Using acoustic measures of hyperarticulation to quantify novelty and evaluation in a corpus of political talk shows

Valerie D. Freeman

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Arts

University of Washington

2010

Program Authorized to Offer Degree:

Department of Linguistics

University of Washington
Graduate School

This is to certify that I have examined this copy of a master's thesis by

Valerie D. Freeman

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Committee Members:

Richard Wright

Betsy E. Evans

Date: Aug. 19, 2010

In presenting this thesis in partial fulfillment of the requirements for a master's degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of this thesis is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Any other reproduction for any purposes or by any means shall not be allowed without my written permission.

Signature ______________________

Date Aug. 20, 2010

University of Washington

Using acoustic measures of hyperarticulation to quantify novelty and evaluation in a corpus of political talk shows

Valerie D. Freeman

Chair of the Supervisory Committee:
Dr. Richard Wright, Associate Professor
Department of Linguistics

Abstract

This study analyzes an episode of a televised political talk show, *Tucker*, for evidence that speakers hyperarticulate concepts about which they express attitudes, a use of hyperarticulation that interacts with the discourse function of signaling new information. Using content analysis, the utterances of five speakers were coded on two dimensions: Evaluation (presence or absence of attitude-expression) and Novelty (new or given information in the discourse segment). To compare the resulting groups, four measures of hyperarticulation were used: speaking rate of the phrases, and duration, pitch excursion, and vowel space expansion (first and second formant values) of stressed vowels in the phrases. Group results show reliable effects for both dimensions and an interaction such that Evaluation has a greater effect than Novelty overall. Attitude-expressing items are hyperarticulated compared to a control group of neutral phrases, and within each group, new information is hyperarticulated compared to given information. Speaking rate and vowel space expansion showed these effects most reliably, followed by vowel duration. Pitch excursion was not a reliable indicator of either dimension. Individual variation contributed to the group results for all four measures, displaying very different patterns between speakers. These findings provide acoustic correlates to attitude-expression, which previous studies have not explicitly investigated and which can be applied in future work on the identification of specific types of attitudes, the perception of attitude-expression, and automatic speech recognition and synthesis.

CONTENTS

List of Figures

List of Tables

Acknowledgements

I am grateful to my advisors, Dr. Richard Wright and Dr. Betsy Evans, without whose guidance, advice, and attention this work would not have been possible. Thank you for being so available, pleasant, and easy to work with, all the way to the end.

Thanks to Hil Lyons for his invaluable statistical advice and patient explanations; to Dr. Emily Bender, Dr. Mari Ostendorf, Brian Hutchinson, and Meghan Oxley for their assistance with the corpus data; and to Chris Martin for his careful proofreading of this manuscript.

To my family: thank you for your confidence in me, for finding my study interesting, and for being impressed with my accomplishments.

Mark Ellison, thank you for your encouragement, understanding, and moral support all along the way. Thank you for helping me through the stressful times and not adding to them, for being a good sounding board and offering good advice, and for knowing when to let me work and when I need a break.

Broadcast audio used with permission from a Linguistic Data Consortium (LDC) corpus produced for the DARPA Global Autonomous Language Exploitation (GALE) project.

Transcript and annotation data provided by the Linguistic Cues of Roles in Conversational Exchanges (LiCORICE) Project, funded by NSF grant IIS-0811210, and by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), Contract No. W911NF-09-C-0131. All statements of fact, opinion or conclusions contained herein are those of the author and should not be construed as representing the official views or policies of IARPA, the ODNI or the U.S. Government.

Chapter 1: Introduction

How do speakers make their attitudes known without explicitly stating them? Changes in intonation contours have been identified as signaling emotion and attitude (Wichmann 2002b), but there may be other cues, including hyperarticulation – slow, distinct, exaggerated, or clear speech (Lindblom 1990). Hyperarticulation has a number of known purposes, but its use as an indicator of attitude has not been explicitly studied. This study seeks to determine whether hyperarticulation is indeed a means to attitude expression, and if so, how this use interacts with the discourse function of signaling new information.

1.1 Hyperarticulation, Novelty, and Evaluation

The first time a term is introduced in a discourse segment, it is considered *new information* and is often hyperarticulated (Aylett 2005; Aylett & Turk 2004), perhaps to increase listener comprehension (Whalen et al. 2004), but after that point, it is understood by both speaker and listener as "on the counter" or *given information* (Prince 1981:235), and its pronunciation can be reduced (said more quickly, with less precision and more contracted, or "centralized" vowels) (Aylett & Turk 2004; Fowler & Housum 1987; Tomita 2008). In this study, this new-given dimension is referred to as Novelty.

On another dimension, speakers may convey their attitudes about a concept by continuing to hyperarticulate it throughout a conversation, rather than reducing their pronunciation as expected for given information. Hyperarticulated material may be intended to attract listeners' attention; if used to express an attitude, repeatedly hyperarticulating a concept sends the message: "I have an attitude about this, and I'll keep reminding you of it every time I mention the concept." In discourse- and conversation analysis, various terms are used to describe speakers' attitudes and opinions (e.g., stance, attitudinal or affective stance) and their expression (e.g., stancetaking, evaluation); this study refers to the dimension of attitude expression as Evaluation (described in detail in section 2.3).

While a number of studies have considered the effect of emotions or moods (e.g., happiness, sadness, anger, fear) on articulation (see Caelen-Haumont & Zei Pollermann 2008 for a summary), not as much has been done to identify the acoustic effects of attitude (Uldall 1960; Wichmann 2002a, 2002b). This study begins to fill in this gap by examining an episode of the political television talk show *Tucker* for evidence that speakers use hyperarticulation to indicate they have attitudes about a concept. Furthermore, this study proposes that by using hyperarticulation to convey their attitudes, speakers supersede the discourse convention of reducing the pronunciation of repeated information.

Chapter 2: Related Literature

2.1 New vs. given

The *new* vs. *given* distinction as a discourse function relies on Grice's (1967) Cooperative Principle, the social agreement under which speakers try – and are expected – to give true, concise, and relevant information. As a practical component, Clark & Haviland (1977) added a "given-new contract" to describe how speakers model their presentation of information based on what they believe their listeners already have in mind:

> "The speaker tries…to make the structure of his utterances congruent with this knowledge of the listener's mental world. He agrees to convey information he thinks the listener already knows as given information and to convey information he thinks the listener doesn't yet know as new information." (p. 4)

Since "new" and "given" are relative terms, it is important to define them carefully. Prince (1981) offers a detailed taxonomy from which the broadest divisions are used for this study. Simply put, an entity is *new* when it is first introduced in a discourse segment, or "put on the counter" (p. 235), or when it is reintroduced after a change in the topic of conversation (to continue the metaphor, it is put back on the counter after being pushed aside by other topics). *Given* information is everything already on the counter, or what the speaker assumes is in the listener's mental discourse model, whether overtly mentioned previously or inferable from logic or context.

Clark & Haviland (1977) detail how the new-given distinction is made syntactically, with each piece of given information having only one antecedent in recent discourse, but there are other ways to make the distinction, including hyperarticulation and reduction. Several studies (Aylett & Turk 2004; Chafe 1974; Fowler & Housum 1987; Tomita 2008) have shown that items that are repeated, predictable, or familiar in the discourse context have shorter durations, more contracted vowels, and weaker stress – in other words, given information is reduced (hypoarticulated) without sacrificing listener comprehension. A few (Aylett & Turk 2004; Bolinger 1963; Tomita 2007) have also found that infrequent and unpredictable items have longer durations and expanded vowel spaces – in short, new information is hyperarticulated.

2.2 Hyperarticulation

Before further discussing the usefulness of hyperarticulation, it is helpful to describe its physical properties. The goal of hyperarticulating is to "exaggerate speech sounds" (Whalen et al. 2004:155) or make phonetic and phonological components more distinct (de Jong, Beckman, & Edwards 1993).

This is accomplished in several ways. As mentioned above, lengthening is a prevalent factor (Aylett & Turk 2004; Soltau & Waibel 2000; Wennerstrom 2001), and expanded vowel space also plays a key role (Tomita 2007). Hyperarticulated items receive heavier stress (Aylett 2005; de Jong et al. 1993), undergo changes in pitch contour and fundamental frequency (Soltau & Waibel 2000), experience less coarticulation, and have more precise places of articulation (de Jong et al. 1993).

While only a few studies have specified hyperarticulation as a means for signaling new information (Aylett 2005; Aylett & Turk 2004), several other uses are commonly discussed, beginning with aiding listener comprehension. As Whalen et al. (2004:156) state, “articulation is made in the most extreme manner possible with the intention that perceptual confusions will be at a minimum.” Similarly, Soltau & Waibel (2000:1779) found that “hyperarticulation often occur[s] as a strategy to recover previous recognition errors” – that is, speakers hyperarticulate when they need to repeat items in order to correct listeners’ misunderstandings. Other studies identify hyperarticulation as a contrastive device (Chafe 1974; de Jong et al. 1993) that even preschool-age children employ (Hornby & Hass 1970). Finally, several researchers have noticed that hyperarticulation is used to signal something important: to “help the listener recognize an important word” (de Jong et al. 1993:210), “give [the] impression of focus” (Aylett 2005:2521), or “bring particular story events to the foreground” (Wennerstrom 2001:200). The reasons speakers choose to emphasize certain items as “something important” vary with context, but one possibility is that speakers use hyperarticulation as a means to express their attitudes.

2.3 Attitude, stance, and evaluation

Across various disciplines, more than one term is used to describe attitudes and their expression. In this study, *attitudes* are treated as in social psychology: “beliefs and opinions which motivate or explain behavior” (Wichmann 2002a:11), not the behavior itself, as is often the case in studies of intonation. In discourse- and conversation analysis, *stance* can be used as a general umbrella term for “personal feelings, attitudes, value judgments, or assessments” (Biber et al. 1999:966). Conrad & Biber (2000:57) divide stance into three categories:

(1) “epistemic stance, commenting on the certainty (or doubt), reliability, or limitations of a proposition, including comments on the source of information;
(2) attitudinal stance, conveying the speaker’s attitudes, feelings, or value judgements;
(3) style stance, describing the manner in which the information is being presented.”

While epistemic stance was included in the coding scheme for this project, the main focus is on attitudinal stance, a concept which is also labeled differently by various researchers. Ogden (2006:1753) refers to attitudinal stance as "how speakers convey their attitude towards what is being said," while others make a distinction between the attitudes and their expression. Haddington (2004:101) defines stance as "speakers' subjective attitudes toward something" but their expression as the social activity of *stancetaking*. Du Bois (2007) describes stancetaking as a three-part act which includes evaluation of an object or proposition, positioning of a speaker in relation to that evaluation, and alignment between two speakers and their evaluations of the same object. The focus of this paper is *evaluation*, "perhaps the most salient and widely recognized form of stancetaking" (Du Bois 2007:142). 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" (Hunston & Thompson 2000:5). This paper uses "Evaluation" specifically to label the dimension of attitude-expression. The process of coding speakers' utterances centers on the question, "Is the speaker expressing an attitude about this concept?" As very little research has been done on acoustic correlates of attitude-expression, this project is a first step with a scope that does not allow any attempt to label what *kind* of attitude is being expressed but only whether *some* attitude is clearly present. Thus, for the dimension of Evaluation, attitude-expressing phrases are coded as *attitude*, and neutral expressions as *control*. (These coding procedures are detailed in section 3.1.2.2 in the next chapter.)

2.4 Televised discussion

Conversations from a televised political talk show were chosen for this study because they contain spontaneous speech that was recorded for a purpose other than linguistic analysis. An important issue to consider is how behavior on talk shows may differ from that in other settings. Haddington (2004, 2007) summarizes the differences many analysts have found between news interviews and everyday conversation. Speakers' turns are longer, being comprised of several turn constructional units (TCUs), between which other speakers are expected to respond in everyday conversations but not during interviews. Turns are also constructed more rigidly, with hosts asking questions and guests answering them. However, this rigidity may not apply as much on political talk shows as it does with news interviews in general. Each program has its own format, with some adhering to structured interviews, but many prefer a looser framework that allows guests to proceed without prompting and introduce or shift topics without deference to the host. The program analyzed in this study, *Tucker*, is an example of the latter, making it a good candidate for studying unscripted discussion.

In talk shows of all formats, the host usually plays the role of moderator, directing the flow of conversation by introducing topics, asking loaded questions, encouraging certain guests to respond, and cutting off the discussion at predetermined times (e.g., for commercial breaks). Although this may cause hosts to behave differently than other participants, this sort of moderated group discussion is not unique to television talk shows. In unobserved settings such as classrooms and business meetings, a teacher or designated leader asks questions and controls the flow of discourse. A similar dynamic is found in group interviews recorded by researchers in the field, who often act as moderators. Such field interviews are similar in nature to talk shows in that speakers know they are being observed and may therefore behave differently than they do in other situations. In public interviews, speakers "display a tacit orientation to the overhearing audience" (Haddington 2004:123) which likely leads them to give more background information and elaborate their thoughts more than necessary to communicate with their fellow commentators alone, who have likely discussed similar issues with them before (e.g., on previous episodes). However, even if speakers are performing for viewers, the group interaction on talk shows is unscripted and spontaneous, and speakers overlap, interrupt, argue, persuade, and deliver opinions just as in everyday talk.

Overall, talk shows which focus on group discussion (rather than a rigid question-and-answer format) are suitable for the purposes of measuring hyperarticulation. Speakers on television can be expected to make an effort to speak "clearly," but this does not imply they hyperarticulate everything. Some items are emphasized and others reduced for the same social and discourse functions as in any similar situation.

