be more likely to sound prominent than stimuli characterized by shorter duration.

Procedure and Design

The procedure and design were the same as in Experiments 1a and 1b, except that participants were now asked to select the sound which made the critical adjective stand out more from the other words around it. In other words, this experiment was a metalinguistic task about perceptual salience/prominence.
Participants

A total of 90 participants were recruited from Prolific for this study. The participants were all between the ages of 18-65 and self-reported as having normal-to-corrected hearing and vision and no reading, hearing or learning impairments. All participants were fluent speakers of American English who had been born in the United States. Due to issues with data submission, five participants were excluded from the contrastive half of the experiment (n = 40) and nine from the intensifying half (n = 36) resulting in a total of 76 participants.

Analysis and Results

Since the goal of this experiment was to test whether differences in the likelihood of choosing a stimulus featuring an intensifying vs. contrastive prosodic profile in Experiment 1a and 1b was due to differences in prominence between the two types of prosodic profiles, I restrict my analysis of this experiment to the across-pitch trials. Once again, I apply two separate Bradley-Terry models to each half of each experiment. The players were entered into the model as a six-level factor (Contrastive Half: CLL, CLS, CSL, CSS, ILL, ISS; Intensifying Half: ILL, ILS, ISL, ISS, CLL, CSS) with the reference level set to CSS (contrastive half) and ISS (intensifying half).
Log ability scores for all eight stimuli in terms of their likelihood of winning a given matchup when participants were asked about relative prominence.

Regarding the model applied to the contrastive half of the experiment, the effect of duration on perceived prominence is in the hypothesized direction. All of the contrastive stimuli featuring any amount of lengthening had a higher chance of being selected than the contrastive stimulus featuring no lengthening (CSS). The effect of duration on prominence was strongest for the CLL and CSL stimuli, both of which had significantly higher chances of being selected than the stimulus featuring lengthening in the onset only (CLL: $\beta = 0.95$, $z=8.59$, $p<.001$; CSL: $\beta = 0.58$, $z=5.296$, $p<.001$). For the intensifying stimuli, the ILL stimulus also had a higher chance of being selected than the ISS stimulus ($\beta=2.47$, $z=28.3$, $p<.001$). Thus, increased duration in both the onset or nucleus, or in the nucleus only for the contrastive stimuli, appeared to result in greater perceived prominence. The effect of prosodic profile on prominence in these experiments is less clear. For example, unlike when participants were asked about intensifying and contrastive
interpretations, the intensifying and prosodic profiles were not clearly separable in terms of their likelihood of being selected as sounding prominent. While each intensifying stimulus had a higher likelihood of being selected as sounding more prominent than the contrastive stimulus of the same duration, the ISS stimulus, for example, had significantly lower odds of being selected than all of the contrastive stimuli featuring any amount of lengthening, suggesting that in this experiment, duration had larger consequences for prominence than prosodic profile.

Moving on to the model applied to the intensifying half of the experiment, I see that there was once again an effect of duration on perceived prominence. Within the stimuli featuring intensifying prosodic profiles, all of the longer stimuli had a higher likelihood of winning a given matchup than the shortest stimulus (ISS). Furthermore, similar to the results for the contrastive stimuli in the contrastive half of the experiment, perceived prominence was highest for lengthening in both the onset and nucleus and lengthening in the nucleus only, as both ILL and ISL had a higher likelihood of winning a given matchup than ILS (ILL: β=0.58, z=4.585, p<.001; ISL: β=0.26, z=2.12, p<.05). Within the contrastive prosodic profiles, CLL was also significantly more likely to win a given matchup than CSS (β=2.04, z=26.37, p<.001). Finally, just as in the contrastive half of the experiment, there was once again not a large effect of prosodic profile on prominence. For example, while each intensifying stimulus still had a higher likelihood of being selected than the contrastive stimulus with the same duration, the shortest intensifying ISS stimulus had a lower likelihood of selection than the longest contrastive CLL stimulus (β = -1.31, z=-13.03, p<.001). This pattern of results suggests that duration had greater consequences for perceptions of prominence than it did for interpretations of semantic intensity and contrast, when
all of the intensifying stimuli had a higher likelihood of being selected than any of the contrastive stimuli.

Considering the results of these two models in tandem, it appears that, although intensifying prosodic profiles were perceived as slightly more prominent than contrastive ones, their effects on prominence were somewhat dependent on their duration. For example, when an intensifying stimulus was short, a long contrastive stimulus could be perceived as more prominent. The effects of duration on prominence as shown in this experiment also differed from the effects of duration on interpretations of semantic intensity and contrast shown in Experiments 1a-b. For example, in the experiment discussed here, for both contrastive and intensifying prosodic profiles, lengthening in the nucleus increased the likelihood of selection compared to lengthening in the onset, while in the experiments addressing semantic intensity and contrast, the two stimuli were either not significantly different, or favored a relation whereby the stimulus with a long onset had a higher likelihood of selection than the one with a long nucleus.

Discussion

If the reason that intensifying prosodic profiles were more compatible with both contrastive and intensifying interpretations (as shown in Experiments 1a and b) was because intensifying profiles sounded generally more prominent, I would expect that, when asked to make explicit judgments about prominence, the stimuli featuring intensifying prosodic profiles would be more likely to win a given matchup than the stimuli featuring contrastive ones. However, the results from this experiment are mixed. On the one hand, the stimuli characterized by intensifying profiles did appear to have a higher likelihood of sounding prominent than their contrastive counterparts, and indeed, the ILL stimulus had the highest likelihood of being
selected as sounding prominent and the CSS stimulus the lowest. However, there was much less separation between the contrastive and intensifying stimuli in general than in the previous experiments, suggesting that, when the pragmatic context is left unspecified, differences in prosodic profile are not clearly mapped onto distinctions in prominence, at least when pitted directly against one another. This finding contradicts previous theories about prominence, which would suggest that the contrastive stimuli would be perceived as more prominent, given their higher pitch profile. Instead, it supports the suggestion that prominence may instead be derived based on markedness (Kügler & Genzel, 2011). Or possibly, it suggests that it cannot be easily mapped onto acoustic cues in the absence of pragmatic context.

Two other possibilities might explain the results here. First, as mentioned in the description of the materials, the stimuli in this experiment differed in a variety of ways, and not just in the scaling of the pitch movement realized on the critical adjective. For example, the intensifying stimuli featured lower pitch maxima but also an overall slower speech rate and less modal voice quality than contrast, in addition to a preamble that featured a higher mean F0. Therefore, it is possible that these additional factors may have contributed to the intensified stimuli’s increased prominence, despite what might be considered a less prominent pitch movement. It is also possible that, because there was not specific pragmatic context instantiated in these experiments, listeners created a context themselves and interpreted prominence as a cue to pragmatic meaning even though they weren’t explicitly forced to do so. If different participants interpreted prominence according to different ‘imagined’ pragmatic contexts, this would explain the more muddied results seen here.
Overall Discussion

In the experiments discussed in this chapter, I investigate the effect of duration on the ability of prosodic profiles originally intended to convey contrast and semantic intensity to convey their respective meanings. Overall, the results from this study support the hypothesis that longer duration is a cue to both contrastive and intensifying interpretations, but that interpretations of semantic intensity show more variation in the effects of duration than contrastive ones do. For example, duration appeared to have a greater influence on log ability scores in Experiment 1a (when participants were asked about semantic intensity) than in Experiment 1b (when participants were asked about contrast), a fact which could be linked to the connection between duration and semantic intensity in production studies in German (Niebuhr, 2010). Furthermore, listeners exhibit sensitivity to which part of the syllable is lengthened when interpreting prominence as a cue to semantic intensity, while they do not show the same sensitivity for contrast. This linkage of duration and semantic intensity may have various sources, including an iconic meaning connection or a possible connection to attitudinal meaning (e.g., Schlenker, 2018 for iconicity; Ladd, 2008 for attitudinal meaning). In this regard, consider Niebuhr’s (2010) suggestion that lengthening in the onset conveys positive attitudes in the context of semantic intensity (i.e., the speaker feels the intensified quality is overall positive), while lengthening in the nucleus conveys a negative attitude. The results from Experiments 1a and 1b suggest that attitudinal meaning may indeed be playing a larger role in the interpretation of cues to semantic intensity as opposed to contrast, as differences between lengthening in the onset vs. nucleus were larger for a) the intensifying the stimuli featuring intensifying prosodic profiles and b) when listeners were asked to interpret prominence as a cue to semantic intensity. These results suggest that listeners are sensitive to the acoustic realization of prominence in ways that differ depending
on whether they are interpreting it for contrast vs. semantic intensity, as they appear to attend to differences in prosodic profile and duration in ways that are specific to each meaning category.

Despite the influence of duration on the likelihood a given stimulus was selected, in all of the experiments conducted here, prosodic profile still served as a stronger cue to interpretations of contrast and semantic intensity than duration. The largest separation between stimuli in terms of their odds of winning a given matchup across all experiments was based on prosodic profile rather than duration. For example, in Experiment 1a, while both contrastive and intensifying stimuli showed distinctions in their log ability scores based on duration, all intensifying stimuli had higher log ability scores than all contrastive stimuli. However, the prosodic profile that was interpreted as being more compatible with a contrastive meaning was not consistent with my hypotheses. For example, it is typically argued that contrast is associated with sharp, high, peaked pitch movements like L+H*, and that the increased prominence of this pitch accent relative to L* or H* accent is what primarily facilitates this interpretation (Pierrehumbert & Hirschberg, 1990; Cole et al., 2010). Yet, in the experiments conducted here, intensifying prosodic profiles, which featured lower pitch, were more consistent with a contrastive interpretation than contrastive profiles, which featured higher pitch.

In Experiment 2, I investigated whether this finding was due to general differences in prominence between the two prosodic profiles, and I found that listeners did not have strong intuitions about the difference in prominence between the contrastive and intensifying stimuli, although intensifying profiles were perceived as slightly more prominent than contrastive ones. This finding has important implications for theories of prominence, as it supports recent suggestions that prominence may be more closely related to markedness, or deviation from an
established pitch norm, rather than a particular direction of pitch change (Kugler & Genzel, 2011). For example, one might argue that, because of the more limited variation speakers have in the lower parts of their pitch range, producing a pitch accent that is scaled lower (as is the case for the intensifying stimuli used in this experiment) can be riskier in terms of being misinterpreted, causing it to be perceived as a more marked choice. Furthermore, lower pitch often co-occurs with non-modal changes in voice quality, such as creaky voice. Thus, making the decision to use lower pitch runs the risk of masking the pitch contour entirely, inhibiting the ability of F0 to convey important distinctions in information structure. Since the intensifying stimuli were more likely to be selected across all three experiments, it does not seem to be the case that lower pitch was only perceived as a more marked choice in a particular pragmatic context. Otherwise, I would expect to see flipped responses in Experiments 1a and b, where participants heard one stimulus as more prominent when interpreting it as semantic intensity and another when interpreting it for contrast.

In all the experiments conducted here, the models suggest a significant effect of order of appearance on the likelihood a given stimulus was selected, with stimuli being more likely to be selected when they appeared on the right than when they appeared on the left. It is possible that this effect is because prominence, and perhaps interpretation as well, is perceived in a relative way (i.e., in relation to the elements around it). In this experiment, I presented participants with utterances containing adjectives that were produced as prominent within a given intonational context/sentence. Yet, in the actual experimental task, participants had to engage with prominence at a higher level: They had to make a relative comparison of prominence or compatibility with a given interpretation across two sentences. This is a difficult task, and not
one that has been used in prior studies. Therefore, it is possible that, in the experiments conducted here, the preference for the stimulus on the right was because, assuming that listeners played the stimuli from left to right, it was easier to determine relative prominence or make a judgment about a stimulus’s compatibility with a given interpretation after hearing an initial baseline in the first audio file. Alternatively, it suggests that participants were simply less likely to replay the two audio files in order to re-compare the stimulus they had initially played with the one appearing on the right, and that they preferred to leave their cursor on the right after playing both stimuli.