2.5 Summary

Cooperative speakers try to highlight new information in a conversation. One way to do this is via hyperarticulation, which can include lengthening, variation in pitch, and vowel space expansion. Several discursive uses have been identified for hyperarticulation; this study proposes that it is also used to call attention to speakers' attitudes. Following Grice's Cooperative Principle, and with the given-new contract in mind, a cooperative speaker's hyperarticulation of given information as well as new must be interpreted not as a violation of the contract but as a signal of some other meaning the speaker intends the listener to grasp. If this is so, an interaction can be expected between the effects of hyperarticulation used to signal Novelty and those used to express attitudes. Political talk shows are suitable for examining hyperarticulation in conversation because their discussion format involves spontaneous speech between multiple participants.

Chapter 3: Procedures and Results

This study uses four measures of hyperarticulation (speech rate, vowel duration, pitch excursion, and vowel space expansion) to examine the interaction of two discourse-functional dimensions: Novelty (whether a phrase is new or given (repeated) information in a conversation) and Evaluation (whether the speaker expresses an attitude about a concept). Utterances taken from conversational segments of a televised political talk show were divided on the Evaluation dimension into *attitude* and *control* groups and on the Novelty dimension into *new* and *given*. Three experiments were conducted to compare the groups using different measures of hyperarticulation. This chapter details the procedures followed for the entire study, including the methods and results of each experiment.

The procedures were divided into two main processes: in Phase I the data set was selected and coded, and in Phase II, the resulting tokens were measured and compared to determine the effects Novelty and Evaluation on hyperarticulation. This chapter describes both phases. For Phase I, section 3.1.1 discusses how the episode and segments of *Tucker* were chosen, and section 3.1.2 explains the procedures for selecting and coding tokens. Section 3.1.3 presents the results of Phase I: the speakers and the numbers and types of tokens identified for them during the coding procedures. In Phase II, section 3.2.1 states general hypotheses and some possible outcomes of the experiments should the data support them, section 3.2.2 outlines general methods used throughout the phase, and section 3.2.3 details each of the three experiments performed, including the specific hypotheses, methods, and results for each measure of hyperarticulation (speaking rate and vowel duration, pitch excursion, and formant values).

3.1 Phase I: Selecting and coding the data set

The data set was selected in Phase I of conducting this study. An episode of the television program *Tucker* was chosen from a corpus of political talk shows, and concepts and tokens were identified for analysis from within conversational segments. These were then coded on the dimensions of Novelty (to indicate whether the information was *new* or *given* in the conversation) and Evaluation (to denote the presence (*attitude*) or absence (*control*) of attitude-expression).

3.1.1 Program and segments

To choose the talk show for this study, five episodes were randomly selected from a corpus of the audio tracks of televised political talk shows (Linguistic Data Consortium 2009). Since conversation is the context to be studied, two episodes were eliminated due to the large proportion of non-conversational reporting they contained. From the remaining three, one was randomly selected: an episode of MSNBC's *Tucker*, dated April 10, 2007 (Geist 2007). The segments (defined as broadcast between commercials) which consisted of conversations with at least one guest were analyzed – reports, announcements, introductions, and clips to prompt discussion were not analyzed. No conversational segment contained highly emotional speech, such as shouting or crying, which would have been excluded for this study, since emotional expression can affect articulation (for a summary of work on vocal cues to emotion, see Caelen-Haumont & Zei Pollermann 2008). The following section describes the process of selecting and coding the words and phrases to be analyzed acoustically in Phase II.

3.1.2 Coding

This section describes the methods for selecting tokens to analyze, followed by the procedures for coding them in terms of attitude expression (Evaluation) and as new or given information (Novelty).

3.1.2.1 Concepts and tokens

All content words (i.e., not "function" words such as articles, connectors, etc.) and phrases containing content words that were repeated at least three times by the same speaker in conversational portions of a segment were selected for analysis. Having three or more repetitions eases the ability to recognize the continued expression of attitude, if present, and lends more power to any pattern differentiating attitude-expressing utterances from neutral utterances or new from given information. Each repetition of lexically identical material (a word or phrase) is a *token*. A group of lexically identical tokens and all references to them (e.g., pronouns, synonyms, etc.) are taken together to form a *concept*. The following example, illustrated in Figure 1, clarifies the terminology. In a four-minute interview, one guest, Eli Pariser, referred to "the war in Iraq" more than three times, so the *concept* of "the war in Iraq" was selected for analysis. Figure 1 displays some examples of the references he made to this concept, including truncations ("the war," "this critical issue of [the war in] Iraq") and pronouns ("it, this"). The only portion common to at least three references is the word "war," which was repeated five times. Thus, each of Eli's repetitions of the word "war" is a token to be measured. However, not every repetition of the word "war" necessarily pertains to the concept of "the war in Iraq." For example, references to "the war in Afghanistan, the war on terror, the Vietnam war, culture wars, drug wars," etc. would form their own concepts, separate from "the war in Iraq." This ensures that speakers' attitudes about different concepts are not lumped together. Eli, for example, expressed very strong attitudes about ending the war in Iraq, but his attitudes about the war in Afghanistan might center on something completely different, which could affect his pronunciation of the word "war" in reference to Afghanistan as opposed to Iraq.

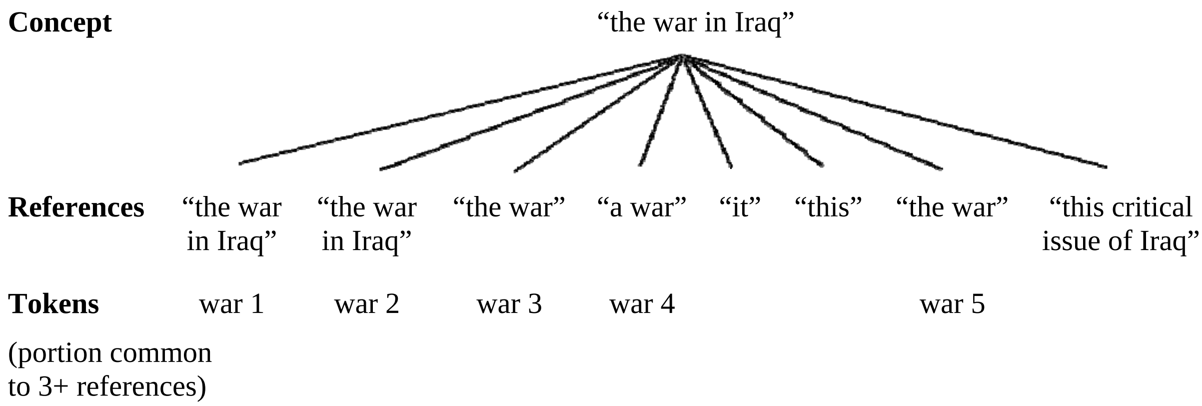


Figure 1. The relationship between concepts, references, and tokens. References and tokens are components of a concept. All are considered together when coding Evaluation, and the resulting code (*attitude* or *control*) applies to the whole concept, but each token is coded separately for Novelty (*new/given*).

Inflectional variations (singular/plural, tenses) were grouped together and considered a single concept, as long as the endings did not shift the word stress or cause a significant difference in meaning. For example, in another segment, "defend, defends, defending" were all considered repetitions of the same concept, but only the portion common to all variants ("defend") was measured (resulting in three "defend" tokens).

As described in the procedures below (section 3.1.2.2), all references to a concept were considered together when coding for Evaluation, so the resulting code (*attitude* or *control*) applies to the whole concept and every member token. Using content analysis to separate concepts into attitude-expressing and control groups makes it possible to address the central proposal of this study, that speakers continue to hyperarticulate concepts about which they express attitudes (as compared to neutral concepts), rather than reducing repeated material as expected for given information (as they should do for neutral/control concepts). To be able to compare *attitude* and *control* concepts on their treatment of new and given information, each token was coded separately as *new* or *given*, using procedures also described below (section 3.1.2.3).

3.1.2.2 Marking Evaluation *(attitude* vs. *control)*

The following content analysis was performed to identify speakers' expressions of attitude regarding each concept. All references a speaker made to the concept, including truncations, abbreviations, ellipses, and pronouns, were considered evidence. One point was given to a concept for each instance in which the speaker performed an action falling under the categories listed below. (The full Evaluation coding protocol is presented in Appendix 1.) Each category contains several actions, so a concept could receive several points per category, and a phrase used as evidence could perform more than one action, possibly in more than one category. To calculate a score for the concept, the number of items of supporting evidence was divided by the number of tokens (repetitions of the concept word/phrase). If the ratio was 2.00 or higher, the concept was marked *attitude*; otherwise, it was marked *control*. This ratio was determined by creating a frequency distribution of all concept scores for all speakers. The distribution was nearly normal, with a mean of 1.92.

Categories of actions used to organize speaker behavior when coding Evaluation

(See Appendix 1 for the complete protocol.)

a. Introduction: Speaker works to keep the topic in play by introducing or returning to it (cf. Labov 1972; Prince 1981).
b. Overt evaluation: Speaker takes a stance by stating an opinion or making a prediction, including what "probably" happened or will happen (cf. Conrad & Biber 2000; Du Bois 2007; Hunston & Thompson 2000).
c. Evaluative description: Speaker uses evaluative modifiers or commentary with a token or when referring to the concept (cf. Hunston & Thompson 2000; Labov 1972).
d. Credibility: Speaker offers support for a stance by expressing certainty, citing experts, personal credentials or experience, or presenting the stance as fact (cf. Biber & Finegan 1989; Conrad & Biber 2000; Hunston & Thompson 2000).
e. Persuasion and recommendation: Speaker attempts persuasion, makes a suggestion or recommendation (e.g., what "should" be done) (cf. Conrad & Biber 2000).
f. Agreement: Speaker agrees or disagrees with another speaker (cf. Conrad & Biber 2000; Du Bois 2007).

3.1.2.3 Marking Novelty *(new* vs. *given)*

As previously mentioned, concepts were selected for analysis when repeated three or more times by the same speaker within a conversational portion of a segment. Following Prince's (1981) taxonomy, the first utterance of a concept in a segment was coded as *new*, even if said in a non-conversational portion or by another speaker. This indicates the instance that introduces the concept into the discourse, making it fresh in all participants' minds. If the first utterance of a concept for one speaker is said in a non-conversational portion and/or by another speaker, it was not measurable as a token, but it was labeled as *new* so that no other repetition would be erroneously marked as *new*. For example, in one segment, the host, Tucker Carlson, introduces the topic to be discussed by reporting that a Democratic primary debate to be held by Fox News was "[effectively] wiped out" because the three major candidates were not going to participate. He said the word "debate" twice in his introduction and again in his first question to the other commentators, but never again in that segment. Although he said "debate" three times, only once was during a conversational portion, so *debate* was not used as a concept for Tucker. However, in a response to Tucker's question about the benefits of refusing to join the debate, the commentator Eugene Robinson repeated "debate" three times, making *debate* an analyzable concept for him. Even so, because it was Tucker who introduced *debate* into the discourse, none of Eugene's repetitions of the concept could be coded as *new*. If Tucker had continued to repeat "debate" within the conversation (making it an analyzable concept for him), he would also have no tokens coded as *new* because the first repetition of "debate" was uttered in a non-conversational portion.

If a concept was dropped in the discussion but later picked up again, the reintroduction was coded as *new*. A concept was considered reintroduced when it followed at least five speaker-turns which spanned 60 seconds or more and did not include any reference to the concept in question. This ensured that the concept was no longer "on the counter" in the discourse (Prince 1981), and a reintroduction was necessary to bring it back to the forefront in speakers' minds.

Finally, all tokens not coded as *new* were marked *given*.

3.1.2.4 Summary

A word or phrase was selected as a *concept* for analysis if a speaker repeated it three or more times within conversational portions of a broadcast segment. Each concept was coded for Evaluation (attitude expression) via content analysis, which resulted in the whole concept (and all of its repetitions) being labeled as *attitude* or *control*. Then, each repetition of the concept, a *token,* was coded for Novelty, i.e., labeled as *new* when introduced or reintroduced within a conversation, or as *given* when subsequently repeated. When the two categories of labels were combined, four types of tokens emerged: *new-attitude, new-control, given-attitude,* and *given-control*.

3.1.3 Resulting Data Set

This section presents the results of the selection and coding processes carried out in Phase I. A brief characterization of the broadcast segments used and each speaker's role is followed by a breakdown of how many tokens fell into each category (*new, given, attitude, control* and combinations thereof) and the proportion each speaker contributed to each combination.

3.1.3.1 Speakers

As mentioned in section 3.1.1, only conversational segments with at least two interlocutors were analyzed. For the episode of *Tucker*, this amounted to six segments with a total of five male speakers. Two of these were interviews approximately four minutes long between the host, Tucker Carlson (age 37, from California) and an expert guest: the executive director of MoveOn.org, Eli Pariser (age 26, from Maine) discussed his organization's hosting of Democratic primary debates, and the chairman of the Center for the Rule of Law, also Dean Emeritus of Boston University Law School, Ron Cass (approximate age: 50s, from Virginia) discussed legal issues surrounding the paternity of the late Anna Nicole Smith's child (Eli Pariser, n.d.; Honorable Ronald A. Cass, n.d.; Tucker Carlson 2010). The other four segments ranged in length from three to ten minutes and included the host (again, Tucker Carlson) and two regular contributors to the show: conservative political commentator Pat Buchanan (age 68, from Washington, DC) and *Washington Post* columnist Eugene Robinson (age 52, from South Carolina) (Eugene Robinson, n.d.; Pat Buchanan 2009). Because they call each other by first name during the program, they are referred to by first name here, as well.