One interesting question regarding the results from this experiment concerns the level of prosodic structure at which semantic intensity is conveyed. Throughout this study, I assume that the distinction between contrastive and intensifying interpretations tested here is based on differences in prominence on the critical adjective. However, I note that, in the productions utilized as stimuli in both Experiments 1a-b and 2, the stimuli produced with the intent to convey semantic intensity featured a lower pitch maximum on the critical adjective but also an increase in the max pitch of the prenuclear material, in addition to an overall slower speech rate. Therefore, it is possible that, rather than targeting an intensifying accent in this study, I more-so targeted an intensifying holistic contour across the phrase. For example, speech acts like exclamatives are sometimes associated with “constructional prosodic defaults” that include in their description of the prosodic realization of exclamatives a prominent pitch accent in addition to information about the shape of the overall prosodic contour (falling) and the speech rate (which is slower) (Repp & Seeliger, 2020). It is therefore an interesting question whether semantic intensity is conveyed at the level of the accent, as appears to be the case for contrast, or
whether it is associated with a prosodic default that spans the entire utterance, thereby operating more similarly to speech act meaning as opposed to information structure or focus.

One potential avenue for further study concerns the contribution of various other acoustic cues to perceptions of contrast vs. semantic intensity. In the experiments discussed here, I isolated the effect of duration in the onset vs. nucleus from an unmanipulated, prosodic profile for contrast and semantic intensity. These contexts involved a variety of differences, including differences in the scaling of the pitch movement on the adjective, the phonological pitch accent realized on the adjective, the pitch scaling of the adjective relative to the preamble, voice quality, and overall speech rate. Any one of these variables could be isolated in future studies to investigate their effects on interpretations of contrast vs. semantic intensity, as they most likely all have a role to play in guiding interpretation. Furthermore, future studies could investigate the shape of the pitch accent realized on the critical adjective, asking whether it featured, for example, an L+H* accent or an H* accent.

**Conclusion**

This study aimed to investigate the association of different acoustic manifestations of prominence with pragmatic meanings of contrast vs. intensity. In Experiments 1a and 1b, I examined participants’ rankings of stimuli that varied in prosodic profile (contrastive vs. intensifying) and duration (long vs. short) as conveying contrast or semantic intensity. The results suggest that a) associate increases in duration with both meaning types, b) associate intensifying prosodic profiles (characterized by lower pitch) with both meaning types, c) associate asymmetric increases in the duration of the onset as opposed to the nucleus with interpretations of semantic intensity but not for contrast, and d) place more weight on prosodic
profile as a cue to either interpretation than they do to duration. In Experiment 2, I confirm that these effects are not driven by overall differences in prominence between the two prosodic profiles and their duration manipulations, as the stimuli that were rated as more prominent in this experiment were not always the stimuli that were the most likely to convey contrast vs. semantic intensity, respectively. These findings have important implications for theories of prominence and meaning, as it suggests that unique realizations of prominence can variably map onto different meaning categories, suggesting a complicated and multifaceted nature to the form-function mapping.
Chapter 3: The Relative Contribution of Onset and Nucleus Duration to Interpretations of Expressive Intensity

Introduction

Previous research suggests that prosodic prominence can convey semantic intensity, increasing the threshold on a measurement scale of degrees above which a given adjective qualifies as being true of an object (see Kennedy & McNally, 2005 for a description for lexical intensifiers; see Bolinger, 2013; Niebuhr, 2010 for suggestions that prosody can convey a similar meaning). However, lexical intensification has been described as a pragmatic and expressive phenomenon in addition to a semantic one, as it can modify metalinguistic scales and convey expressive content (e.g., Beltrama & Bochnak, 2015; Beltrama, 2018; Beltrama & Trotzke, 2018). Similarly, prosodic prominence has been shown to convey paralinguistic meaning related to emotions and attitudes, and especially so when simultaneously conveying semantic intensity (Bolinger, 1986; Ladd, 2008; Niebuhr, 2010). Existing investigations into the acoustic realization of prominence conveying intensification do not clearly distinguish between the acoustic realizations most typical of one type of intensity vs. another, and instead tend to target the realization of prominence as a cue to intensity more generally, without teasing apart the roles of different cues in targeting one dimension vs. another.

In this chapter, I investigate the effect of changes in the duration of the onset/nucleus of an adjective, in addition to changes in pitch scaling, on interpretations of strong opinions. In this sense, I target intensity related to the expressive dimension, probing the specific realization of intensity targeting attitudinal meaning as opposed to semantic meaning related to a measurement scale denoted by a gradable adjective. I compare these results to those from Chapter 2, when I
targeted interpretations of semantic intensity by asking participants to compare stimuli based on whether they conveyed a high degree value on the scale in question, denoted by a pitch-accented adjective (e.g., stripedness, cloudiness, etc.). The results from the present study show that the effect of duration and pitch scaling on interpretations of intensity differs for intensity targeting semantic vs. expressive scales, although the two types of meanings do show considerable overlap in their preferred realizations. The findings contribute to the literature by providing a clearer picture of the specific realization of prominence conveying expressive intensity as opposed to semantic intensity. Furthermore, they provide initial insights into the mapping of variation in the acoustic dimension with variation in the dimension of intensity, just as previous investigations do for the mapping between prominence and pragmatic meaning categories like focus and givenness.

**Semantic/Expressive Intensity**

The term ‘intensity’ is typically applied to any type of meaning computation that involves degrees on some sort of scale. The scale can be semantic, pragmatic, or expressive. Below, I discuss each of these scales in turn. However, I note that, to date, only intensity targeting pragmatic and expressive scales has been modeled in terms of its relation to prosody. Therefore, I begin by introducing formal models developed for lexical intensifiers before moving on to discuss the possible connection between prosody and all the scale types discussed below.

**Semantic intensity** involves measurement scales of ordered degree values that are introduced in the denotation of gradable predicates. For example, gradable adjectives like tall “map their arguments onto abstract representations of measurement, or degrees, which are formalized as points or intervals partially ordered along some dimension” (Kennedy & McNally,
2005: 349). According to the framework of Threshold Semantics (Kennedy, 2007), there is a degree value for each scalar adjective, known as the threshold (θ), above which a given entity qualifies as being true of that adjective. In other words, a gradable adjective is true of an object if it manifests the relevant property to a degree at least as great as the threshold. These thresholds are not part of the denotation. Rather, they are fixed by constructions like comparatives and measure phrases and adjusted by intensifiers. For example, researchers have proposed that adjective phrases modified by intensifiers have the same semantics as unmodified adjectives, but with new and increased thresholds (Bennett & Goodman, 2018). Specifically, intensified thresholds are determined by collecting the set of objects in the comparison class for which the bare adjective is true (for example, all objects that are above the threshold for tall), and then using that as the comparison class to infer a new threshold (Bennett & Goodman, 2018).

Calculating a threshold for a very tall tree, then, would first involve determining what counts as a tall tree and then determining a new threshold of 'tallness' for this class, so that very tall now comes to denote something akin to tall for a tall tree (Bennett & Goodman, 2018: 149).

Scalar adjectives that are not modified by degree morphology (e.g., tall as opposed to very tall) introduce ambiguity about the threshold’s location. For example, the degree of height that counts as tall varies greatly depending on what is being talked about. To capture this type of behavior, semanticists propose that bare adjectives have thresholds that are inferred from the context, formalizing this type of contextual restriction with a phonologically null ‘positive’ morpheme pos (Kennedy & McNally, 2005). For example, one aspect of the context that can influence the thresholds of these bare adjectives is the degree structure of the scale itself.

Kennedy and McNally (2005) propose that whether a given adjective introduces a scale that is relative, with thresholds that are highly variable and context-dependent (big, small) or absolute,
with default minimum or maximum thresholds (*full, straight*), influences the type of inferences listeners draw about the likely value of the threshold. For example, absolute adjectives with minimum thresholds like *open* or *bent* characterize objects as having a non-zero degree of the relevant property, while maximum adjectives like *straight* or *closed* characterize objects has having a maximal degree of the relevant property. As a result, the thresholds for these types of absolute adjectives are much less variable than those for relative adjectives like *tall*.

Furthermore, the type of scale structure influences the felicity of a gradable adjective with specific types of intensifiers. For example, while absolute adjectives are felicitous with intensifiers like *completely* or *entirely* (e.g., *completely full*), relative adjectives are not (e.g., *completely tall*).

Beltrama and Bochnak (2015) argue that some intensifiers can also function as expressives, conveying what I refer to in this chapter as EXPRESSIVE INTENSITY. Expressives are defined by Potts (2007) as expressions whose main function is to display some kind of evaluative attitude or emotion, mostly on the part of the speaker. Their meaning is independent of the descriptive content, and the attitude is mostly speaker-oriented. For example, Beltrama and Bochnak argue that the Italian suffix *-issimo* conveys that “the speaker is in a state of maximal, emotional involvement about the fact that an individual x has property P…While a particular individual may not be the tallest in the context, taller individuals would not trigger any higher emotional involvement” (2015: 870). Furthermore, modifiers like *fucking* can boost the intensity of emotional involvement of the speaker, targeting a speaker-oriented scale and conveying a high degree of frustration/emotive involvement (Potts, 2005) on the part of the speaker. Irwin (2014) also suggests that ‘drama so’ (e.g., *I am so taking a nap later today*) can express emotive, rather than knowledge-based commitment of the speaker towards the proposition.
These two different types of intensity are not entirely independent. For example, Gutzmann (2011) and McCready (2010) discuss the notion of mixed expressives, which are expressions that contribute both expressive and descriptive content. In their analysis of expressive *much* (e.g., *Rude, much*?), they argue that *X-much* targets a gradable predicate while also adding a speaker’s evaluative attitude about the degree to which an individual stands out on the relevant scale (namely, that the degree is exceptionally high). Similarly, Beltrama and Bochnak (2015) argue that the Italian suffix *-issimo* contributes expressive content related to the speaker’s emotional response, but also descriptive content related to precision. For example, a tree that is *altissimo* is a tree that counts as tall in every context, but also one that provokes an emotional response.

Thus, overall, based on previous models of lexical intensity, intensity as conveyed by lexical intensifiers can be a semantic and expressive phenomenon. Considering observations that prosody can also convey intensity, it is possible that prosody is capable of targeting both these dimensions, and that the connection between certain realizations of prominence and intensification is distinct depending on whether an individual is conveying semantic intensity, expressive intensity, or both. In the following section, I provide an overview of prior research on the realization of prominence conveying semantic intensity and expressive intensity, before moving on to describe the experimental design and questions targeted in this chapter.

**Prosody and Intensity**

As already discussed in Chapter 2, experimental investigations into the specific realization of prominence conveying semantic intensity remain limited to German, and they tend to elicit this meaning category in an attitudinally charged context. For example, in Niebuhr’s (2010) production study, he elicits semantic intensity in both positive and negative contexts by asking...
pairs of participants to read aloud short dialogues and monologues. The dialogues and monologues were constructed so as to trigger the production of positive and negative intensification on a particular word. In both contexts, the words used always matched semantically with the positive vs. negative context. Niebuhr (2010) concludes from this study that asymmetric lengthening of the onset vs. nucleus can be mapped onto a distinction in attitudinal valence, with the latter being consistent with negative intensity and the former with positive. Furthermore, he investigates several acoustic properties that might be characteristic of each type. For example, his analysis suggests that positive intensity is characterized by quickly articulated onset consonants and lengthened vowel nuclei, extended high F0 plateaus featuring steep rises and shallow falls, and sonorous, slightly breathy voice. Negative intensity, on the other hand, was characterized by lengthening of the sound preceding the accented syllable and the onset consonant, along with quick and abrupt articulation of the vowel nucleus, a pointed F0 peak that fell sharply right after the vowel’s onset and pressed and irregular voice quality. Niebuhr’s (2010) results provide a foundation on which to base hypotheses about the realization of semantic intensity, but because he elicits semantic intensity in combination with attitudinal meaning, his findings make it difficult to isolate the contribution of any acoustic dimension to one dimension of intensity over another.

On the expressive side, although not specifically described as expressive intensity, it is well-established that prosody conveys emotions and attitudes. For example, Ladd writes that prosodic prominence “conveys some sort of special emotional involvement on the part of the speaker” (2008: 129). Bolinger argues that this affective side of prosody cannot be separated from propositional meaning. For example, he writes that, “though intonation is indispensable to
grammar, the grammatical functions of intonation are secondary to the emotional ones; speakers FEEL differently about what they say, and the feelings manifest themselves in pitch changes that serve as cues” (1986: 27).