3.1.3.2 Concepts, tokens, and vowels

After the coding procedures from section 3.1.2 were completed, a total of 65 concepts were identified: 32 *attitude* and 33 *control*, which broke down to 109 *attitude* and 109 *control* tokens (repetitions of the concept word/phrase), a total of 218 tokens analyzed. Because some token phrases contained more than one stressed vowel, the total number of vowels measured was higher: 112 *attitude* and 125 *control*, for a total of 237. Table 1 summarizes these totals. When each token was coded as *new* or *given*, the *attitude* tokens broke down into 36 *new* and 73 *given* (37 *new* and 75 *given* vowels), and the *control* tokens divided into 27 *new* and 82 *given* (31 *new* and 94 *given* vowels). Table 2 displays these totals. Overall, the sample is well balanced: there are a similar number of *attitude* and *control* items in each category, and there are two to three times as many *given* as *new* tokens, which is expected since a token is likely to be repeated more frequently as given information than to be introduced as new.

Table 1. Total number of concepts, tokens, and vowels coded as *attitude* or *control*.

Token type	Concepts	Tokens [1]	Vowels [2]
Attitude	32	109	112
Control	33	109	125
Total	**65**	**218**	**237**

[1] Tokens are repetitions of the concept word/phrase.
[2] Some token phrases contained multiple measurable (stressed) vowels.

Table 2. Total number of *attitude* and *control* tokens and vowels coded as *new* or *given*.

Token type	Tokens			Vowels [1]		
	New	***Given***	**Total**	***New***	***Given***	**Total**
Attitude	36	73	**109**	37	75	**112**
Control	27	82	**109**	31	94	**125**
Total	**63**	**155**	**218**	**68**	**169**	**237**

[1] Some token phrases contained multiple measurable (stressed) vowels.

Although the totals for all speakers combined are well balanced, not all speakers contributed to each cell equally. Figure 2 is a series of pie charts illustrating the proportions each speaker's contribution of each combination of *new/given* and *attitude*/*control* tokens. For *new-attitude* tokens, Tucker contributed 6%, Pat 33%, Eugene 9%, Ron 37%, and Eli 15%. For *given-attitude* tokens, Tucker contributed 6% of the total, Pat 49%, Eugene 9%, Ron 20%, and Eli 16%. Tucker contributed 38% of *given-control* tokens, Pat 19%, Eugene 8%, Ron 23%, and Eli 12%. Finally, Tucker contributed 44% of *new-control* tokens, Pat 30%, Ron 15%, and Eli 11%; Eugene had no tokens in this category. To summarize, the host, Tucker, contributed the largest proportion of both *new-* and *given-control* tokens but the smallest proportion of both *new-* and *given-attitude* tokens. Pat and Ron varied in their contributions to each cell, while Eli was fairly constant. Eugene contributed a small amount in each cell except *new-control*: he had no tokens in this cell. For this reason, his data had to be removed from measures that compare data from all four cells.

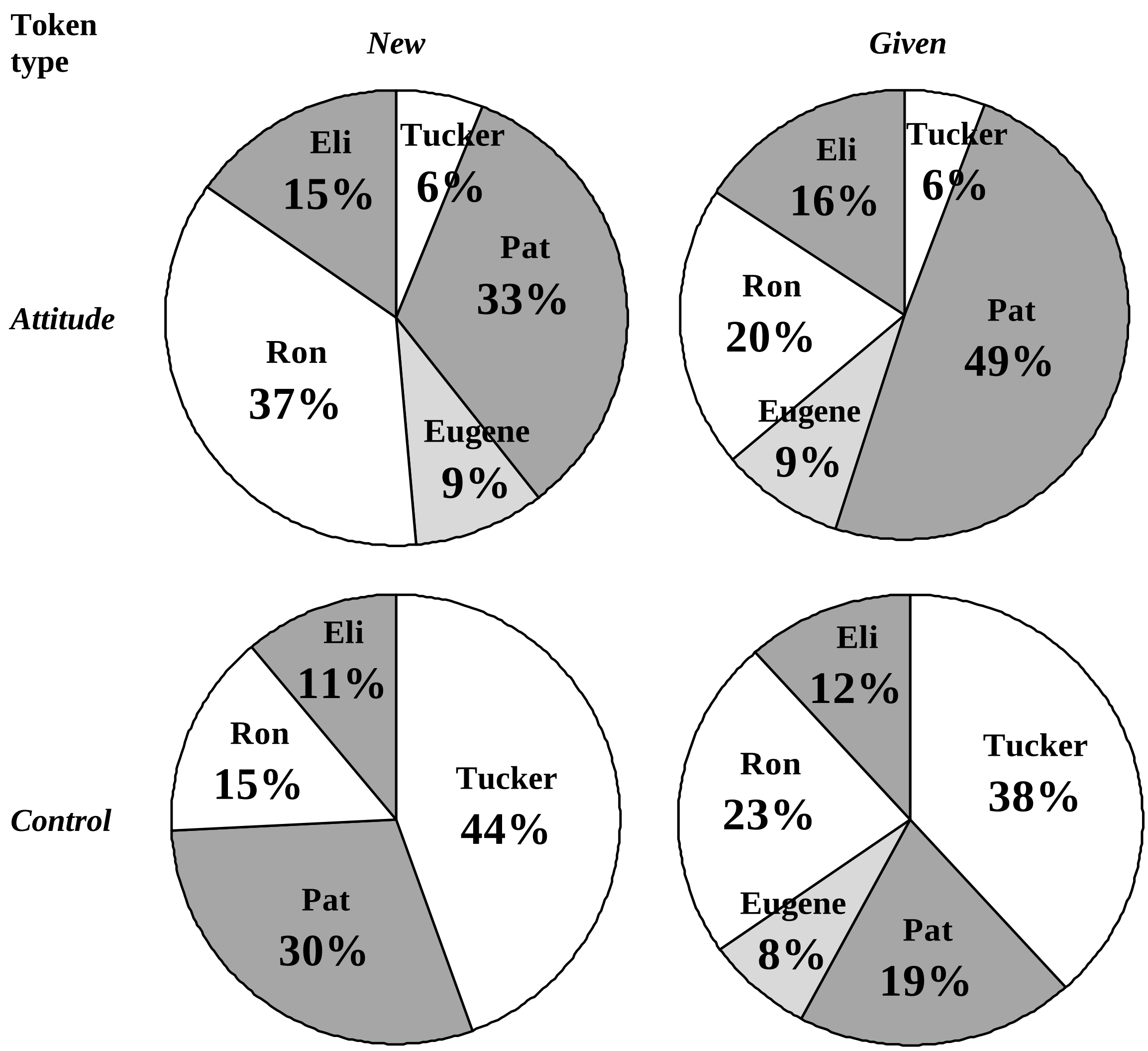


Figure 2. Proportion of each type of token contributed by each speaker.

3.1.3.3 Summary

The data set was drawn from six broadcast segments of an April 2007 episode of the televised political talk show *Tucker* involving conversation between a total of five male speakers. A total of 65 concepts were identified, which amounted to 218 tokens containing 237 measurable vowels. One-quarter to one-third of tokens were coded as *new* information, the rest as *given*, and there were roughly equal numbers of *attitude* and *control* concepts, tokens, and vowels. However, the proportion of each token type contributed by each speaker varied; notably, the host contributed a much larger proportion of *control* than *attitude* tokens, and one speaker did not have any *new-control* tokens.

3.2 Phase II: Measurements

Phase I of conducting this study involved selecting conversational portions of an episode of the political talk show *Tucker*, from which tokens were identified for analysis and coded on the dimensions of Novelty (indicating whether the information was *new* or *given* in the conversation) and Evaluation (denoting the presence (*attitude*) or absence (*control*) of attitude-expression). Phase II states the hypotheses and methods for measuring the data set selected in Phase I.

3.2.1 General Hypotheses

This study aims to determine how Novelty and Evaluation interact as effects on hyperarticulation, as measured by speech rate, vowel duration, pitch excursion, and formant values of stressed vowels. Three general hypotheses were tested for each measure of hyperarticulation:

H1: There is a reliable effect for Novelty. As in previous studies (Aylett 2005; Aylett & Turk 2004), speakers are expected to hyperarticulate new information.

H2: There is a reliable effect for Evaluation. Speakers are expected to hyperarticulate words and phrases about which they express attitudes (*attitude* tokens) along one or more dimensions of the acoustic signal, relative to their articulation of phrases about which they do not express attitudes (*control* tokens).

H3: There is an interaction between Novelty and Evaluation. The interaction may be additive, or some types of tokens may be affected more than others by one or both dimensions. It is predicted that Evaluation will have a greater effect than Novelty overall, such that all *attitude* tokens will be hyperarticulated compared to all of their *control* counterparts. Individual variation is also expected in that speakers may employ each measure of hyperarticulation to different degrees or in different directions.

If all three hypotheses are supported, many different interactions are possible. Each combination of *new/given* and *attitude/control* tokens might be articulated to a different degree. For example, Novelty may have a greater effect for *attitude* than *control* tokens, or Evaluation may affect *new* more than *given* tokens, or both, resulting in a combination where *new-attitude* tokens are hyperarticulated more than any other type of token.

3.2.2 General Methods

To test the above hypotheses, three experiments were performed, each using a different measurement of hyperarticulation. This section describes the methods used to prepare for all three experiments, which are detailed individually in section 3.2.3.

3.2.2.1 Measures chosen

One goal of this study is to establish that speakers' expression of their attitudes have measurable acoustic correlates. With the hypothesis that attitudes are signaled by some form of hyperarticulation, several acoustic measures become candidates for testing the hypothesis. As previously mentioned, lengthening is a prominent feature of hyperarticulated speech which is also easy to measure. While intonation is considered an important cue to underlying mental positions, matching specific intonational patterns to easily labeled attitudes is difficult (Uldall 1960; Wichmann 2002a, 2002b). Pitch excursion (the amount a pitch deviates from a speaker's baseline mean or median pitch) is a coarser measure, but it can be used to detect whether pitch is responsive to attitude expression. Changes in loudness are also linked to hyperarticulation, but intensity could not be used in this study because factors such as microphone distance and background noise were not explicitly controlled in the recordings. Vowel space is often used to define the continuum between reduced (contracted) and hyperarticulated (expanded). Precision of consonant place of articulation is also a feature of hyperarticulated speech, but it can be more difficult to measure from audio recordings than vowel formants. Of these possible measures of hyperarticulation, this study used speech rate over a word/phrase, and the duration, pitch excursion, and changes in first and second formant values (F1, F2) of stressed vowels within the word or phrase. Stressed vowels were chosen as the focus for measurement because they carry useful information: they are targets for lengthening in English, they often carry the most prominent pitch in a word, and their formant structure reflects degrees of reduction in an easily measured way.

3.2.2.2 Vowel measuring procedures

Measurements were taken using a script in *Praat* (Boersma & Weenink 2008). Vowel formants were measured using *Praat*'s LPC (linear predictive coding) formant tracker with the following settings: a formant range of 0-5500 Hz with a window length of 25 ms, dynamic range of 30 dB, and 14 formant coefficients (6 formants). Pitch was measured using *Praat*'s autocorrelation algorithm with a pitch range of 60-500 Hz. The beginning and end of all tokens were marked manually on one text tier in

Praat's annotation interface, and the stressed vowels within each token phrase were marked manually on another tier. Methods for marking vowel onsets and offsets are detailed below. The script recorded the following measurements: the duration in seconds of each token and stressed vowel, the pitch in Hertz at vowel midpoint, and the first and second formants in Hertz at vowel midpoint.

3.2.2.2.1 Before/after stops

The onsets of vowels after stops were marked just after the consonant release burst, as indicated by increased energy throughout the range of the spectrogram and/or a short period of increased amplitude in the waveform. In Figure 3, dark vertical strips are visible in the spectrogram for the release bursts of initial /t/ and following /k/, which are also accompanied by increased amplitude in the waveform. The vowel begins directly after the burst of /t/; the aspiration that follows is included as part of the vowel duration so that vowels following voiced and voiceless stops are measured in the same way, as can be seen for the vowel following a voiced /b/ in Figure 4. The offsets of vowels before stops are marked at the consonant closure, as indicated by the loss of energy of F2 and/or a drop in amplitude in the waveform. Both indicators are visible in the offset before /k/ marked in Figure 3.