The expressive contribution of prominence can be both emotional and attitudinal. These two terms are sometimes conflated in the literature, but they are most often distinguished by the fact that one characterizes a speaker’s state (emotion) while the other characterizes a stance that motivates a speaker’s actions (attitude) (Couper-Kuhlen, 1986; Wichmann, 2000; Wickens & Perry, 2015). For example, Arndt and Janney (1991) suggest a distinction between emotional communication and emotive communication, which differ according to the speaker’s intent. While emotional communication is spontaneous, unplanned, and instinctive, emotive communication is defined as “the intentional expression of feelings and attitudes via culturally learned affective displays” (Arndt & Janney, 1991: 531). Scherer similarly describes emotions as being expressed in response to highly significant events and expressed/identified in ways that are largely universal (Scherer et al., 2001, Scherer, 2003), while attitudes represent longer-term beliefs and predispositions that are less intense and more socially and culturally dependent (Scherer, 2003).

Attitudes and emotions are generally assumed to be conveyed at different levels of prosodic structure. For example, emotions tend to be expressed via acoustic changes at the phrasal level (such as via voice quality or intonational contour). In fact, Niebuhr (2010) argues that emotions cause longer-term changes in physiology that influence the phonetic signal at the level of F0. Attitudes, on the other hand, are often conveyed at the level of accents or feet (Niebuhr, 2010; Wickens & Perry, 2015). Perception studies support the idea that attitudes and emotions can be distinguished based on differences in the acoustic signal. For example,
Grichkovtsova, Morel, and Lacheret (2012) illustrate that voice quality is important for the perception and identification of emotions in speech, while prosodic contour is more important for the perception of attitudes.

Many research studies have aimed to identify a standard mapping between aspects of the prosodic signal such as single acoustic parameters (Lieberman & Michaels, 1962), overall-intonation contours (Uldall, 1964), or tone groups/nuclear tones (O’Connor & Arnold, 1973) and emotional/attitudinal meaning. For example, previous studies illustrate that pitch is the most important cue to speaker attitude (Blanc & Dominey, 2003; Bulut & Narayanan, 2008), although amplitude, loudness and pauses may also play a role (Gelinas-Chebat & Chebat, 1992). However, the mapping between the prosodic signal and emotional/affective meaning is not one-to-one. Rather, it seems more likely that paralinguistic meaning is inferred based on a combination of prosody, speech act, event, participants, knowledge, etc. (Wichmann, 2002). For example, both Halliday (1994) and Ladd (2008) argue that attitudinal meaning results from the combination of a specific pitch contour with a particular sentence type.

Overall, these prior studies suggest that prosody is capable of conveying intensity targeting semantic, pragmatic, and expressive scales, although its realization may differ slightly for each type, or each type may be conveyed at different levels of prosodic structure. In Chapter 2 of this dissertation, I examined the effect of duration and PROSODIC PROFILE (a term I use to refer to the pitch and duration characteristics evident in natural productions of adjectives intended to convey contrast vs. semantic intensity) on interpretations of semantic intensity and contrast, asking participants to choose which of two adjectives differing in prosodic profile and in the duration of their onsets vs. nuclei (long vs. short) more strongly conveyed the intended
meaning, specified as contrast or semantic intensity in a between-subjects design. For example, participants in the intensity condition were asked which of two stimuli sounded like the speaker was conveying a higher degree of the quality being denoted by a given gradable adjective (stripedness, happiness, windiness). The results of that study suggested that stimuli with increased duration and an intensifying prosodic profile had an increased likelihood of being selected as sounding more intensifying when compared to stimuli featuring contrastive prosodic profiles and shorter duration, a configuration more typically associated with contrastive focus. In this chapter, I build on the results of that study by using the same experimental paradigm and the same stimuli to test the effect of duration and prosodic profile on interpretations of expressive intensity related to stronger opinions. Findings from the present study will be compared to findings from the semantic intensity condition of the prior study to identify any differences in the interpretation of duration and pitch for interpretation related to pragmatic vs. attitudinal meaning. The results from this study will provide us with a better understanding of a) the realization of prosodic intensification more generally and b) the mapping of acoustic variation in duration and pitch onto different types of intensification.

**Experiment 1**

*Introduction*

In this experiment, I test whether the acoustic cues that I found in Chapter 2 to be more strongly associated with semantic intensity (intensifying prosodic profiles and longer duration) are also the stronger cues to attitudinal meaning. Specifically, I test whether these cues predict an increase in the likelihood that an adjective is selected as sounding like it conveys a stronger opinion when compared to stimuli with acoustic cues more strongly associated with contrast.
Alternatively, I entertain the notion that duration may have different effects on interpretation when intensity is targeting the expressive dimension, specifically. In the following section, I describe the stimuli in detail before discussing the specific predictions for this experiment.

**Materials**

The stimuli used in this experiment were the same stimuli used in the experiments discussed in Chapter 2. Specifically, six gradable adjectives (*bright, mad, sour, happy, salty,* and *windy*) were naturally produced in the sentence frame “The NOUN is ADJ.” Since the research question in Chapter 2 concerned interpretations of contrast and semantic intensity, each adjective was originally produced in two versions: one version that featured a prosodic profile intended to convey contrast, and one that featured a prosodic profile intended to convey semantic intensity. The contrastive and intensifying adjectives exhibited the same duration and pitch characteristics discussed in Chapter 2 (see Section 2.5 for more detail).

Although duration and intensity properties of these recordings were manipulated in creating the stimuli, the original, naturally produced pitch productions over the entire sentence were retained in order to preserve all aspects of the signal that may be linked to prominence, including voice quality and the overall shape of the F0 trajectory. These original productions differed in their F0 trajectories in a variety of ways, including in the shape of the F0 movement on the adjective in addition to the scaling of the prenuclear material (Figure 18).

**Figure 18**

*Average F0 Across the Adjectives Produced with Intensifying and Contrastive Prosodic Profiles*
Note. The raw F0 trajectories for all six adjectives produced with the intent to convey contrast (left panel) vs. semantic intensity (right panel). The blue portion of the line represents the F0 trajectory over the material spanning from the beginning of the sentence through the word immediately preceding the critical adjective (e.g., ‘the shirt is’). The portion of the contour outlined in red represents the F0 trajectory over the critical adjective itself. The solid black line illustrates the average F0 trajectory (with the average duration) for the contrastive and intensifying conditions.

Eight types of stimuli were created for each adjective. These stimuli were created by manipulating the onset and nucleus of each adjective realized with each prosodic profile to create a long and short version of each. The short duration condition equaled a decrease in duration of 40% from the duration of the original recording of the sentence produced as contrastive, while the long duration equaled an increase of 40% from the original intensifying duration. Two pitch conditions crossed with two duration manipulations in two segments (onset and nucleus) led to the creation of eight unique stimuli for each adjective, as shown in Table 3.
Table 3

Stimuli Utilized in Experiment 1

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Prosodic Profile</th>
<th>Onset Dur</th>
<th>Nucleus Dur</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CLL</strong></td>
<td>Contrastive</td>
<td>Long</td>
<td>Long</td>
</tr>
<tr>
<td><strong>CLS</strong></td>
<td>Contrastive</td>
<td>Long</td>
<td>Short</td>
</tr>
<tr>
<td><strong>CSL</strong></td>
<td>Contrastive</td>
<td>Short</td>
<td>Long</td>
</tr>
<tr>
<td><strong>CSS</strong></td>
<td>Contrastive</td>
<td>Short</td>
<td>Short</td>
</tr>
<tr>
<td><strong>ILL</strong></td>
<td>Intensifying</td>
<td>Long</td>
<td>Long</td>
</tr>
<tr>
<td><strong>ILS</strong></td>
<td>Intensifying</td>
<td>Long</td>
<td>Short</td>
</tr>
<tr>
<td><strong>ISL</strong></td>
<td>Intensifying</td>
<td>Short</td>
<td>Long</td>
</tr>
<tr>
<td><strong>ISS</strong></td>
<td>Intensifying</td>
<td>Short</td>
<td>Short</td>
</tr>
</tbody>
</table>

Design and Procedure

Unlike in Chapter 2, when participants completed a 2 AFC task that required them to select which of two stimuli conveyed a stronger degree of the quality being denoted by the critical adjective, in the experiment discussed here, participants were asked in which of the two sentences presented to them the speaker sounded like they had a stronger opinion about the quality they were describing. Four of the comparisons were between stimuli with different prosodic profiles (hereafter, across-pitch): one with a contrastive prosodic profile and one with
an intensifying prosodic profile. The other two comparisons were between stimuli with the same prosodic profile (within-pitch): They compared the SS (short onset and short nucleus) and LL (long onset and long nucleus) conditions, or the SL and LS conditions for a single prosodic profile (e.g., CSS vs. CLL, CSL vs. CLS, ISS vs. ILL, ISL vs. ILS). These within-pitch trials allowed us to test for a difference between overall lengthening (SS vs. LL) or between lengthening of the onset vs. nucleus, respectively (SL vs. LS), for each prosodic profile.

The experiment was between-subjects. Half of the participants heard trials that compared a full range of manipulations for intensifying prosodic profiles (e.g., ISS, ILL, ISL, and ILS) with two levels of duration for the contrastive prosodic profiles (e.g., CSS and CLL). The other half of the participants heard trials that compared a full range of duration manipulations for the contrastive prosodic profiles (e.g., CSS, CLL, CSL, CLS) with two levels of duration for the intensifying prosodic profiles (ISS and ILL). This design yielded the critical comparisons indicated in Figure 19.

**Figure 19**

*Comparisons Tested in Experiment 1*
### Intensifying Half

<table>
<thead>
<tr>
<th>Across Pitch</th>
<th>Within Pitch</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ILL</strong></td>
<td><strong>ISS</strong></td>
</tr>
<tr>
<td><strong>CLL</strong></td>
<td><strong>X</strong></td>
</tr>
<tr>
<td><strong>CSS</strong></td>
<td><strong>X</strong></td>
</tr>
</tbody>
</table>

### Contrastive Half

<table>
<thead>
<tr>
<th>Across Pitch</th>
<th>Within Pitch</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ILL</strong></td>
<td><strong>CSS</strong></td>
</tr>
<tr>
<td><strong>CLL</strong></td>
<td><strong>X</strong></td>
</tr>
<tr>
<td><strong>ISS</strong></td>
<td><strong>X</strong></td>
</tr>
</tbody>
</table>

*Note.* The stimuli comparisons in the across-pitch and within-pitch trials in each half of the experiment. Only cells marked with an X were tested in the experiment.

The comparisons were divided into two blocks based on the duration condition tested: one block had trials comparing stimuli in all four duration levels of one prosodic profile with the LL duration level of the other prosodic profile, while the other block had the same trials but substituting the SS duration level of the other prosodic profile. The same within-pitch trials were included in both blocks. For example, in the intensifying half of the experiment (testing the full range of intensifying manipulations), Block 1 trials compared all four of the duration manipulations for the intensifying stimuli (ILL, ISS, ISL and ILS) to the CLL condition in addition to trials with the two within-pitch comparisons, while Block 2 compared all four of the intensifying durations to the CSS condition, in addition to the same two within-pitch comparisons.
comparisons. This resulted in a total of six comparisons for each adjective in each block: four across pitch (e.g., ILL v. CLL, ISS v. CLL, ISL v. CLL, ILS v. CLL) and two within (e.g., ILS v. ISL, ISS v. ILL). Whether the intensifying or contrastive stimulus appeared on the left or the right was counterbalanced across block. Participants heard each block twice: once with each of the presentation orders.

Within each block, the stimuli were organized into six sub-blocks based on adjective, so that each adjective occurred once in each block, each in a different condition. The adjectives were then rotated through all conditions over the course of the six blocks. Six adjectives produced in six sub-blocks over 4 blocks resulted in a total of 144 trials.

Participants completed the experiment remotely using their own laptop or desktop computer and headphones. Before beginning the experiment, they were required to test their headphones for sound quality. On each trial, participants were presented with two audio files and told to select the sound in which they thought the speaker sounded like they had a stronger opinion about the quality of the object they were describing (e.g., for the sentence the shirt is bright the participant was asked, In which sentence does the speaker sound like they have a stronger opinion about the brightness of the shirt?).