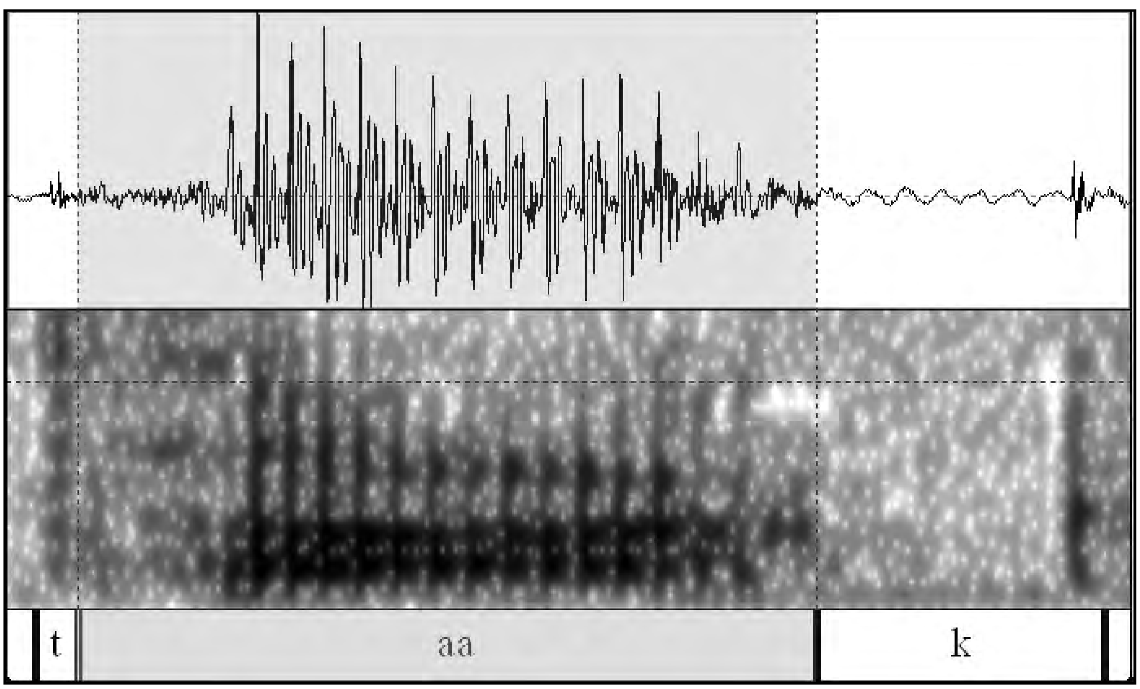


Figure 3. Measuring the vowel of "talk" [tɑk] said by Eli. Vowel begins after the release burst of the initial consonant /t/ and ends at the closure of the closing consonant, /k/, indicated by the loss of energy in F2 and a drop in amplitude in the waveform.

3.2.2.2.2 Before/after nasals

Vowels adjacent to nasals were marked just inside the depression in F2 that indicates the nasal consonant, which coincides with a less dense, less complex waveform. Both features are visible in Figure 4, where a vowel precedes /m/.

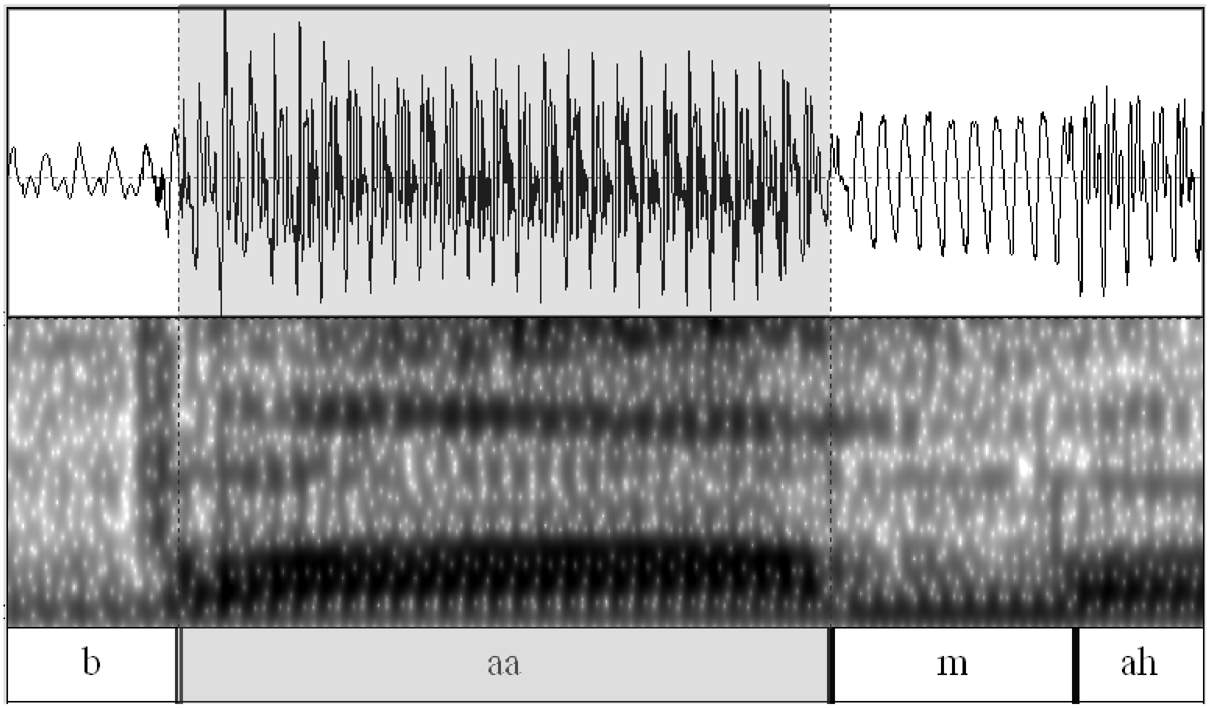


Figure 4. Measuring the stressed vowel of "Obama" [ˈbɑː.mə] said by Eli. Vowel begins after the release burst of the initial consonant /b/ and ends at the loss of energy in F2 before the nasal /m/, which is also indicated by a less dense waveform.

3.2.2.2.3 Before/after fricatives

Vowels neighboring fricatives were marked just up to but not including the frication noise, as indicated by energy in high frequencies and a dense, medium-amplitude waveform. Figure 5 illustrates marking a vowel before /s/.

3.2.2.2.4 Before/after approximants

Vowels between glides and liquids often relied heavily on changes in waveform in combination with changes in F2 and/or F3. The rhotic /ɹ/ is characterized by a lower-amplitude, less complex waveform and a dip in F3 which sometimes leaves an "empty bowl" shape, as in Figure 5. Rhoticization often carries over to nearby vowels; to avoid the majority of this coarticulation, vowels adjacent to rhotics were marked at a point where the slope of F3 flattens, coinciding with lower

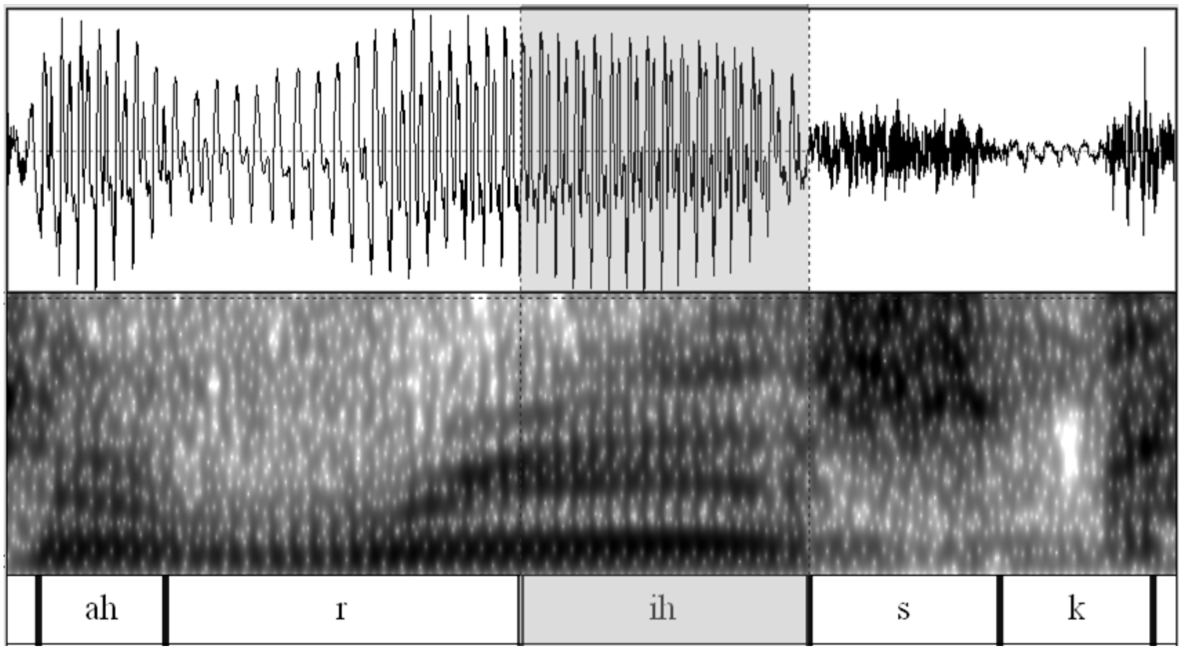


Figure 5. Measuring the stressed vowel of "a risk" [əˈɹɪsk] said by Pat. Vowel begins as F3 finishes its rise out of the rhotic approximant /ɹ/, which coincides with a reduction in the amplitude of the waveform. Vowel ends before the frication of /s/, indicated by dense energy in high frequencies and a dense, medium-amplitude waveform.

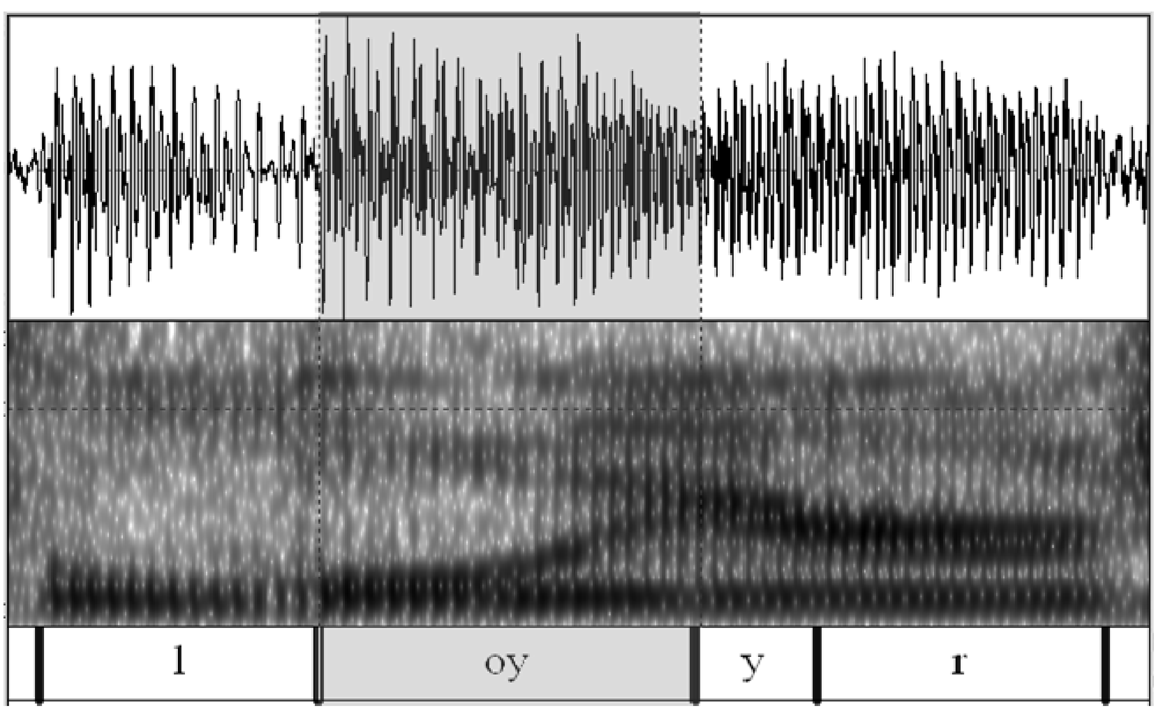


Figure 6. Measuring the vowel of "lawyer" [ˈloɪ.jɹ̩] said by Tucker. Vowel begins where waveform changes and ends where F2 peaks, which coincides with another change in waveform.

amplitude in the waveform, as illustrated in Figure 5. If necessary, the waveform was magnified and visually inspected for a change in the complexity of the wave. The lateral /l/ is often characterized by a rise in F3, but this is not always visible, as in Figure 6. As with the rhotic, it was often necessary to rely on changes in the waveform to measure the lateral, such as the change in amplitude visible in Figure 6.

Glides are characterized by changes in F2: a rise for /j/ and a fall for /w/. The offsets of diphthongs were marked at the extreme point of the glide, which often coincides with a change in waveform amplitude, as in Figure 6, as were vowels following or preceding consonantal glides, which is also the case in Figure 6, where the diphthong /oɪ/ precedes the onset of the next syllable, /j/.

3.2.2.3 Vowel labels

Once marked, vowels were broadly transcribed using ARPAbet transcription codes (cf. Arpabet 2010). Table 3 shows the ARPAbet codes used and their corresponding IPA symbols.

Table 3. ARPAbet transcription codes used to label vowel nuclei.

Monophthongs	
ARPAbet	**IPA**
iy	i
ih	ɪ
eh	ɛ
ae	æ
aa	a / ɑ
ao	ɔ
uh	ʊ
uw	u
ah	ʌ / ə
er	ɝ / ɚ

Diphthongs	
ARPAbet	**IPA**
ey	eɪ
ay	aɪ
aw	aʊ
ow	oʊ
oy	ɔɪ
+r (offglide)	ɹ

Consonants	
ARPAbet	**IPA**
r-, -r	ɹ
l-, -l	l

Note: These are not measured as part of the vowel but are included in its label to indicate that the nucleus is preceded or followed by a liquid, which may affect its quality.

Because liquids are known to affect adjacent vowel quality, preceding and following liquids were linked to the vowel label by a hyphen; e.g., “l-ey” indicates that the nucleus /eɪ/ is preceded by /l/, as in the word “lay,” and “ay-l” indicates /aɪ/ followed by /l/, as in “aisle.” When a liquid was vocalized to such a degree that it behaved as a glide, it was included in the vowel and linked to the nucleus label with a plus sign; e.g., “ao+r” describes the nucleus /ɔ/ with the glide /ɹ/, as in “war.”