Hypotheses and Predictions

In this chapter, I hypothesize that semantic intensity and expressive intensity are conveyed by the same acoustic realizations of prominence and are thus non-separable in terms of their meaning contribution. In the experiment conducted here, I predict that, because of the close connection between semantic and expressive intensity, the same stimuli will be selected as conveying the speaker’s strong opinion that were selected in Chapter 2 as conveying semantic
intensity: the stimuli with longer duration and an intensifying pitch profile are predicted to be more often selected under both interpretive criteria.

Participants

A total of 119 participants were recruited for this experiment from the Northwestern University Linguistics Department Undergraduate Subject Pool. The participants were all between the ages of 18-22 and self-reported as having normal-to-corrected hearing and vision and no reading or learning impairments. All participants were fluent speakers of American English who had been living in the United States for at least 10 years. Because the participants were between the ages of 18 and 22, this meant that about 50% of their life had to have been spent in the United States. Twenty-three participants were excluded for not fitting these criteria, resulting in a total of 96 participants (47 in the intensity condition and 52 in the contrast condition).

Analysis and Results

For each experiment, I conducted two separate analyses: one for the within-pitch trials and one for the across-pitch trials. The within-pitch analysis (e.g., CLL vs. CSS, ILL vs. ISS) tests whether increased duration increases the likelihood of an interpretation of expressive intensity for both the contrastive and intensifying prosodic profiles. The across-pitch analysis (e.g., CLL vs. ILL, CLL vs. ISS) tests whether stimuli with longer duration provide a stronger cue to interpretations of expressive intensity for contrastive vs. intensifying prosodic profiles when the two prosodic profiles are directly pitted against one another. Both analyses also probe

---

5 The decision to analyze these trials separately was made because of the difference in the type of comparisons applied to each stimulus. For example, the CLL stimulus was compared to the ILL, ILS, ISS, ISS and CSS stimulus, while the CLS stimulus was compared to the ILL, ILS, ISL, ISS and CSL stimulus.
the interaction between duration and prosodic profile (i.e., whether increased duration increases the likelihood of an interpretation of expressive intensity when it is paired with an intensifying vs. contrastive prosodic profile). The two analyses and their respective results are presented in turn below.

**Within-Pitch Analysis.** The within-pitch trials allow us to test for a main effect of duration on the likelihood of an expressive meaning, independently for each prosodic profile. To analyze the effect of overall longer duration (i.e., lengthening in both the onset and nucleus) on the interpretation of speaker opinion, I fit a mixed effects logistic regression model to the LL vs. SS trials for both the contrastive and intensifying prosodic profiles. The model predicted the likelihood of picking the LL stimulus from the fixed effects of prosodic profile and stimulus order. Both prosodic profile and stimulus order were sum-coded, with reference levels set to the contrastive prosodic profile and order 1, respectively. The model also included random intercepts for participant and adjective. As shown in Figure 20, the model results indicate, first, that when holding prosodic profile and stimulus order constant, the LL stimuli had a slightly higher likelihood of being selected as conveying that the speaker had a stronger opinion than the SS stimuli. This result is seen in the model intercept, showing that the LL stimuli had a selection probability of about 60%, although this intercept was not significant ($\beta_0=0.38, z=1.57, p=0.11$). The model results also suggest that the probability of selecting the LL stimulus as sounding like it conveyed a stronger opinion was significantly higher when the prosodic profile was intensifying ($\beta_1=1.522, z=4.493, p<.001$) and when it appeared on the right as opposed to the left ($\beta_1=0.326, z=3.125, p<.001$). This finding suggests that lengthening only significantly increased
the likelihood of selection when the prosodic profile was the one most typically associated with semantic intensity (i.e., the prosodic profile featuring lower pitch).

Moving on to the SL vs. LS trials, the model predicted the likelihood of selecting the SL stimulus (i.e., the stimulus with nucleus lengthening) from the fixed effects of prosodic profile (contrastive vs. intensifying) and stimulus order. Once again, the model included random intercepts for both participant and adjective. As shown in Figure 20, the model shows, first, that holding all other factors constant, the SL stimulus had a lower likelihood of being selected as conveying the speaker’s stronger opinion than the LS stimulus. This result is seen in the significant intercept of the model, showing that the SL stimulus has about a 40% chance of being selected ($\beta_0=0.38$, $z=-2.12$, $p<0.05$). The model results also suggest that the likelihood that the LS stimulus was selected decreased even further when paired with an intensifying prosodic profile ($\beta_1=-0.88$, $z=-4.694$, $p<.001$), suggesting that, when the pitch profile was the one conveying semantic intensity, lengthening in the onset was even less compatible with interpretations of expressive intensity. Finally, on these trials, the effect of stimulus order was not significant ($\beta_1=0.035$, $z=0.371$, $p=0.711$).
Figure 20

*Predicted Probability of Selection for the Within-Pitch Trials*

*Note.* Predicted probability and the 95% confidence interval of selecting the LL (left) and SL (right) stimulus in the within-pitch trials.

Overall, the results from the within-pitch trials suggest that longer duration increases the likelihood that an adjective is selected as sounding like it conveys a stronger opinion only when that adjective is also accompanied by an intensifying prosodic profile. Furthermore, the model results suggest that lengthening in the onset is a much more powerful cue to interpretations of expressive intensity than nucleus lengthening, as the LS stimulus was selected more than 50% of the time regardless of whether it featured a contrastive or intensifying prosodic profile.
Across-pitch Analysis. The analysis of the across-pitch trials involved fitting a Bradley-Terry model (Bradley & Terry, 1952) to each half of the experiment using the BradleyTerry2 package in R (Firth & Turner, 2012). Just as in Chapter 2, in the two models used for this analysis, the ‘players’ were entered into the model as a six-level factor (Contrastive Half: CLL, CLS, CSL, CSS, ILL, ISS; Intensifying Half: ILL, ILS, ISL, ISS, CLL, CSS). In the intensifying half of the experiment, ISS was set to the reference level, while in the contrastive half of the experiment, CSS was set to the reference level. The model also included a contest-specific predictor for stimulus order (1 vs. 2).

The model applied to the contrastive half of the experiment tests the effect of duration on contrastive prosodic profiles for judgments of expressive intensity. As shown in Figure 21, for stimuli featuring contrastive prosodic profiles, the model results illustrate that duration did not have a large effect on the likelihood that contrastive prosodic profiles conveyed stronger opinions than intensifying prosodic profiles. Among the contrastive stimuli, the only significant difference was between the CSL and CSS stimuli, with the likelihood of selecting the CSL stimuli being significantly less than the likelihood of selecting the CSS stimuli ($\beta=-0.29$, $z=-2.27$, $p<.05$). The model also shows that the stimuli characterized by intensifying prosodic profiles were much more likely to be selected (in this half of the experiment) than contrastive prosodic profiles, regardless of their duration manipulation. For example, both the ILL ($\beta=2.07$, $z=20.86$, $p<.001$) and ISS ($\beta=1.70$, $z=18.2$, $p<.001$) stimuli had significantly higher chances of winning a given matchup than the CSS stimuli. The model further shows that the longest intensifying stimulus (ILL) had a significantly higher likelihood of being selected as conveying a stronger opinion across all of durational variants of the contrastive prosodic profiles, compared to the ISS stimulus in the same comparisons ($\beta=0.372$, $z=4.12$, $p<.001$).
Figure 21

Log Ability Scores for an Interpretation of Expressive Intensity vs. Semantic Intensity

Note. Log ability scores for each unique stimulus when participants were asked about expressive intensity (top) as opposed to semantic intensity (bottom) in both halves of each experiment.
The model applied to the intensifying half of the experiment tests the effect of duration on intensifying prosodic profiles for judgments of expressive intensity. The model results once again illustrate that intensifying prosodic profiles were more likely to be selected as conveying stronger opinions compared to both levels of duration of the contrastive stimuli, regardless of their duration. However, they also suggest that the likelihood that an intensifying stimulus was selected was higher for adjectives featuring increased duration in both the onset and nucleus (ILL) ($\beta=0.56$, $z=4.57$, $p<.001$) or lengthening on the onset only (ILS) ($\beta=0.01$, $z=6.80$, $p<.001$) compared to ISS. Furthermore, the stimulus featuring lengthening only in the onset (ILS) was significantly more likely to be selected than the stimulus featuring lengthening in both the onset and nucleus (ILL) ($\beta=0.34$, $z=2.43$, $p<.05$), suggesting that onset lengthening for intensifying prosodic profiles is a particularly effective cue of expressive intensity.

Putting the results from both models together, it appears that prosodic profile has the greatest effect on the interpretation of expressive intensity, as intensifying prosodic profiles were overall more likely to be selected as sounding like they conveyed stronger opinions than the contrastive prosodic profiles. However, the results also suggest that, for intensifying prosodic profiles, lengthening can boost a meaning of expressive intensity, and especially when that lengthening occurs in the onset alone, rather than overall lengthening of onset and nucleus.

To address my experimental question about the interpretation of acoustic cues as conveying semantic vs. expressive intensity, I can qualitatively compare the results from the experiment conducted here with those from Experiment 1a in Chapter 2, when I asked participants to make judgments about semantic intensity (pictured again in Figure 20), for the same stimuli. Overall, the patterns of results are largely the same: In both experiments,
intensifying prosodic profiles were much more likely to be chosen as sounding like they conveyed semantic/expressive intensity than contrastive prosodic profiles. Furthermore, a durational boosting effect was evident for the intensifying prosodic profiles in both experiments, with longer versions of the intensifying prosodic profiles being more likely to be selected than those featuring short duration. The primary difference between the results from these two studies appears to concern the effect of lengthening in the nucleus only vs. nucleus and onset for the intensifying prosodic profiles. For example, when asked about expressive intensity, onset lengthening increased the likelihood of an interpretation of expressive intensity to a greater degree than lengthening in the onset and nucleus together. However, in the experiment testing semantic intensity, the ILL and ILS stimuli did not have significantly different chances of winning a given matchup. This finding suggests that onset lengthening may be more tied to expressive meaning than overall syllable lengthening.

Slight differences in the effect of duration on contrastive prosodic profiles between the two experiments were also evident. For example, when participants were asked about expressive intensity, the shortest contrastive stimulus (CSS) actually had a higher likelihood of being selected than the longest contrastive stimulus (CLL). This contrasts with the findings from the experiment testing interpretations of semantic intensity, when the CLL stimulus was much more likely to be selected than the CSS stimulus (though none of the contrastive stimuli were selected more often than any intensifying stimulus).

Discussion

The results indicate that, overall, the stimuli featuring intensifying prosodic profiles sounded like they conveyed stronger opinions than the stimuli featuring contrastive prosodic
profiles. These results echo those discussed in Chapter 2, when listeners were asked to make judgments about semantic intensity, which were also facilitated by intensifying prosodic profiles. The results also suggest an effect of duration on interpretations of stronger opinions, although this effect was specific to the prosodic profile. For example, while the LL stimulus tended to be perceived as more attitudinally charged for the intensifying prosodic profiles, the same was not true for the contrastive profiles.

There were also interesting differences in the influence of duration on the likelihood of being selected for both the contrastive and intensifying prosodic profiles when participants were asked to interpret prominence as a cue to expressive intensity as opposed to semantic intensity in Chapter 2. For example, I previously found (in Chapter 2) that there was a decrease in the odds of winning a given matchup moving from the longest intensifying stimulus to the stimulus featuring lengthening in only the onset to the stimulus featuring lengthening in only the nucleus to the stimulus featuring no lengthening at all: ILL > ILS > ISL > ISS. In this experiment, however, the odds of being selected were greatest for the stimulus featuring lengthening in the onset only (ILS), followed by the stimulus featuring lengthening in both the onset and nucleus (ILL): ILS > ILL > ISL, ISS. Niebuhr (2010) suggests that onset lengthening is especially linked to negative attitudes in the context of semantic intensity. Thus, it is interesting that onset lengthening boosts the interpretation for both dimensions of intensity, which aligns with an intuitive sense that negative attitudes are more salient than positive ones. It is also interesting that this cue (onset lengthening) becomes even more salient in the prosodic context most strongly associated with expressive intensity (the lower pitch profile), and that this effect vanishes for the stimuli featuring contrastive (higher pitch) prosodic profiles, where participants did not distinguish
between lengthening in the nucleus vs. onset. This result suggests that when the prosodic profile most typically conveys a meaning unrelated to intensity (such as contrast), asymmetric lengthening of the onset has less of an effect on interpretation.