3.2.2.4 Statistics

Inferential statistics were calculated using *R* (R Development 2008) and its standard statistical package. T-tests and ANOVAs (analyses of variance) performed for the measures in each experiment are described below.

3.2.2.5 Summary

The measures of hyperarticulation used in this study – speech rate, stressed vowel duration, pitch excursion, and changes in stressed vowel formants – were determined from acoustic measures taken using *Praat*. Token words and phrases and their stressed vowels were delimited and their durations recorded. Speech rate (ms/syllable) was calculated over the word/phrase. The pitch and first and second formant values (F1, F2) were taken at the midpoint of each stressed vowel. Statistics such as ANOVAs and t-tests were performed in *R*.

3.2.3 Experiments

Three experiments were performed in order to measure the effects of Novelty and Evaluation on hyperarticulation. Experiment 1 uses speech rate and vowel duration as measures of hyper-articulation, Experiment 2 uses pitch, and Experiment 3 uses formant values.

3.2.3.1 Experiment 1: Speaking rate and stressed vowel duration

As previously mentioned, lengthening is a prevalent component of hyperarticulated speech. In this experiment, lengthening of whole tokens is measured as speech rate in milliseconds per syllable, and lengthening of stressed vowels is measured in their duration in milliseconds. The dimensions of Novelty and Evaluation were examined separately and in combination for their effects on the two measures of lengthening.

3.2.3.1.1 Hypotheses

H1: There is a reliable effect for Novelty as measured by speaking rate and stressed vowel duration. If speakers hyperarticulate new information as expected, both measures will be greater for new than given information.

H2: There is a reliable effect for Evaluation as measured by speaking rate and stressed vowel duration. If speakers hyperarticulate *attitude* tokens as expected, both measures will be greater for *attitude* than *control* tokens.

H3: There is an interaction between Novelty and Evaluation as measured by speaking rate and stressed vowel duration. The interaction could be additive, or some types of tokens (e.g., *new-attitude* or *given-control*) may be lengthened to different degrees.

3.2.3.1.2 Methods

Two measures were used for this experiment: the speech rate of all 218 tokens and the duration of all 237 stressed vowels. Speech rate was calculated by dividing a token's duration in milliseconds by its length in syllables. Longer words generally have shorter syllables, but calculating speech rate controls somewhat for the length of the token, so that long and short tokens can be compared. The duration of stressed vowels can also be compared across all lengths of tokens (with the same caveat that longer words have shorter stressed vowels). For both measures, the means of *new* and *given* tokens were compared, followed by *attitude* and *control*, and finally combinations of the two (*given-control, new-control, given-attitude, new-attitude*). This was first done for all speakers combined to show group patterns and then with each speaker separated to show individual variation.

3.2.3.1.3 Results

Group means for speech rate offer support for H1 and H2 (effects for Novelty and Evaluation), and those for stressed vowel duration support H2, an effect for Evaluation. Individual results are mixed, again displaying very different patterns between speakers.

3.2.3.1.3.1 Group results

Repeated measures ANOVAs were performed to determine the factors which have significant effects on speech rate and stressed vowel duration. The number of syllables in the token was included as a factor because this is known to affect rate (i.e., syllables in longer words are shorter than those in shorter words). For speech rate, all three factors had significant effects: Novelty ($p < 0.01$), Evaluation, and number of syllables ($p < 0.001$ for both), but the interaction between Novelty and Evaluation was not significant (Table 4).

Table 4. ANOVA showing effects on speech rate (Novelty, Evaluation, number of syllables). All three factors show significant effects. Nov = Novelty, Eval = Evaluation, * p < 0.01, ** p < 0.001.

Factor	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Novelty	1	0.0213	0.0213	3.9133	0.0491	*
Evaluation	1	0.0652	0.0652	11.9766	0.0006	**
syllables	1	0.3408	0.3408	62.5627	0.0001	**
Novelty:Evaluation	1	0.0004	0.0004	0.0724	0.7881	
Nov:syllables	1	0.0002	0.0002	0.0366	0.8484	
Eval:syllables	1	0.0517	0.0517	9.4931	0.0023	**
Nov:Eval:syllables	1	0.0001	0.0001	0.0014	0.9706	
Residuals	229	1.2474	0.0055			

The ANOVA for stressed vowel duration produced different results: only Evaluation was a significant factor (p < 0.01). No significant effect was found for Novelty, number of syllables, or any interaction (Table 5).

Table 5. ANOVA showing effects on stressed vowel duration (Novelty, Evaluation, number of syllables). Only Evaluation shows a significant effect. Nov = Novelty, Eval = Evaluation, * p < 0.01.

Factor	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Novelty	1	0.0018	0.0018	0.6996	0.4038	
Evaluation	1	0.0155	0.0155	6.1702	0.0137	*
syllables	1	0.0001	0.0001	0.0352	0.8514	
Novelty:Evaluation	1	0.0019	0.0018	0.7346	0.3923	
Nov:syllables	1	0.0001	0.0001	0.0371	0.8473	
Eval:syllables	1	0.0008	0.0008	0.3084	0.5792	
Nov:Eval:syllables	1	0.0002	0.0002	0.0820	0.7749	
Residuals	229	0.5764	0.0025			

With all speakers combined, evidence from mean speech rates and stressed vowel durations shows the predicted effects for Novelty and Evaluation. Figures 7-10 show speech rate and stressed vowel duration for *new* and *given* tokens (Figures 7, 9) and *attitude* and *control* tokens (Figures 8, 10). Rate or duration appears on the vertical axes, token type on the horizontal. For both measures, *new* tokens are said more slowly than *given*, and *attitude* more slowly than *control*. Both Novelty and Evaluation show significant effects on speech rate (p < 0.01, p < 0.001, respectively, see Table 4), supporting H1 and H2, but only Evaluation shows a significant effect ($p < 0.01$, Table 5) on vowel duration, supporting H2.

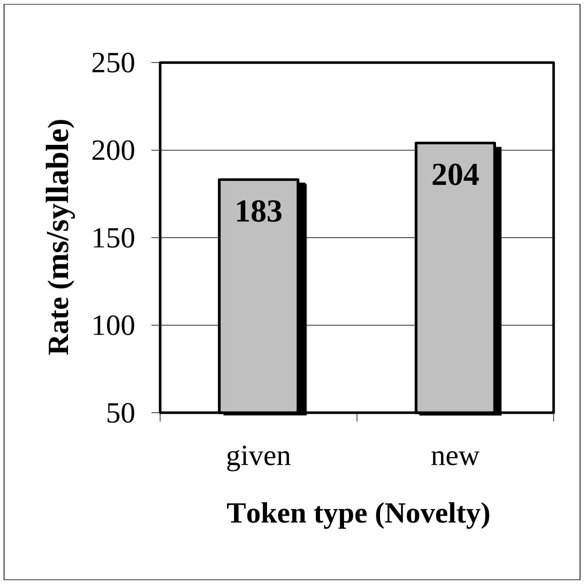


Figure 7. Mean speech rate for *new* and *given* tokens, all speakers combined. Rate (ms/ syllable) shown on vertical axis, token type on horizontal. The difference is significant ($p < 0.01$).

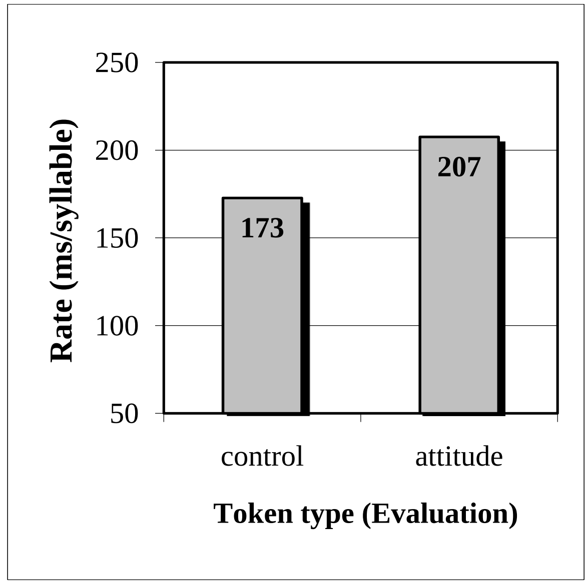


Figure 8. Mean speech rate for *attitude* and *control* tokens, all speakers combined. Rate (ms/syllable) shown on vertical axis, token type on horizontal. The difference is significant ($p < 0.001$).

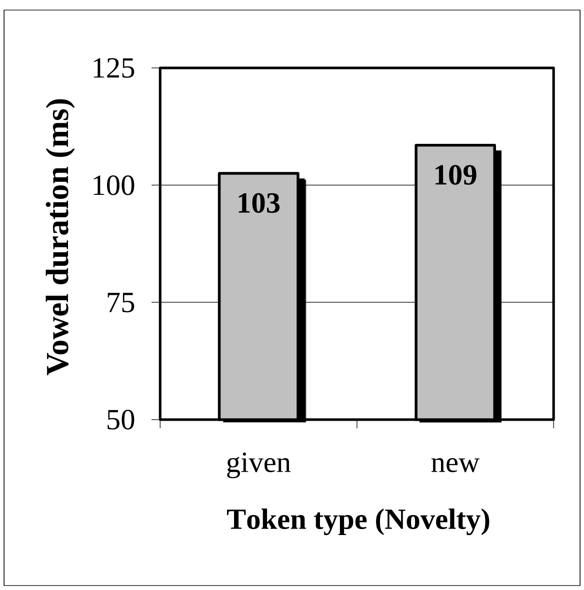


Figure 9. Mean stressed vowel duration for *new* and *given* tokens, all speakers combined. Vowel duration (ms) shown on vertical axis, token type on horizontal. This difference is not significant.

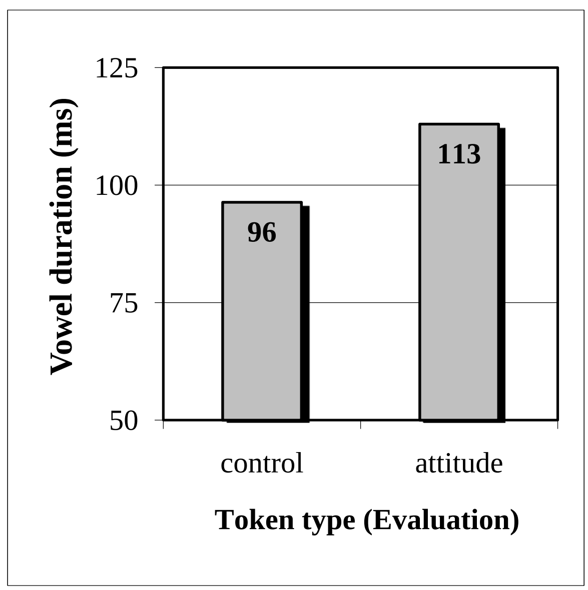


Figure 10. Mean stressed vowel duration for *attitude* and *control* tokens, all speakers combined. Vowel duration (ms) on vertical axis, token type on horizontal. The difference is significant ($p < 0.01$).

When Novelty and Evaluation labels are combined, an interaction appears to be present, supporting H3, but it is not significant for either measure. Figures 11-12 show each token type horizontally with mean speech rate (Figure 11) and stressed vowel duration (Figure 12) vertically. In both, Evaluation has a greater effect than Novelty: all *attitude* tokens are said more slowly than all of their *control* counterparts. Novelty behaves in the expected direction (rate and duration are greater for *new* than *given* items), except for the vowel duration of *attitude* tokens, in which *given* vowels are slightly longer than *new* (114 vs. 112 ms, not a significant difference).

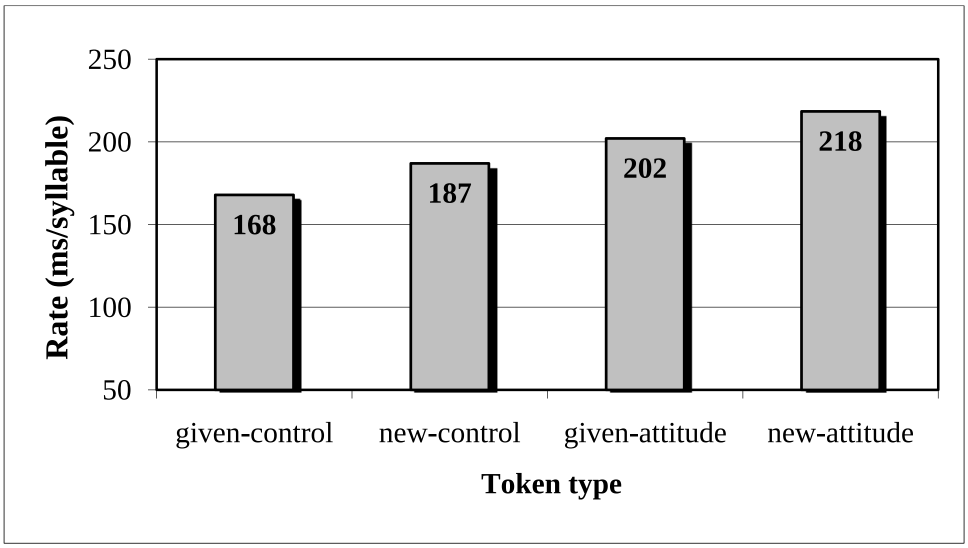


Figure 11. Mean speech rate for each type of token, all speakers combined. Rate (ms/syllable) shown on vertical axis, token type on horizontal. Significant effects were Novelty ($p < 0.01$) and Evaluation ($p < 0.001$).

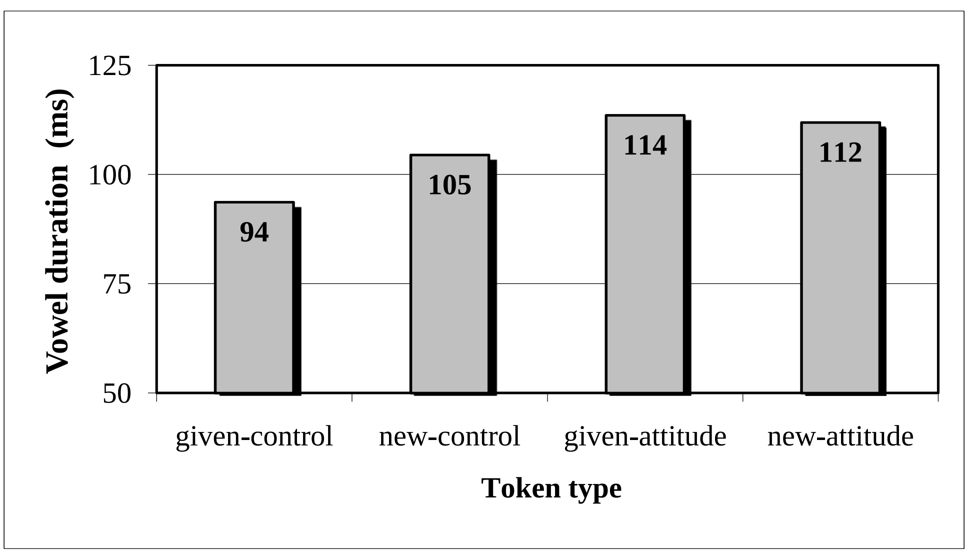


Figure 12. Mean stressed vowel duration for each type of token, all speakers combined. Vowel duration (ms) shown on vertical axis, token type on horizontal. Only Evaluation is a significant factor ($p < 0.01$).

3.2.3.1.3.2 Individual results

Because individual variation contributed to the group results, repeated measures ANOVAs were performed for each speaker separately, resulting in a different combination of significant factors for each speaker (Table 6). As with the group ANOVAs, number of syllables was included as a factor but was significant only for speech rate ($p < 0.01$) for all speakers except Eugene. For speech rate, Pat and Ron showed a significant effect for Evaluation ($p < 0.001$), while only Ron and Tucker showed significant effects for Novelty ($p < 0.001$ and $p < 0.05$, respectively). The interaction between Novelty and Evaluation was not significant for any speaker. For stressed vowel duration, no speaker showed a significant effect for Novelty. Only Eugene and Ron showed significant effects for Evaluation ($p < 0.001$ and $p < 0.01$, respectively), and only Pat had a significant interaction between Novelty and Evaluation ($p < 0.05$). No factor was significant for Eli. Table 6 displays the p-values of the significant factors for each speaker.

Table 6. P-values for significant factors for speech rate and stressed vowel duration, by speaker. The number of syllables in the tokens was also included in the ANOVAs but was a significant factor only for rate, for all speakers except Eugene.

	Rate			**Duration**		
Speaker	**Novelty**	**Evaluation**	**Interaction**	**Novelty**	**Evaluation**	**Interaction**
Eli	--	--	--	--	--	--
Eugene	--	--	--	--	$p < 0.001$	--
Pat	--	$p < 0.001$	--	--	--	$p < 0.05$
Ron	$p < 0.001$	$p < 0.001$	--	--	$p < 0.01$	--
Tucker	$p < 0.05$	--	--	--	--	--

Individual variation contributes to the combined-speaker means in different ways for each speaker. Figure 13 shows the mean speech rate for each type of token said by each speaker. Rate is shown on the vertical axis, speaker on the horizontal, and token type is indicated by bar color. For Novelty, Pat, Ron, and Tucker said *new* tokens more slowly than *given*, as expected, but Eli was the opposite; Eugene said *given-attitude* tokens more slowly than *new-attitude* (like Eli) but had no *new-control* tokens to compare. These differences for Novelty are only significant for Ron and Tucker ($p < 0.001$ and $p < 0.05$, respectively; see Table 6). For Evaluation, Ron behaved as expected, saying *attitude* tokens much more slowly than *control*, but Pat showed the opposite pattern (both significant, $p < 0.001$). Eli, Eugene, and Tucker show no significant difference between *attitude* and *control*.

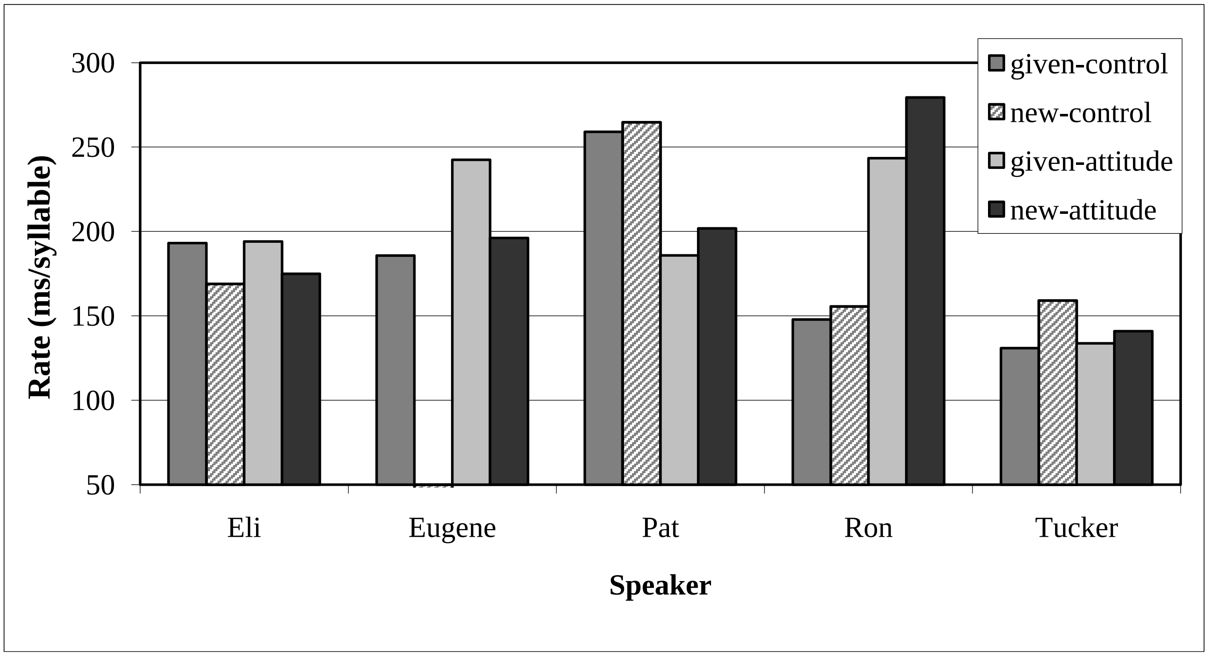


Figure 13. Mean speech rate for each type of token said by each speaker. Rate (ms/syllable) shown on vertical axis, speaker on horizontal; token type indicated by bar color.

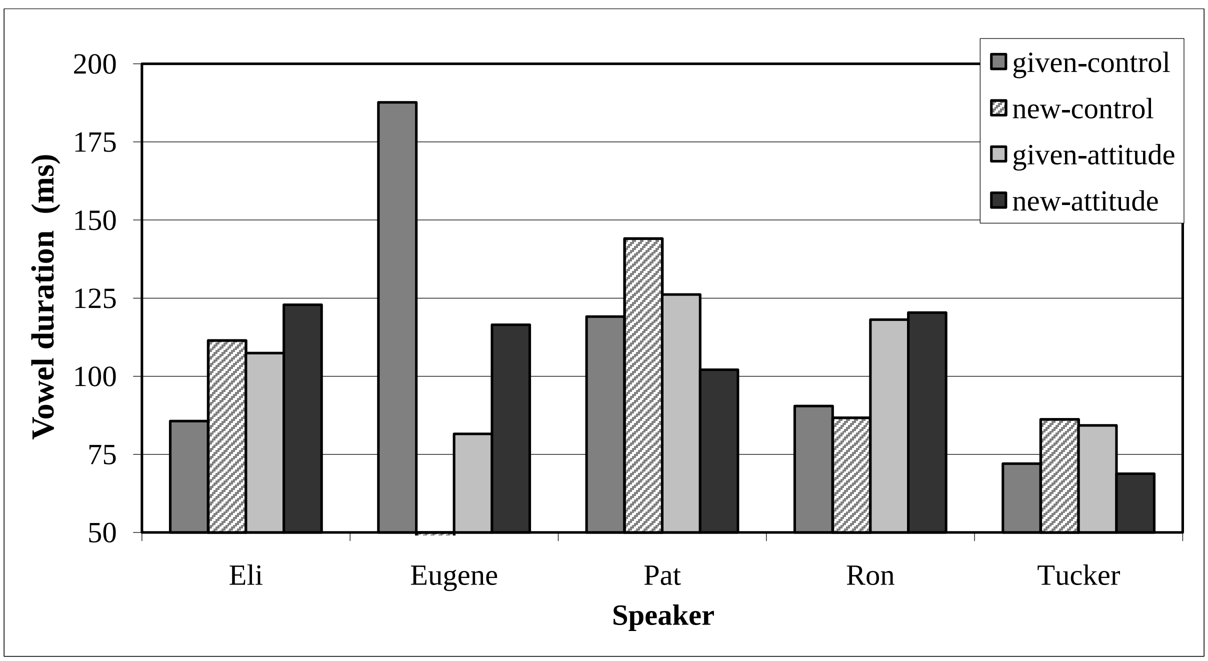


Figure 14. Mean stressed vowel duration for each type of token said by each speaker. Duration (ms) shown on vertical axis, speaker on horizontal; token type indicated by bar color.

Figure 14 shows individual variation for stressed vowel duration. Duration is shown on the vertical axis, speaker on the horizontal, and token type is indicated by bar color. In contrast to the configurations for rate (Figure 13), Eli behaves as expected, with *new* vowels longer than *given*. This is also true for Eugene's *attitude* tokens and Pat's and Tucker's *control* tokens. However, the effect of Novelty was not significant for any speaker. Eli and Ron show clear effects for Evaluation, with *attitude* vowels said more slowly than *control*, as expected, but this is only significant for Ron ($p < 0.01$; Table 6). Pat and Tucker only show this pattern for *given* tokens, but neither difference is significant. Eugene's *given-control* vowels are longer than the mean of any other type said by any speaker; with no *new-control* tokens, the effect for Novelty on *control* tokens is unknown for him, but the effect for Evaluation is starkly opposite from expected, although not significant. Only Pat shows a significant interaction between Novelty and Evaluation ($p < 0.05$).

3.2.3.1.4 Summary

Speech rate (ms/syllable) and stressed vowel duration (ms) were measured for all tokens and vowels in the sample. Means were compared for *new* and *given* tokens, *attitude* and *control*, and combinations of the two (*given-control, new-control, given-attitude, new-attitude*), for all speakers combined and with each speaker separated. Group means for speech rate support H1 and H2 (effects for Novelty and Evaluation), but for stressed vowel duration, the only significant effect is Evaluation, supporting only H2. Novelty and Evaluation both behave in the expected direction such that *new* and *attitude* tokens are said more slowly and with longer stressed vowels. For both measures, Evaluation has a greater effect than Novelty.

Individual variation contributes to the group means, but speakers do not behave the same way for both measures. Overall, Tucker speaks more quickly than the others in most situations. For speech rate, Ron and Pat show clear effects for Evaluation, but acting in opposite directions. Pat, Ron, and Tucker behave as expected for Novelty (although this is not significant for Pat). For stressed vowel duration, the majority of *new* vowels are longer than their *given* counterparts, but the effect of Novelty was not significant for any speaker. Evaluation was a significant factor for Eugene and Ron, and the interaction between Novelty and Evaluation was significant for Pat.

3.2.3.2 Experiment 2: Pitch

The relationship between pitch and hyperarticulation is slightly different than for the other measures in this study. In general, changes in pitch may attract listener attention because they contrast with surrounding material, and such changes are not unidirectional. Pitch excursion – the amount a pitch deviates from a speaker's baseline mean pitch – is used in this study because a contrasting pitch may extend above *or* below a speaker's baseline, but there is another consideration. Speakers may expand or contract their pitch *ranges* in order to contrast words or phrases from surrounding speech. For example, if Speaker A generally has a small pitch range, deviating little from his mean pitch, a relatively large excursion would stand out to listeners and serve as a signal of something important or interesting. Conversely, if Speaker B has a large pitch range, deviating widely from his mean pitch throughout normal conversation, a sharp reduction in this range would also stand out to listeners. Thus, in the usual sense of hyperarticulation as exaggerated speech, pitch excursion can be expected to be larger for highlighted words/phrases, as is hypothesized below, but it must be kept open as a possibility that the opposite may occur, i.e., speakers may show *less* excursion for highlighted items.

3.2.3.2.1 Hypotheses

H1: There is a reliable effect for Novelty as measured by the pitch excursion of stressed vowels. If speakers hyperarticulate as expected, the pitch excursion of new information will be greater than that of given information.

H2: There is a reliable effect for Evaluation as measured by the pitch excursion of stressed vowels. If speakers hyperarticulate as expected, the pitch excursion of *attitude* tokens will be greater than that of *control* tokens.

H3: There is an interaction between Novelty and Evaluation as measured by the pitch excursion of stressed vowels. The interaction could be additive, or each dimension may vary in its effects; i.e., Evaluation may affect *new* and *given* tokens differently, and/or Novelty may affect *attitude* and *control* tokens differently.