Moving on to the contrastive prosodic profiles, I previously observed that, when participants were asked about interpretations of semantic intensity, the CSS condition always had the lowest odds of winning a given matchup. However, in this experiment, CSS no longer had the lowest odds of being selected. Rather, CSS seemed to have slightly higher odds of being selected than the stimulus featuring lengthening in the nucleus and onset (CLL). In his investigation of the realization of prosodic intensification in German, Niebuhr (2010) associates the ‘barked’ quality of shortened vowel nuclei with negative attitudes (Niebuhr, 2010). Therefore, the reason a CSS stimulus might have been particularly compatible with stronger opinions, as suggested in this experiment, might be because of its shortened nucleus. Thus, it is possible that the increase in the odds for the SS stimuli may have to do with a connection between expressive meaning and shortening, rather than lengthening.

In this experiment, I asked participants to assess which stimulus was conveying a stronger opinion, but I did not address the valence of that opinion (i.e., whether that opinion was negative or positive). It is possible that information about the valence and strength of an attitude has implications for the way that listeners associate different realizations of prominence with expressive intensity. Intuitively, negative attitudes seem stronger than positive ones, and this intuition supports the finding here that ILS, which Niebuhr (2010) links to negative attitudes, is the stimulus that is the most compatible with interpretations of stronger opinions. In Experiment 2, I attempt to explore the influence that attitude strength and valence have on interpretations of
expressive intensity by asking participants to provide information about the perceived valence and strength of the attitude conveyed by each of the unique stimuli tested in Experiment 1. This experiment will enable us to determine whether the reason listeners find stimuli featuring lengthening in the onset only to be the most compatible with interpretations of stronger opinions in Experiment 1 is because of its connection to negative attitudes as opposed to positive ones.

**Experiment 2**

In this experiment, I investigate the valence and strength of each of the unique stimuli used in Experiment 1. The results will shed light on the extent to which information about the strength and valence of an attitude influences the degree to which duration and prosodic profile impact interpretations of expressive intensity. Specifically, I predict that, based on Niebuhr’s (2010) observations for positive vs. negative intensity in German, stimuli characterized by intensifying prosodic profiles and lengthening in the onset will convey negative attitudes, while stimuli characterized by contrastive prosodic profiles and lengthening in the nucleus will convey more positive attitudes.

**Materials**

The materials for this experiment were the same materials used in Experiment 1. There were six critical adjectives produced with two prosodic profiles (contrastive vs. intensifying) and crossed with four duration conditions (LL, LS, SL, SS), resulting in a total of 48 unique stimuli.

**Procedure and Design**

The stimuli were organized with a hierarchical block structure. At the lower level, the stimuli were organized into four sub-blocks, with each adjective occurring once in each block: three with intensifying prosodic profiles and three with contrastive prosodic profiles. Across the
four sub-blocks, each adjective appeared in each of the four duration conditions. At the higher block level, the stimuli were organized into four blocks, so that each adjective appeared twice with each prosodic profile. In blocks 1 and 3, three of the adjectives appeared with contrastive prosodic profiles and three with intensifying, and in blocks 2 and 4, each of these adjectives appeared with the opposite prosodic profile. Each of the 48 stimuli was presented twice, resulting in a total of 96 trials.

Participants completed the experiment remotely using their own desktop computers. They were instructed to wear headphones for the duration of the experiment and to complete it in a quiet location. At the beginning of the experiment, they were given the opportunity to test the quality of the sound recordings and adjust the volume. On each trial, participants were presented with a single audio file and asked to rate how they thought the speaker felt about the quality of what they were describing on a scale from negative to positive. For example, after hearing that the shirt was striped, the participant would be asked: “Based on the way they said this sentence, how do you think the speaker feels about the sourness of the food?” Participants provided their answer using a five-point Likert scale spanning from 1-5, with 1 representing the most negative attitude, 3 representing a neutral attitude, and 5 representing the most positive attitude.

Participants

Thirty-six undergraduate students currently enrolled in an introductory linguistics courses at Northwestern University were recruited for this experiment. The participants were all between the ages of 18-65 and reported no vision, hearing, reading, or learning difficulties. All participants were fluent speakers of American English who were born in the United States. Six
participants were excluded for not satisfying these conditions, resulting in a total of 30 participants.

**Analysis and Results**

Figure 22 illustrates the average deviation from a neutral rating (in this case, a rating of 3, re-labeled on this graph with a label of 0) for each prosodic profile x duration condition. From a visual inspection of the data, stimuli featuring intensifying prosodic profiles were overall rated as sounding more negative (i.e., lower) than those featuring contrastive prosodic profiles. Furthermore, ratings tended to increase (become more positive) as duration decreased.

**Figure 22**

*Mean Deviation from a Neutral Rating*
A Bayesian mixed effects model was applied to the data using the brms package in R (Bürkner, 2017). The model predicted variation in stimulus ratings from the fixed effects of prosodic profile (contrastive vs. intensifying), onset duration (long vs. short), and nucleus duration (long vs. short) in addition to the interaction between nucleus and onset duration and the three-way interaction between nucleus/onset duration and prosodic profile. The model also included by-participant and by-adjective intercepts. All predictors were scale sum-coded with contrastive prosodic profiles and short duration set as reference levels. Rating was entered into the model as an ordered factor.

There were not significant main effects but the model did indicate a significant interaction between onset and nucleus duration ($\beta=-0.07$, CI[-0.23, 0.09]), suggesting that a long onset and short nucleus (i.e., the LS condition) decreased ratings, meaning that it made the speaker sound like they were conveying a more negative opinion. This finding supports the hypothesis that lengthening in the onset is associated with more negative attitudes. There was also a significant interaction between intensifying prosodic profiles and a short nucleus ($\beta=-0.09$, CI[-0.25, 0.07]), suggesting that asymmetric shortening of the nucleus while holding onset duration constant in a stimulus featuring an intensifying prosodic profile further lowered ratings in comparison to shortening in the nucleus for contrastive prosodic profiles. Finally, there was a significant three-way interaction between intensifying prosodic profiles, a long onset, and a short nucleus ($\beta=-0.09$, CI[-0.57, 0.10]), suggesting that the combination of having a long onset and short nucleus decreased ratings more for intensifying prosodic profiles than for contrastive ones. Overall, the results from this model suggest that, consistent with my hypothesis, the ILS stimulus was heard as conveying a particularly negative attitude. However, contrary to my original
hypothesis, as shown in Figure 21, this stimulus did not necessarily sound more negative than stimuli featuring lengthening in both the nucleus and onset. For example, both the ILL and ILS stimuli appeared to have similar ratings.

Although not statistically significant in my model, the empirical results suggest that, contrary to my hypothesis, stimuli featuring contrastive prosodic profiles were not rated as conveying particularly positive attitudes. Rather, on average, they were rated at or slightly above neutral, while intensifying stimuli were rated consistently below neutral. Thus, the results confirm my hypothesis that intensifying prosodic profiles are more compatible with negative attitudes, although they do not support the hypothesis that contrastive prosodic profiles are more compatible with positive attitudes insomuch as they suggest that they are less compatible with negative ones.

It is important to note that the adjectives included in the experimental materials may have been differentially affected by the duration and prosodic profile manipulations, considering they may have been benefitting from a more positive/negative semantic starting point. However, as shown in Figure 23, it did not appear that any of the adjectives differed substantially from the others in terms of its overall likelihood to be rated as positive or negative.
Figure 23

*Average Rating Per Adjective Across Duration and Prosodic Profile*

![Graph showing average rating per adjective across duration and prosodic profile.](image)

*Note.* The average rating of each of the six adjectives used in this study, collapsing across duration and prosodic profile.

For example, while *bright* and *happy* were rated overall more positively than a more negative adjective like *mad*, the differences are not large, and in fact, even the adjectives with more positive semantics are not rated particularly positively.

*Discussion*

In this experiment, I investigated the valence and strength of the attitude being conveyed by adjectives featuring long and short duration in their onsets and nuclei in addition to
contrastive vs. intensifying prosodic profiles. The results suggest that stimuli featuring intensifying prosodic profiles and lengthening in either the onset only or in both the onset and nucleus were rated as more negative than stimuli featuring contrastive prosodic profiles or shorter duration. Stimuli featuring contrastive prosodic profiles, on the other hand, were not necessarily rated as sounding positive, but were certainly heard as sounding less negative.

Taking these results into consideration as I re-consider the findings in Experiment 1, the rating data discussed here suggest that interpretations of expressive intensity are influenced by the strength of the attitude being conveyed, as the stimuli that were rated as deviating the most from a neutral rating were also more likely to be selected as sounding compatible with interpretations of stronger opinions. For example, the stimuli that were rated a) the most negatively and b) as deviating the most from neutral (i.e., the ILS and ILL stimuli) were also the most likely to convey expressive intensity in Experiment 1. Thus, when asked to interpret combinations of prosodic profiles and duration as conveying stronger opinions as in Experiment 1, participants appeared to consider stimuli conveying negative attitudes (ILL, ILS) as conveying stronger opinions than stimuli conveying more neutral attitudes.

**Overall Discussion**

Overall, the results support the hypothesis that listeners associate different prosodic profiles and configurations of duration with interpretations of semantic vs. expressive intensity, as the prosodic profile and duration characteristics that boost interpretations of semantic intensity are somewhat different from the characteristics that boost interpretations of expressive intensity, operationalized here as expressing stronger speaker opinion. For example, in Chapter 2, stimuli featuring intensifying prosodic profiles and any amount of lengthening (ILL, ILS, ISL) had a
higher likelihood of being chosen (over a contrastive stimulus) as conveying semantic intensity, compared to the intensifying stimuli featuring no lengthening (ISS). Furthermore, the ILL stimulus had the highest likelihood of being selected as conveying semantic intensity when compared to a contrastive stimulus. However, in the experiments conducted here, participants judged intensifying stimuli with lengthening only in the onset (ILS) as more strongly conveying expressive intensity, compared to stimuli featuring lengthening in both the onset and nucleus. This result suggests that duration functions differently in cueing to expressive intensity vs. semantic intensity, as participants preferred different configurations of duration over the onset and nucleus as being the most representative of expressive intensity than for semantic intensity.

However, while onset and nucleus duration pattern differently as cues to different types of intensity, they can also operate in tandem. For example, in Chapter 2, when participants were asked about semantic intensity, onset lengthening (LS) boosted the interpretation of semantic intensity, more so than nucleus lengthening (SL). Based on the results for expressive intensity as discussed in this chapter, it is possible then that in the semantic intensity experiment, onset lengthening increased the likelihood that a stimulus was selected as sounding like it conveyed semantic intensity specifically because of its connection to expressive meaning. Intuitively, expressive meaning would seem to be highly intertwined with the meaning category of intensification, as particularly high degrees of adjectival qualities often lead to an emotional response. Thus, when listeners heard onset lengthening in the semantic intensity experiment, it is possible that they made an inference about semantic intensity based on the strong attitudinal meaning being conveyed by this realization of prominence. The fact that lengthening both the onset and nucleus (LL) produced an even greater boost in the interpretation of semantic intensity
than for expressive intensity suggests that lengthening overall may be directly connected to semantic intensity, but asymmetric lengthening in the onset alone is only associated with semantic intensity mainly by virtue of its connection with attitudinal meaning.

The results also suggest an interaction with prosodic profile. For example, in the experiment testing expressive intensity, onset lengthening (LS) boosted selection only for stimuli featuring Intensifying prosodic profiles. Similarly, when participants were asked about semantic intensity, onset lengthening (LS) was more compatible than nucleus lengthening (SL) only when accompanied by an intensifying prosodic profile. This finding suggests that the effects of expressive meaning on duration as a cue to intensity is only licensed given a compatible prosodic profile. As Niebuhr (2010) suggests, and as the results from my rating experiment show, onset lengthening was perceived as sounding more negative than nucleus lengthening. According to Niebuhr (2010), intensification conveying negative attitudes is also conveyed by lowered and narrowed pitch courses. Therefore, the fact that onset lengthening in my experiment was boosted interpretations of semantic or expressive intensity only when accompanied by an intensifying prosodic profile (which featured lower pitch) suggests that durational effects on intensifying interpretations are dependent on a compatible prosodic profile.