As noted above (section 3.2.3.2), although pitch excursion is hypothesized to be greater for *new* and *attitude* tokens, the opposite is also a possibility. That is, speakers may contract rather than expand their pitch ranges to signal Novelty and/or Evaluation.

3.2.3.2.2 Methods

Speakers' pitches were normalized using z-scores before the pitch excursions (deviations from the mean pitch) of each token type (*new, given, attitude, control*) were averaged and compared, both for the group as a whole and for each speaker individually.

3.2.3.2.2.1 Vowels examined

For this experiment, nine vowels had to be excluded because background noise or another speaker's voice masked the audio signal: 228 stressed vowels were measured from 210 tokens (of the total 237 vowels and 218 tokens shown in Table 2, section 3.1.3.2). Table 7 shows the total number of vowels measured for pitch, divided by how they were coded for Novelty and Evaluation.

Table 7. Total number of *attitude* and *control, new* and *given* vowels suitable for measuring pitch.

Token type	**Vowels**		
	New	***Given***	**Total**
Attitude	36	73	**109**
Control	30	89	**119**
Total	**66**	**162**	**228**

3.2.3.2.2.2 Measures

Pitch was measured at the midpoint of stressed vowels using *Praat*'s autocorrelation algorithm with a pitch range of 60-500 Hz. Each speaker's mean pitch, shown in Table 8, was calculated as the mean of all pitch values at midpoint for all stressed vowels. The standard deviation was calculated for the same distribution. Because each speaker has a different mean pitch and range, values need to be normalized across speakers before they can be pooled or compared between speakers. Pitches were normalized by calculating z-scores – the absolute value of: the difference between the vowel's pitch at midpoint and the speaker's mean pitch, divided by the standard deviation (Figure 15; Rose 1987). Z-scores were then averaged for each combination of token type (*new/given, attitude/control*) and compared for all speakers combined and for each speaker separately.

Table 8. Mean pitch at midpoint for each speaker.

Speaker	**Mean Pitch (Hz)**	**Std. Deviation**
Eli	140	55
Eugene	125	28
Pat	206	60
Ron	125	37
Tucker	163	61

$$\text{z-score} = \left\| \frac{\text{vowel's pitch at midpoint} - \text{speaker's mean pitch}}{\text{speaker's standard deviation}} \right\|$$

Figure 15. Z-score formula. Used to normalize pitch values across speakers.

3.2.3.2.3 Results

No statistical support is found for any of the three hypotheses when all speakers are examined as a group, so speakers were separated in order to examine individual behavior. Speakers displayed very different patterns in their use of pitch excursion, although only one speaker showed any statistically significant result.

3.2.3.2.3.1 Group results

When all speakers are combined, there is no reliable effect for Novelty (*new/given*) or Evaluation (*attitude/control*) on pitch. A repeated measures ANOVA was performed, with no significant result for Novelty, Evaluation, speaker, or any interaction between any of the three factors. As can be seen in Figure 16, there is very little difference between *new* and *given* or between *attitude* and *control*; the group pitch results offer no support for any of the three hypotheses.

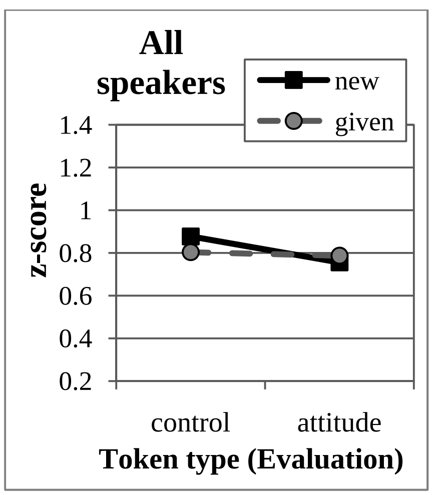


Figure 16. Pitch excursion interaction between Novelty and Evaluation for all speakers together. Mean z-scores are shown on the vertical axis. *Control* and *attitude* tokens are on the horizontal axis; *new* tokens are represented by black squares connected by solid lines; *given* tokens are shown as grey circles connected by dashed lines. There is no difference between *new, given, attitude,* or *control* tokens.

3.2.3.2.3.2 Individual results

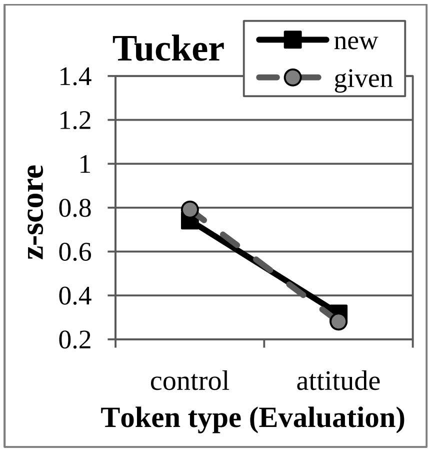
Because there is a large amount of individual variation in how each speaker uses pitch, it is worth examining within-speaker effects. When speakers are separated, it appears that the relatively flat lines shown in Figure 16 result from differing individual behavior (Figure 17). ANOVAs were performed for each speaker separately, but the only significant effect was Novelty for Eugene ($p < 0.01$). He does not have any *new-control* tokens, but there is clearly no effect for Evaluation on *given* tokens (Figure 17e), while the difference between *new* and *given* for *attitude* tokens is large. With only a partial data set for Eugene, minimal support is found for H1, but overall, individual results do not support the hypotheses.

Tucker and Pat show no effect for Novelty: neither uses pitch to signal any difference between *new* and *given* information for either *attitude* or *control* tokens. They do show an effect for Evaluation, but not in the expected direction. Greater pitch excursion was hypothesized to be an indicator of attitude, but Tucker and Pat appear to do just the opposite, perhaps suppressing their pitch ranges when expressing an attitude, as discussed in the opening remarks for this experiment (section 3.2.3.2).

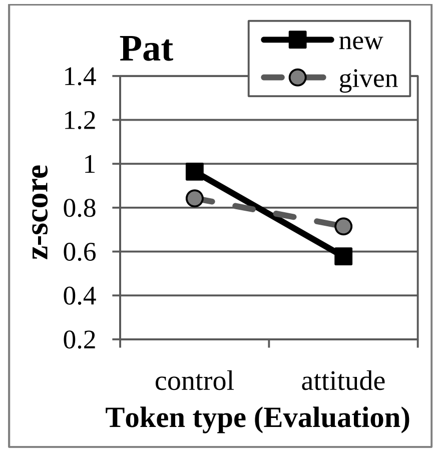
Ron shows a more expected pattern, suggesting some support for H3 (without statistical significance) in that Novelty affects *attitude* but not *control* tokens. *New-attitude* tokens have a much greater pitch excursion, while all other combinations have similar, lower values. Thus, Ron only uses pitch to signal the combination of Novelty and Evaluation, such that there is a large difference between *new-* and *given-attitude* tokens but no difference between *new-* and *given-control* tokens.

Eli shows the oppostite pattern from Ron. All tokens are treated similarly except *given-control* tokens, which have a smaller pitch excursion. In other words, Evaluation affects only *given* tokens, and Novelty affects only *control* tokens. As a combination of interactions, this configuration suggests support for H3, but no factor or interaction was significant. A possible interpretation is that Eli uses pitch to signal both Novelty and Evaluation, but he reaches his maximum pitch excursion with *new-control* tokens, so that he cannot deviate any more to indicate *new-attitude* tokens.

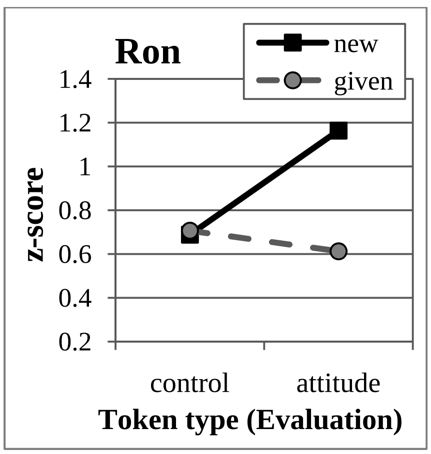
These widely differing patterns help explain the null result obtained when all speakers are combined (Figure 16). It should also be taken into account that each speaker contributes a very different number of tokens to the total. Table 9 shows the number of tokens of each type said by each speaker.



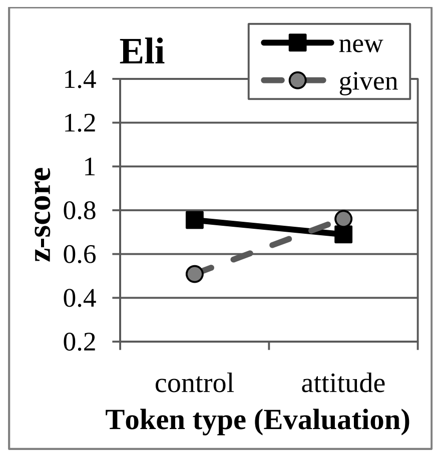
(a) Pitch excursion interaction: Tucker



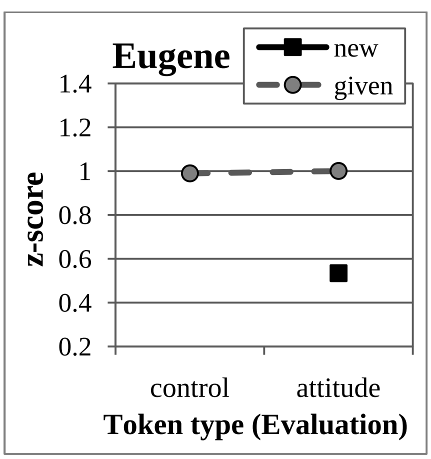
(b) Pitch excursion interaction: Pat



(c) Pitch excursion interaction: Ron



(d) Pitch excursion interaction: Eli



(e) Pitch excursion interaction: Eugene

Figure 17. Pitch excursion interaction between Novelty and Evaluation for each speaker individually. Z-scores are shown on the vertical axis. *Control* and *attitude* tokens are on the horizontal axis; *new* tokens are represented by black squares connected by solid lines; *given* tokens are shown as grey circles connected by dashed lines. Tucker and Pat treat *new* and *given* tokens similarly, while the others show different patterns.

Table 9. Number of tokens coded as combinations of *new/given* and *attitude/control* for pitch measures, organized by speaker.

Tucker	***New***	***Given***
Attitude	2	3
Control	12	31

Pat	***New***	***Given***
Attitude	10	34
Control	8	14

Ron	***New***	***Given***
Attitude	12	14
Control	6	30

Eli	***New***	***Given***
Attitude	6	14
Control	4	13

Eugene	***New***	***Given***
Attitude	6	8
Control	0	6

3.2.3.2.4 Summary

With all speakers combined, no effect is seen for either Novelty or Evaluation as measured by pitch excursion, but each speaker shows a different pattern and contributes a vastly different number of tokens to each combination of types (*new/given, attitude/control*). Tucker and Pat treat *new* and *given* information equally, with *attitude* tokens deviating less from their mean pitches than *control* tokens in an apparent compression of their pitch ranges to signal attitude. Ron shows minimal difference between *new-control* and all *given* tokens, while *new-attitude* tokens deviate largely from his mean pitch. Eli shows little difference between all combinations except *given-control*, which deviates the least from his mean pitch. Eugene shows no effect for Evaluation but does show a wide difference between *new-* and *given-attitude* tokens: his result for Novelty is the only significant effect found for any speaker ($p < 0.01$). Overall, pitch excursion results do not support any of the hypotheses.

3.2.3.3 Experiment 3: Formant values

As a two-dimensional representation, vowel space is a more measure difficult to explore. Each speaker's system is different, and each phoneme may move in a number of directions and overlap with other vowels as it is hyperarticulated or reduced. Therefore, each vowel quality said by each speaker must be examined separately. In order to investigate the effects of Novelty, Evaluation, and both together, only vowel qualities which had tokens of all four combinations (*new-attitude, given-*

attitude, new-control, given-control) said by one speaker were used in this experiment. The Euclidean distances between tokens of each type combination were calculated and compared to determine the behavior of each dimension on each type of token.

3.2.3.3.1 Hypotheses

H1: There is a reliable effect for Novelty as measured by the formant values of stressed vowels. If speakers hyperarticulate as expected, the first and second formants (F1, F2) of new information will be more expanded than those of given information.

H2: There is a reliable effect for Evaluation as measured by the formant values of stressed vowels. If speakers hyperarticulate as expected, the first and second formants (F1, F2) of *attitude* tokens will be more expanded than those of *control* tokens.

H3: There is an interaction between Novelty and Evaluation as measured by the formant values of stressed vowels. The interaction could be additive, or each dimension may vary in its effects; i.e., Evaluation may affect *new* and *given* tokens differently, and/or Novelty may affect *attitude* and *control* tokens differently.

3.2.3.3.2 Methods

Vowel space configurations are extremely variable between speakers, so only a limited set of vowels could be examined. Euclidean distances between combinations of token types (*new/given, attitude/control*) were calculated in order to quantify the effects of Novelty and Evaluation.