Support for the idea that onset lengthening is a particularly productive cue to semantic intensity because of its ability to convey expressive intensity can be found by comparing the results from the experiments addressing semantic intensity vs. contrast as discussed in Chapter 2 (and pictured again in Figure 23).

**Figure 24**

*Log Ability Scores for Interpretations of Semantic Intensity vs. Contrast*
Note. Log ability scores for each unique type of stimulus when asked about interpretations of semantic intensity (top panel) and contrast (bottom panel).

In the presence of a semantically intensifying context, where attitudes and expressive meaning are presumably more salient/at play than in a more neutral context like contrast, the stimuli that were a stronger cue to expressive intensity, such as ILS, had a higher likelihood of
being selected. The same was not true for interpretations of contrast, when the two stimuli featuring lengthening in only the onset or nucleus were not significantly different from one another in terms of their likelihood of being selected, in either the contrastive or intensifying conditions. This finding suggests that cues that are closely linked to expressive intensity only influence interpretations of meaning types that are closely intertwined with expressive meaning (such as semantic intensity), while these effects do not shine through for more attitudinally neutral contexts.

If prosodic prominence can indeed convey semantic intensity partly because of its ability to convey stronger expressive meaning, then one interesting question to consider is how an intensifying accent differs from accents that have been previously described as characteristic of exclamative sentence types and of exclamations more generally. Both *wh*-exclamatives (*What a tall tree that is!* and declarative exclamations (*The tree is tall!* are generally analyzed as conveying a violation of expectations on the part of the speaker (Rett, 2011). For *wh*-exclamatives, this violation of expectations is specifically argued to involve degrees (e.g., Bolinger, 2013; Espinal, 1995; Michaelis & Lambrecht, 1996). For example, according to Rett (2011), a *wh*-exclamative expresses that the speaker considers it noteworthy that the degree property corresponding to the exclamative is instantiated by some value that exceeds the relevant contextual standard, although it does not modify that degree value directly. Beijer (2002) similarly argues that, to qualify as an exclamative, an expression must a) convey strong positive or negative emotions but not propositionalize those emotions, b) indicate high positions on degree or quantity scales that are either explicit or inferable, and c) express a deviation from a norm through a generalized implicature.
Exclamatives are reported to be associated with intonational ‘extremeness’ (Bolinger, 2013). In this sense, exclamatives may be characterized by high or low pitch, as long as that pitch can be considered extreme. Similarly, the contour itself can be either extremely varied or extremely monotone. According to Bolinger (2013), wh-exclamatives universally have an intonation contour in which the degree-word has the (exclamatory) accent and much higher pitch than the rest of the utterance. Similarly, Seeliger and Repp (2023) suggest that German exclamative utterances are characterized by a) falling contours, b) exclamative accents (which are attracted to degree words) and c) a slower speaking rate. Limited research has been done on exclamative accents in English, but in German, Altmann (1993a, 1993b) suggests that exclamative accents differ from accents conveying focus/contrast, as they are characterized by greater maxima with respect to the basic frequency, greater length and possibly a higher intensity (Oppenrieder, 1987, 1989; Batliner, 1988). Rett and Sturman (2020) find that, in a production experiment, both exclamations and exclamatives featured a higher frequency of L+H* accents compared to neutral declarative utterances.

The stimuli used in this study feature declarative syntax and are thus more compatible with descriptions of exclamations rather than exclamatives. However, the fact that my stimuli also involved prominence on a degree word makes it interesting to consider how the prosodic realization of exclamatives, exclamations, and intensity differs. Prosodically, the stimuli used in the study conducted here did indeed exhibit some of the acoustic characteristics of exclamative sentence types, including a slower speaking rate, a falling contour, and an accent that appeared on a degree word. Adjectives that were characterized by listeners as sounding more compatible with interpretations of expressive intensity in this study were not characterized by higher pitch,
but it could be argued that the lower pitch maxima could be considered extreme in its own right, making it possible that intensifying accents as produced in this experiment were more akin to an accent associated with exclamations and exclamatives. Future studies should clarify the relationship between intensification accents and exclamative accents to determine whether the two types of prominence constitute the same or different meaning categories. However, these prior theories of exclamations and exclamatives provide potential support for the hypothesis that prosody may convey information about semantic intensity directly (e.g., through a lower pitch profile and a lengthened onset) but may sometimes also convey information about degrees indirectly via a connection to expressive meaning or mirativity (i.e., surprise). Thus, the fact that when compared to a contrastive stimulus, ILS stimuli boosted the interpretation of semantic intensity more than ISL stimuli in Chapter 2 may be because listeners were considering the linkage between onset lengthening and expressive meaning when asked about semantic intensity.

As discussed in the introduction, prosody conveys expressive meaning related to both attitudes and emotions, and it is often the case that these two dimensions interact. For example, if a speaker has a very strong opinion about something, it is possible that they are also having a strong emotional response. In the experiments here, I did not operationalize the paralinguistic context as an emotional one. However, emotions tend to influence aspects of the phonetic signal such as F0, voice quality, acoustic energy, and the general dynamics of articulatory and phonatory movements (Banse & Scherer, 1996; Mozziconacci, 1998; Chasaide & Gobl, 2004), while attitudes are conveyed at a shorter time frame, perhaps at the syllable level. In my experimental materials, I directly manipulated the duration of the prominent syllable, with the intent to target attitudes rather than emotions. However, as mentioned in my description of the
materials, there were also differences between the contrastive and intensifying stimuli at a more global level. For example, the intensifying and contrastive prosodic profiles exhibited differences in the speech rate and average F0 of the preamble, in addition to differences in voice quality on the critical adjective. Because of the conflation of emotional and attitudinal cues in my stimuli, a better term to describe the type of meaning I tap into in this chapter might be the notion of evaluation, which encompasses both the concept of stance (attitudes) and emotions (Escandell-Vidal, Aguilar, & Oncón, 2014). For example, according to Hunston and Thompson (2000:5), evaluation “[…] is the broad cover term for the expression of the speaker or writer’s attitude or stance towards, viewpoint on, or feelings about the entities or propositions that he or she is talking about.” Future studies might attempt to tease apart the cues that are more closely connected to emotions vs. attitudes in the context of expressive or semantic intensity, perhaps by controlling the emotional context. Furthermore, they might investigate the impact that more global changes have on interpretations of intensity, such as voice quality, as this cue has been reported to play an important role in emotional signaling.

Future studies might also investigate the ability of different realizations of prominence to convey expressive intensity when prominence is realized on words that have degree readings vs. those that do not. For example, lexical intensifiers are only acceptable with gradable adjectives, and their acceptability is further influenced by the underlying structure of the scale introduced by that adjective (e.g., Kennedy & McNally, 2005). Furthermore, Bolinger (2013) points out that intensifiers like quite and rather are only acceptable modifiers of verbs when those verbs are capable of a degree meaning, such for verbs like hesitate or struggle, as opposed to wait or perform. When intensifiers modify non-gradable predicates, they are argued to modify
expressive or pragmatic scales (e.g., Beltrama & Bochak, 2015; Beltrama, 2018). It is an interesting question whether a similar characterization is true for prosody. For example, in the experiment conducted here, when prominence was always realized on a gradable adjective, I saw that the ILS and ILL stimuli were the most compatible with interpretations of expressive intensity. However, it is possible that the reason that the ILL stimulus was so compatible with this interpretation was primarily because of its compatibility with semantic intensity, which might be presupposed in the case of expressive intensity, as has been suggested for exclamatives. Therefore, it would be interesting to know whether the compatibility of changes in pitch and duration with interpretations of expressive intensity differs when an interpretation of semantic intensity is not available (i.e., when prominence is not realized on a degree word). For example, in this case, I might predict that the ILS stimulus would be far-and-above the most compatible stimulus with interpretations of expressive intensity.

An important caveat to the studies conducted here is the fact that the variation in the stimuli was relatively limited. The original, contrastive and intensifying stimuli were produced by a single speaker, and the differences that characterized them were meant to mirror limited observations about prosodic intensification noted for German (e.g., Kohler, 2006; Niebuhr, 2010). Yet, it is certainly possible that expressive intensity is realized in more ways than just those produced here, and in fact there might be considerable variation between speakers in terms of the methods employed. For example, perhaps some speakers choose to convey semantic intensity primarily by conveying information about their emotions. In this case, acoustic qualities like voice quality might be more important than changes in duration on the critical adjective itself. Furthermore, speakers may choose to convey intensity using lexical means only rather
than prominence, and prominence may be less important for a derivation of an intensified meaning than lexical intensifiers more generally, meaning that some speakers do not use it at all to convey intensity. Considering the variability in the possible acoustic realization of prominence conveying intensity, future studies should investigate the acoustic realization of intensity across speakers and contexts, perhaps using a production paradigm.

**Conclusion**

In this chapter, I test the effect of duration and prosodic profile on interpretations of expressive intensity, examining how these effects differ for those for semantic intensity as discussed in Chapter 2. In Experiment 1, I played participants sentences featuring adjectives with contrastive and intensifying prosodic profiles that differed in their duration and asked them to select in which sound file the speaker sounded like they had a stronger opinion about what they were describing, comparing these findings to those from Chapter 2. The results suggest that, similar to interpretations of semantic intensity, listeners consider increased duration and lowered pitch courses to be a particularly strong cue to interpretations of expressive intensity. However, unlike for interpretations of semantic intensity, the strongest cue configuration to interpretations of expressive intensity was lengthening only in the onset accompanied by intensifying prosodic profiles, rather than lengthening in both the onset and nucleus. In Experiment 2, I ask whether this finding has something to do with the valence and/or strength of the opinion being expressed by a given stimulus, and I ask participants to rate how positive or negative each stimulus sounded. The results suggest that stimuli featuring intensifying prosodic profiles and longer duration were also rated as sounding more negative than stimuli featuring contrastive prosodic profiles and shorter duration. Considering the results from both of these experiments in tandem,
the findings from this study suggest that realizations of prominence that convey a) stronger and b) more negative attitudes are strong cues to interpretations of expressive intensity. Overall, the results shed light on the realization of prominence intended to convey intensification more generally, in addition to suggesting a difference between prominence intended to convey intensity targeting semantic scales vs. expressive ones.

**General Conclusion**

In this dissertation, I investigated listener’s associations of different realizations of prominence with interpretations of contrast (focus) vs. semantic/expressive intensity. Since these meanings have all been linked to increases in prosodic prominence, I sought to investigate whether listeners can distinguish between them by attending to the specific acoustic realization of a given instance of prominence and, if so, which realizations of prominence are more associated with intensity as opposed to contrast. I hypothesized that, based on prior observations for the realization of prominence conveying semantic intensity, listeners might be sensitive to differences in the realization of prominence, and that they might use these differences to determine a speaker’s intended meaning. Specifically, I predicted that increased duration would be a reliable cue to both intensified and contrastive meanings (albeit a stronger cue to intensified meanings, especially when that duration was particularly long), while salient pitch movements would more reliably cue interpretations of contrast.

In Chapter 1, I investigated the effect of variation in duration on the ability of prosodic profiles originally intended to convey contrast vs. semantic intensity to convey their respective meanings. Prior studies suggest that increasing prominence via a salient pitch movement can increase the likelihood that a listener arrives at a contrastive interpretation independently (e.g.,
Watson, Tanenhaus & Gunlogson, 2008; Tomlinson, Gotzner & Bott, 2017), but no studies have tested how a listener might distinguish a contrastive meaning from an intensified meaning given a single instance of prominence. Therefore, in this experiment, I played participants sentences featuring prominent gradable adjectives that differed in their pitch and duration and asked them to choose which of two pictures representing interpretations of contrast or semantic intensity they thought the speaker was looking at when they had uttered that sentence. The results suggest that, when asked to make a binary judgment about contrast vs. semantic intensity, listeners associate increases in duration with interpretations of semantic intensity as opposed to contrast. In fact, when the duration was long enough, it did not matter whether an adjective was realized with contrastive (high) vs. intensifying (low) pitch: rather, participants were always more likely to arrive at an intensified interpretation. In a follow-up experiment, I introduced more subtle variation in duration and in the scaling of the pitch contour across the entire utterance to investigate how variation in either dimension might impact the ability of intensifying and contrastive prosodic profiles (characterized by differences in the pitch maxima of the critical adjective but also by differences in speech rate and voice quality) to signal semantic intensity vs. contrast. The results suggest that, for intensifying prosodic profiles, increasing duration increased certainty about interpretations of semantic intensity, while decreasing duration decreased certainty. Conversely, increasing duration decreased certainty about the association of contrastive prosodic profiles with interpretations of contrast, while decreasing duration increased certainty. Modifying the pitch scaling of the entire phrase had no effect on certainty about the meaning being conveyed by intensifying and contrastive prosodic profiles. Overall, the results discussed in this chapter suggest that a) large increases in duration are a stronger cue to interpretations of semantic intensity than intensifying prosodic profiles (which featured lower pitch) and b)
intensifying prosodic profiles can also cue interpretations of semantic intensity when the changes in duration are more gradual, although manipulating duration still has a significant impact on the ability of intensifying profiles to signal semantic intensity.

The results discussed in Chapter 1 suggest that, when asked to make a binary choice about interpretations of contrast vs. semantic intensity, increased duration and lower pitch encourage intensifying interpretations over contrastive ones. In Chapter 2, I approach this question from a different angle, investigating which of two realizations of prominence are most preferred for a single interpretation. For example, I played participants two sound files and asked them to choose which sounded more intensifying (Experiment 1a) or more like a correction (Experiment 1b). Furthermore, I introduced additional variations in duration by including conditions that manipulated the duration of the onset and nucleus independently, asking whether participants prefer lengthening in one over the other when asked about contrast vs. semantic intensity. The results indicate slight differences in the ability of configurations of pitch and duration to cue interpretations of contrast and semantic intensity. For example, when asked about interpretations of semantic intensity, listeners preferred stimuli that featured intensifying prosodic profiles (lower pitch) on the critical adjective in addition to increased duration, and especially stimuli featuring increased duration in both the onset and nucleus. The results were similar for interpretations of contrast, although listeners did not distinguish as much between the stimuli featuring lengthening in the onset vs. nucleus when interpreting prominence as a cue to contrast. In fact, the likelihood that a stimulus was selected as sounding more contrastive appeared to be less affected by changes in duration than it was for interpretations of semantic intensity. Furthermore, in the intensifying experiment, the distinction between lengthening in the
onset vs. nucleus differed for stimuli featuring intensifying vs. contrastive prosodic profiles. For example, a stimulus with a long onset was interpreted as sounding more intensifying when accompanied by an intensifying prosodic profile (lower pitch). However, when accompanied by a contrastive (higher) prosodic profile, the stimulus featuring lengthening in the onset rather than the nucleus had a higher likelihood of being selected as sounding like it conveyed semantic intensity.

In a follow-up experiment, I investigated whether these differences were primarily driven by differences in overall prominence, asking whether the two meanings are simply mapped onto different degrees of prominence as opposed to specific differences in acoustic realization. When participants were instead asked to judge which of two stimuli sounded more prominent, the results suggest that the most prominent stimuli were also the most likely to be chosen as sounding both intensifying and contrastive independently, suggesting that compatibility with either meaning is partly driven by prominence. However, subtle differences suggest that prominence does not tell the whole story. For example, the stimulus featuring lengthening only in the onset was rated as sounding less prominent than the stimulus featuring lengthening only in the nucleus but was chosen as sounding more intensifying than the stimulus featuring lengthening only in the nucleus. Furthermore, the two prosodic profiles were not as clearly separable in terms of their prominence as they were for interpretation, when participants clearly preferred stimuli featuring intensifying profiles with either interpretation. For example, the ISL and CLL stimuli had similar likelihoods of being chosen as sounding more prominent in Experiment 2, while they had very different likelihoods of being selected as sounding more contrastive or intensifying in Experiments 1a and 1b. Overall, these results suggest that the
ability of intensifying and contrastive prosodic profiles paired with long vs. short duration to convey contrast vs. semantic intensity is partly driven by prominence, but listeners appear to show sensitivity to different dimensions of prominence depending on the specific interpretation, suggesting that prominence cannot entirely explain the results discussed here.

Finally, in Chapter 3, I built on the experiments conducted in Chapter 2 by investigating whether participants were associating the realizations of prominence tested in Chapters 1 and 2 with semantic intensity, or whether they were more likely making a secondary inference about semantic intensity based on prosody’s ability to convey expressive intensity. For example, I asked participants to select in which of two stimuli the speaker sounded like they had a stronger opinion about what they were describing. If the stimuli that had the highest likelihood of being selected as sounding like they conveyed stronger opinions were the same as the stimuli that had the highest likelihood of being chosen as sounding like they conveyed semantic intensity in Chapter 2, then this would provide partial evidence for the hypothesis that prosody does not convey semantic intensity directly. However, if different stimuli cued interpretations of stronger opinions compared to the stimuli that cued interpretations of semantic intensity in Chapter 2, then it would suggest that the realizations of prominence conveying semantic vs. expressive intensity differ slightly, although they might show some degree of overlap in their realizations. The results support the latter hypothesis by indicating slight differences among the ranking of stimuli in terms of their ability to cue interpretations of expressive intensity as opposed to semantic intensity. For example, the ILS stimulus (characterized by lower pitch and a long onset) had the highest likelihood of being selected as sounding like it conveyed expressive intensity, while the same was not true in the experiment testing interpretations of semantic intensity,
although this stimulus was still more likely to be chosen when participants were asked about semantic intensity than it was when participants were asked about contrast. This finding suggests that prominence conveying expressive intensity differs slightly in its realization from prominence conveying semantic intensity, although both meanings may be capable of being conveyed at the same time. In the second part of this chapter, I investigated whether the valence of the attitude conveyed by a prominence influenced its compatibility with interpretations of stronger opinions. Specifically, I asked participants to rate the degree to which they thought a given realization sounded positive or negative. The valence ratings suggest that the stimuli that were most likely to be selected as sounding like they conveyed stronger opinions were also the stimuli that were more likely to be selected as conveying more negative attitudes, suggesting that the strength and valence of the attitude being conveyed impacts the ability certain realizations of prominence to convey expressive intensity.

Overall, the findings discussed in this dissertation have several important implications. First, they provide a clearer picture of the prosodic realization that is most typical of prominence conveying intensity (both expressive and semantic). Before this study, support for generalizations about the prosodic realization of intensity was limited to small production experiments and corpus experiments in German (Niebuhr, 2010; Kohler, 2009) in addition to qualitative observations in English (Bolinger, 2013; Armstrong & Ward, 1926). The experiments conducted here bolster previous characterizations of the prosodic realization of intensity by suggesting that listeners do in fact associate increases in duration and lowered pitch courses with an intensified meaning as opposed to a contrastive one. Furthermore, despite arguments that intensity and contrast differ in their acoustic manifestation, to the best of my knowledge this is
the first experiment that directly pits the two meanings against one another, asking whether prominence intended to convey different pragmatic meanings is indeed distinguishable by listeners. Finally, as far as I am aware, this is the first study that attempts to separate out the cues involved in expressive semantic vs. expressive intensity, as previous investigations tend to elicit semantic intensity that is also attitudinally loaded.

While the results discussed here provide insight into a ‘canonical’ prosodic realization of intensity, they also indicate that there is variability within this meaning category in terms of its realization. The output of the Bradley-Terry models allows us to compare a variety of realizations of prominence in terms of their ‘representativeness’ of an intensified meaning. For example, while an ILL stimulus appeared to be the most likely stimulus to convey an intensified meaning in Chapter 2, the ILS stimuli was the next highest ranked. Thus, the output of these experiments provides insight into not only which realizations of prominence are associated with contrast vs. intensity, but also which realizations are more or less typical of either meaning type.

The findings also have implications for previous theories regarding the mapping between prominence and meaning more generally. For example, it is typically argued that increases in pitch and more dynamic pitch courses universally lead to increases in prominence (e.g., Gussenhoven, 2002). Yet, as shown in the prominence experiment conducted in Chapter 2, intensifying prosodic profiles (which featured lowered pitch courses) were more likely to be chosen as sounding more prominent compared to contrastive (higher) ones. This result supports recent suggestions in the literature that judgments of prominence are more influenced by deviation from an unmarked norm rather than a simple association between prominence and higher pitch. For example, Kügler and Genzel (2011) find that pitch register lowering can be
used to convey corrective focus in Akan, a tone language spoken in Ghana. Seeliger and Repp (2023) also illustrate that, in questions, when rising contours are more common, distinctions in information structure can be marked by lower pitch rather than higher pitch, as the unmarked baseline in this case is higher pitch. The sentences used in my experiments were all declaratives, which should have an unmarked baseline that is low, given their realization with falling pitch. Yet, once again, the results from the experiments discussed in this dissertation suggest that, even in declaratives, decreases in pitch can be heard as more prominent. This result is curious. It may stem from the fact that lowered pitch often co-occurs with non-modal changes in voice quality, thus inhibiting the ability of F0 to convey important distinctions in information structure and making sufficient pitch lowering a particularly marked choice. Alternatively, it is possible that other aspects of the signal present in the stimuli used in my study contributed to the marked nature of this pitch course, as it was not only the height of the pitch movement on the critical adjective that distinguished the contrastive and intensifying stimuli, but also differences in the max F0 of the preamble, voice quality, and speech rate.

One question that is not made clear by my experiments is whether contrast and intensity occupy positions on the same meaning scale, or whether they fall on different scales entirely. For example, information-structural categories like givenness and focus are assumed to constitute different but complimentary meaning categories, as both given and new information can be focused. It is therefore interesting to consider where intensification lies in relation to these two scales. For example, is it the case that intensified information is always being focused and is thus a particularly strong type of contrast? Or is it possible for the two meanings to operate independently? One way to investigate this question would be to continue to probe the
relationship between the two meanings in terms of their prominence. For example, as discussed throughout this dissertation, focus appears to be distinguished from givenness by its prominence. Pitch accents marking higher degrees of contrast (i.e., correction) are assumed to be more prominent than those marking higher degrees of givenness (i.e., newness). If contrast and intensity constitute truly independent scales of meaning, then they should map onto the prominence hierarchy in distinct ways. For example, intensity might be able to be conveyed by pitch accents that are lower in prominence than those that are typical of contrast. Alternatively, if intensity is simply a strong subset of contrast, then it should always be realized with the high degrees of prominence typical of contrast. Future studies might test these hypotheses by combining contrast and intensity in the same production experiment or by asking participants to rate the prominence of information that is clearly being intensified, contrasted, or both in a Rapid Prosody Transcription Task.

This dissertation is not without its limitations. The primary limitation concerns the nature of the auditory stimuli, which 1) featured only the voice of a single speaker, and 2) varied in a great number of acoustic dimensions while featuring only a basic manipulation of duration. The decision not to manipulate pitch directly was made in an effort to maintain the naturalness of the stimuli. Furthermore, because of the variety of cues that may play a role in the portrayal of intensity or contrast, it was not possible to control and/or manipulate all potential factors. Therefore, decisions on the creation of the stimuli were based on initial generalizations about mappings between intensity and specific realizations of prosodic prominence (i.e., lowered pitch courses and increased duration) from German (Niebuhr, 2010) and limited observations in English (Bolinger, 2013). Further production studies in English are required to confirm that
intensity is realized in a similar way cross-linguistically. However, I consider the findings from Chapters 2 and 3 to provide initial support for the idea that prominence conveying semantic intensity is realized similarly in English as it is in German. Future studies should consider a variety of other acoustic factors in the perception and interpretation of prominence conveying intensity, including voice quality, which many researchers point out as an important cue to intensification (Niebuhr, 2010; Landgraf, 2014).

The experiments conducted here also did not closely control for the semantics of the adjectives on which prominence was realized, which may influence a participant’s expectations about the likelihood that a given adjective would be intensified. For example, Kennedy and McNally (2005) propose that whether a given adjective introduces a scale that is relative, with thresholds that are highly variable and context-dependent (*big, small*) or absolute, with default minimum or maximum thresholds, influences the type of inferences listeners draw about the likely value of the threshold. Absolute adjectives with minimum thresholds like *open* or *bent* characterize objects as having a non-zero degree of the relevant property, while maximum adjectives like *straight* or *closed* characterize objects has having a maximal degree of the relevant property. As a result, the thresholds for these types of absolute adjectives are much less variable than those for relative adjectives like *tall*. The adjectives used in the experiments conducted here varied in the salience of the measurement scale they were associated with (e.g., *striped*, which is perhaps less typically associated with a scale than a weather term like *rainy*) and in the nature of the scale itself (*striped* is an absolute adjective with a minimum threshold while *rainy* is a relative adjective). Although I did not find significant differences among the adjectives used in my studies in terms of the ability of pitch and duration to signal intensity and
contrast, future studies should examine whether the underlying structure of an adjective’s scale influences listener’s expectations about that adjective’s likelihood of being intensified, and therefore the degree to which they associate changes in pitch and duration with intensified meanings. This dissertation also specifically looked at intensification of adjectives. Yet, as Bolinger (2013) points out, it is also possible to intensify other parts of speech, including nouns and verbs. Thus, future studies might also investigate whether listeners associate the same realizations of prominence with interpretations of intensity for nouns as they do for adjectives, or whether there exist specific realizations for each intensity being conveyed on each unique part of speech.

As discussed in the limitations of the acoustic stimuli, there are many other acoustic factors that could be accounted for or directly tested in future studies. For example, it would be interesting to investigate how vowel type influences the association of patterns of pitch and duration with interpretations of intensity vs. contrast as discussed here, as the stimuli featured in this dissertation were not necessarily balanced in terms of whether the adjectives featured high/low or front/back vowels. Guerrinni (2020) identifies a potential interaction between the felicity of prosodic lengthening and vowel quality. Specifically, he argues that there must be a match between the iconic contribution of the lengthening (longer = bigger) and the sound symbolism (back vowels = large-type meanings and front vowels = small-type meanings) for lengthening to sound felicitous. If this proposal is true, then lengthening of the nucleus in adjectives featuring back vowels may be a stronger cue to interpretations of semantic intensity than the lengthening of a front vowel. Future studies could also control for the sound featured in the onset of the stimuli. In this study, the adjectives sometimes featured stops and sometimes featured continuants, meaning that the stimuli varied in whether the lengthened onset involved
what was essentially a short pause before the beginning of the word or an extended onset consonant. In future investigations, this factor could be explicitly manipulated in order to investigate whether lengthening in either type of sound in the onset serves as a stronger cue to interpretations of intensity/contrast.

An interesting point of discussion for this project concerns the relationship between intensifying prosody and the prosody that is typical of expressives and exclamations more generally. Rett (2011) argues that both exclamations (e.g., *The tree is tall!*) and exclamatives (e.g., *What a tall tree!* ) convey a violation of the speaker’s expectations. In the case of exclamatives, this violation of expectations has to do with the fact that the object being referenced exhibits a particularly extreme or noteworthy degree of the adjective describing it, while for exclamations, this violation of expectations does not necessarily involve degrees. For example, Rett argues that a *wh*-exclamative like *What delicious desserts John bakes!* implies that the speaker expected the desserts John bakes to be less delicious, while an exclamation like *John bakes delicious desserts!* conveys that the speaker wouldn’t have baked delicious desserts (2011: 412). The stimuli used in my studies feature declarative syntax and thus appear to be more consistent with exclamations more generally. However, it is interesting to consider what the relationship is between prosody intended to convey semantic intensity vs. expressive meaning as conveyed in exclamations.

Although my stimuli cannot be considered true exclamatives from the perspective of syntax, I consider previous suggestions for the realization of both exclamations and exclamatives, pointing out a need for further clarity on the specific realization of exclamatives, exclamations, and intensifying accents. For example, *wh*-exclamatives are often associated with exclamative
accents, which are characterized by a higher and later pitch peak when compared to nuclear accents in assertions (German: Batliner, 1988a, 1988b; English: Rett and Sturman, 2020). Rett and Sturman (2020) illustrate that both exclamations and exclamatives are characterized by a greater frequency of L+H* pitch accents, extra-high pitch targets, and additional intermediate phrase boundaries. However, in their production experiment, exclamations and exclamatives did differ slightly in their macrorhythm and in terms of which words specifically were marked. More globally, exclamations are also characterized by a slower speech rate than other speech acts (Altmann, 1933) in addition to rising contours (Rett, 2011), while wh-exclamatives are associated with falling contours (Repp and Seeliger, 2020). Overall, the existence of exclamative speech acts and their unique acoustic realization call to question whether prominences that convey intensity are distinct from those that convey exclamations more generally, as they seem to be for contrast and givenness, or whether intensity and exclamations constitute a single meaning category. Thus, future studies might compare the realization of prosody in exclamations, exclamatives, and intensifying accents, to better understand their differences.

Lexical intensifiers and other degree modifiers can also convey what I refer to as PRAGMATIC INTENSITY, operating at a metalinguistic level and targeting scales related to precision (Morzycki, 2011; Bochnak & Csipak, 2014), preference (Giannakidou & Stavrou, 2009; Giannakidou & Yoon, 2011), certainty (Romero & Han, 2004), and commitment (Beltrama, 2018). For example, Beltrama (2018) argues that when an intensifier like totally combines with a non-gradable predicate (e.g., I am totally taking a nap later), it conveys a propositional attitude related to the speaker’s commitment towards the utterance. Similarly, Romero and Han (2004) argue that really can operate as an epistemic conversational operator, expressing the speaker’s certainty...
about whether a proposition should be added to the Common Ground—the shared set of background knowledge between the two participants (Stalnaker, 1978). Other measurement expressions such as comparatives (Giannakidou & Stavrou 2009, Giannakidou & Yoon 2011, Morzycki 2011) and hedges (Bochnak & Csipak 2014) have also been shown to operate at the metalinguistic level, and in most cases appear to manipulate a gradable attitude of the speaker towards the appropriateness of the expression (see Krifka, 2008 and Horn, 1985, for similar notions of expression focus and metalinguistic negation, respectively). For example, Beltrama (2016) argues that extreme degree modifiers like flat-out or downright indicate that the chosen form of the expression is superior to all alternative forms.

Prosody has also been argued to target pragmatic scales, especially those related to speaker commitments. Most of these accounts are modeled using Farkas & Bruce’s Table Model (Farkas & Bruce, 2010), which formalizes conversational updates as involving the Question Under Discussion, the individual discourse commitments of the interlocutors, the Common Ground, and projection of future possible Common Grounds. According to this model, both the speaker and the addressee must commit themselves to the truth of a proposition before it can be added to the CG. Therefore, enrichment of the CG happens incrementally, with no single speaker adding directly to it. Rather, speakers add Issues to the Table, defined as a stack of Issues that have been proposed to be added to the CG. Once these issues are ‘resolved’ by the addressee, they then become a part of the CG and are removed from the top of the stack on the Table. Issues are resolved when participants make commitments to the truth of the Issue raised. If no commitments are made, the Issues remain on the Table and do not enter the CG.
Many accounts argue that intonational tunes convey the nature of the speaker’s commitment to an issue. In some proposals, this commitment is binary. For example, Rudin (2022) argues that rising declaratives are distinguished from falling declaratives based only on the presence or absence of full assertive commitment, which is called off by rising intonation. However, other accounts consider commitments to be continuous and gradable. For example, Heim (2019) examines the mapping between sentence-final intonation and a pragmatic scale of commitment, asking how variation in acoustic prominence varies as a function of the degree to which the speaker publicly commits to the truth of a proposition. He argues that the degree of speaker commitment correlates with the duration of the sentence-final intonation, with increased duration conveying stronger commitment. Similarly, Farkas and Roelofsen (2017) argue that degree of commitment varies as a function of sentence type and intonational tune. For example, while falling declaratives express full commitment, polar interrogatives, either falling or rising, express neutrality, and rising declaratives express bias but not a full commitment. Overall, investigations of the use of prosody to convey intensity along pragmatic scales remain limited to scales of commitment, and do not ask about the contribution of specific acoustic dimensions to these interpretations, instead remaining focused on holistic intonation contours and their combination with different sentence types. Future studies might target this question, investigating the realization of prominence that is most typical of pragmatic intensity.

A final potential avenue of exploration involves the interaction of acoustic prominence conveying pragmatic meaning with lexical items. For example, Gotzner and Spalek (2014) and Gotzner, Spalek, and Wartenburger (2013) investigate prosodic prominence in combination with focus operators like only, illustrating that participants are equally likely to derive a contrastive/exhaustive inference when exposed to contrastive L+H* accents as when exposed to
explicit only. Furthermore, the effects of accenting and explicit particles did not seem to be additive. As Beltrama and Trotzke (2019) point out, intensity can be conveyed in a variety of ways, including syntactically via non-canonical word order and through the use of a variety of intensifiers. Future work should look at the use of lexical intensifiers and their combination with prosodic prominence. In other words, one might ask whether producing prominence on an adjective already modified by an intensifier would increase the likelihood of an intensifying interpretation even further, or whether the two cues would cancel one another out.

Overall, the results from this dissertation contribute to the wider discussion about the mapping between prominence and meaning, suggesting that the specific realization of prominence may differ depending on the meaning being conveyed. Furthermore, they suggest that listeners can use these differences in production to draw inferences about a speaker’s intended meaning, suggesting that listeners do not simply attune to whether a word is categorically more prominent, but how specifically that prominence is being achieved. Finally, this dissertation provides initial insights into listeners’ associations of different realizations of prominence with the novel meaning categories of semantic and expressive intensity, suggesting that listeners associate unique realizations of prominence with these two types of meanings compared to the realizations that they associate with contrast.
References


DOI: [http://doi.org/10.1515/9783110874020](http://doi.org/10.1515/9783110874020)


[https://doi.org/10.1093/acprof:oso/9780199249633.003.0002](https://doi.org/10.1093/acprof:oso/9780199249633.003.0002)


DOI: [http://doi.org/10.1080/01690965.2010.504378](http://doi.org/10.1080/01690965.2010.504378)


Appendix A

In this appendix, I discuss the results from the trials involving neutral prosodic profiles from Experiment 2 of Chapter 1. To analyze the likelihood of correctly matching a neutral prosodic profile with a neutral image on the filler trials included in Chapter 1, Experiment 2, two mixed effects logistic regression models were applied to the trials comparing contrastive and neutral images and intensifying and neutral images, respectively. These models predicted the likelihood of correctly matching a neutral prosodic profile to a neutral response from the fixed effects of continuum step, condition (whether the manipulation involved duration or pitch), and the interaction between continuum step and condition.

Figure 25
Predicted Probability of a Match for Neutral Prosodic Profiles

![Predicted Probability of a Match](image)
Note. Predicted probability of matching a neutral prosodic profile with a neutral image when the comparison image was contrastive (left panel) or intensifying (right panel).

As shown in Figure 25, the model results suggest that, when the comparison image was intensifying (right panel), there was a significant effect of continuum step, such that increasing the step significantly decreased the likelihood of a match ($\beta_1=-0.12$, $z=-2.134$, $p<0.05$). There were no other significant effects in the model. A visual inspection of the data suggests that the frequency of a match was overall was very low. When the comparison image was contrastive (right panel), the model results indicate that there was once again a significant effect of continuum step ($\beta_1=-0.29$, $z=-4.655$, $p<.001$), suggesting that, as continuum step increased, the likelihood of a match decreased. On these trials, there was also a significant interaction between continuum step and condition ($\beta_1=0.25$, $z=2.379$, $p=-0.01$), suggesting that, when the manipulation involved duration, the likelihood of a match was not as affected as it was for pitch. In fact, matches tended to increase as pitch increased (meaning that neutral prosodic profiles were more likely to be heard as sounding like they conveyed a neutral meaning), but they tended to decrease as duration increased (meaning that neutral prosodic profiles sounded less neutral and more intensifying the longer that they were). Furthermore, the likelihood of a correct match was overall much higher on these trials than on the trials comparing neutral and contrastive images.

Overall, the findings from the neutral trials suggest that neutral prosodic profiles were not easily mapped onto a meaning that was distinct from contrast, regardless of their pitch scaling or of the duration of the critical adjective’s nucleus. However, they were clearly mapped onto a
meaning that was distinct from intensity, and especially so when their duration was short.
Increasing the duration of the critical adjective’s nucleus seemed to significantly decrease the
likelihood of a match, suggesting that increased duration is especially linked to the meaning
category of intensity.