3.2.3.3.2.1 Vowels examined

Each speaker's vowels differ in their locations in two-dimensional vowel space, as do the locations of token types in relation to each other, resulting in a different pattern for each vowel said by each speaker. With such individualized patterns, vowel qualities could not be combined for many measures. Therefore, to determine the interaction of Novelty and Evaluation in vowel space, only vowel qualities with all four combinations of token types were examined. Table 10 shows the vowels said by each speaker, and Table 11 shows the numbers of vowels in each coding category. Of the total 62 vowels measured, 9 were coded *new-attitude*, 12 *given-attitude*, 12 *new-control*, and 29 *given-control*. This sub-sample is less balanced than the data set as a whole: although there are still about twice as many *given* as *new* tokens overall and for the *control* category, there are proportionately few *attitude* tokens, and more than half the sample was coded as *given-control*.

Table 10. Vowels with all four combinations of token types said by each speaker.

Speaker	Vowels	
Pat	ɪ	ʌ
Ron	aɪ	ʌ
Tucker	ɛ	ʌ
Eli	ɛ	

Table 11. Total number of *attitude* and *control, new* and *given* vowels measured for formant values.

Evaluation code	Vowels		
	New	***Given***	**Total**
Attitude	9	12	**21**
Control	12	29	**41**
Total	**21**	**41**	**62**

3.2.3.3.2.2 Measures

First and second formants (F1, F2) were measured in Hertz at the midpoint of stressed vowels. Within each speaker, the F1 and F2 values for each combination of token types were averaged for each vowel, resulting in plots resembling Figure 18, where each node represents the mean of all tokens of that type. (The actual plots produced are presented with the individual speaker results in section 3.2.3.3.3.2 below.) Next, the Euclidean distance (Figure 19) was calculated between the nodes to obtain lengths for the lines in Figure 18. These distances quantify the effects of Novelty and Evaluation on each type of token. To examine all speakers together, these distances were averaged to create the graph in Figure 20, to be discussed in the next section.

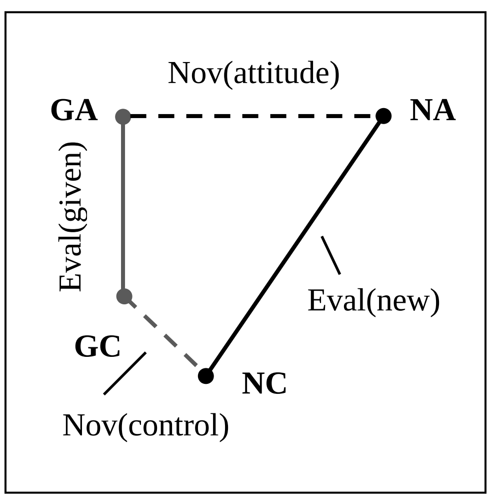


Figure 18. Interactions between Novelty and Evaluation. Nodes represent tokens: N = *new*, G = *given*, A = *attitude*, C = *control*. Lines represent effects of the dimensions on types of tokens: Eval(new) = the effect of Evaluation on *new* tokens, Eval(given) = the effect of Evaluation on *given* tokens, Nov(attitude) = the effect of Novelty on *attitude* tokens, Nov(control) = the effect of Novelty on *control* tokens.

$$\text{Euclidean distance (Hz)} = \sqrt{\Delta F1^2 + \Delta F2^2}$$

Figure 19. Euclidean distance formula.
$\Delta F1$ and $\Delta F2$ are the differences between the F1 and F2 values (respectively) of the endpoints of the distance being measured.

3.2.3.3.3 Results

Evidence was found to support all three hypotheses: Evaluation was found to have a greater effect than Novelty overall, Evaluation affects *new* more than *given* tokens, and Novelty affects *attitude* more than *control* tokens. However, the only statistically significant difference is that between the first and the last of these groups, Evaluation(new) and Novelty(control), and individual speakers show very different configurations in two-dimensional vowel space.

3.2.3.3.3.1 Group results

For vowels which had all four combinations of token types (*new/given and attitude/control*), a two-sample t-test was performed to determine whether the mean Euclidean distances between token types were significantly different with all speakers combined; a significant difference was found only between Novelty(control) and Evaluation(new) (t = -3.407, df = 11.976, p = 0.005). However, the mean distances between token types are arranged in an expected pattern, shown in Figure 20. Effects on token types are shown on the horizontal axis, e.g., Nov(control) = the effect of Novelty on *control* tokens, followed by the combinations of types used to determine the size of the effects, e.g., NC-GC = the distance between *new-control* and *given-control* tokens. This distance is represented on the vertical axis in Hertz.

As was the case for speech rate (Figure 11, section 3.2.3.1.3.1), Evaluation has a greater effect than Novelty overall – it affects both *new* and *given* tokens more than Novelty affects any token. Evaluation also has a greater effect on *new* than *given* tokens, and Novelty has a greater effect on *attitude* than *control* tokens – this configuration was also found for Ron with pitch excursion (Figure 17c). This pattern holds for five of the seven individual speakers' vowels, with the exception of Pat's /ʌ/ and Tucker's /ɛ/.

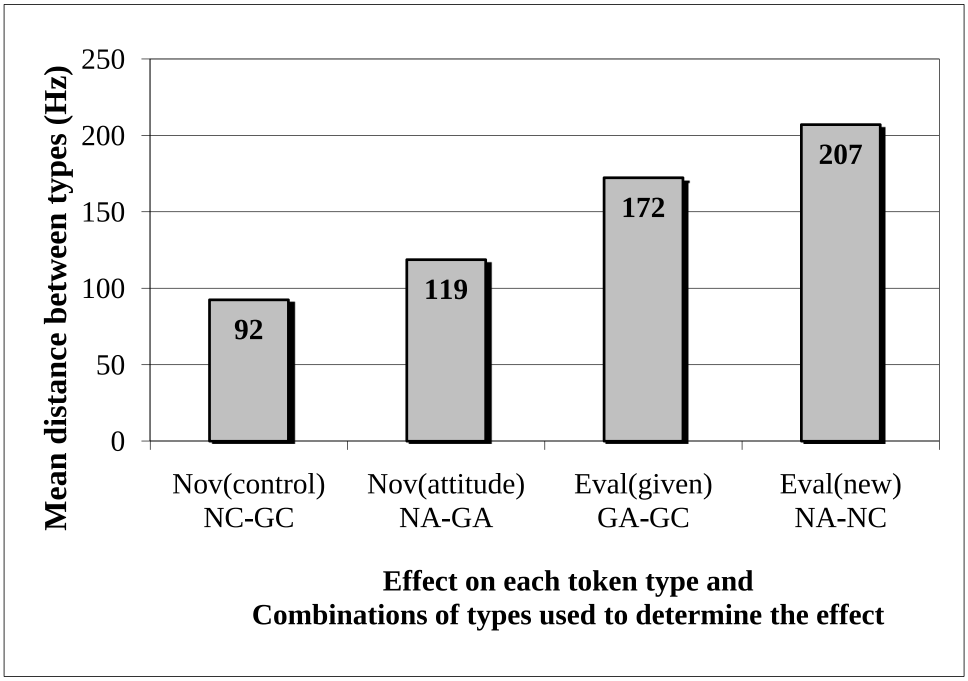


Figure 20. Effects of Novelty and Evaluation on each type of token for all speakers together. Effects on token types are shown on the horizontal axis, e.g., Nov(control) = the effect of Novelty on *control* tokens, followed by the combinations of types used to determine the size of the effects, e.g., NC-GC = the distance between *new-control* and *given-control* tokens. This distance is represented on the vertical axis in Hertz. Evaluation has a greater effect than Novelty in general. Evaluation has a greater effect on *new* than *given* tokens; Novelty has a greater effect on *attitude* than *control* tokens.

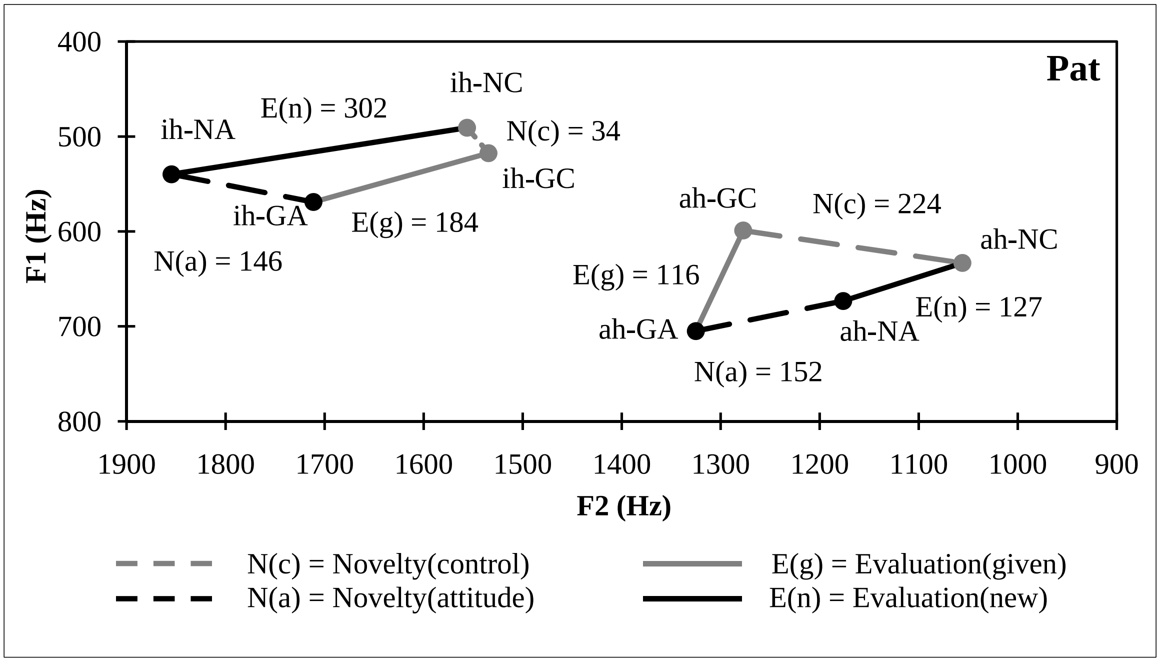
3.2.3.3.3.2 Individual results

Because speakers have different vowel qualities contributing to the overall pattern in Figure 20, the direction of movement for each token type must be considered for each speaker and vowel individually. Figure 21 displays plots for each speaker of the mean formant values of combinations of token types (*new/given* and *attitude/control*) and effects of Novelty and Evaluation on them as measured by the distances between them. F1 is shown vertically, F2 horizontally, both measured in Hertz. Nodes represent token types: N = *new*, G = *given*, A = *attitude*, C = *control*. Lines represent effects on types of tokens: Eval(new) = the effect of Evaluation on *new* tokens, Eval(given) = the effect of Evaluation on *given* tokens, Novelty(attitude) = the effect of Novelty on *attitude* tokens, Novelty(control) = the effect of Novelty on *control* tokens. Values for these lines are their lengths (Euclidean distance) in Hertz.

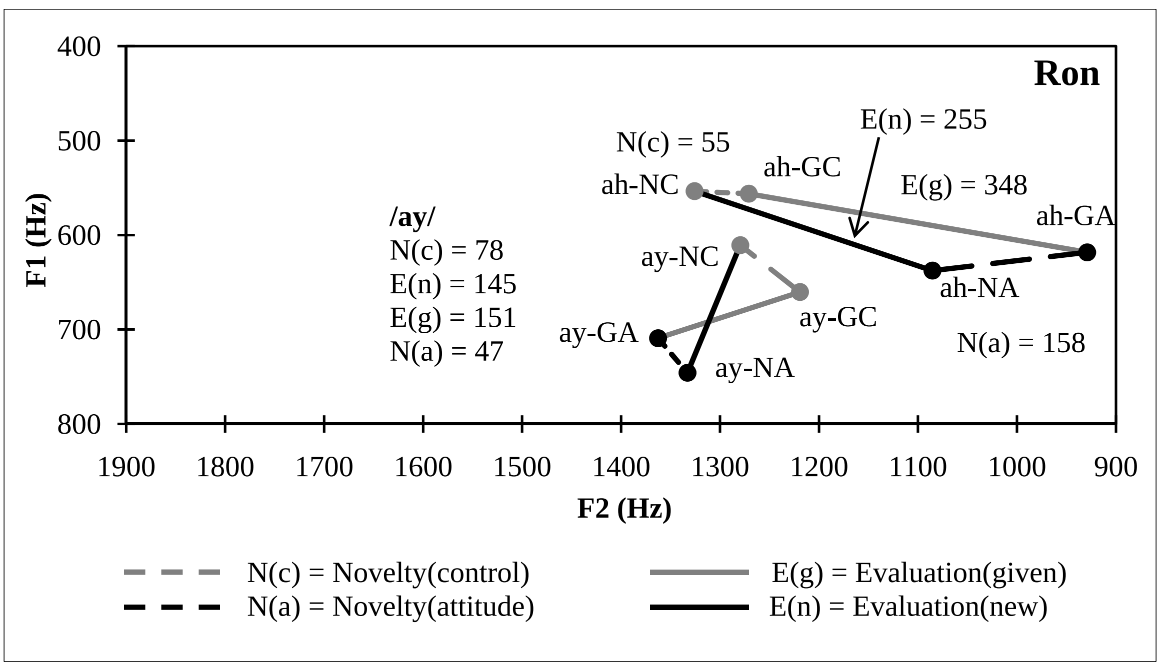
Looking at Figure 21, there is no consistent configuration or orientation common between vowel qualities or within any speaker's vowel plot. Hyperarticulated vowels should move toward the edges of the vowel space. This can be seen for the effect of Novelty with Pat's /ɪ/ (Figure 21a), in which *new* tokens are more peripheral than *given* for both *attitude* and *control* tokens. It could also be argued that *control* tokens are more contracted (often called “centralized”) than *attitude* tokens for this high vowel, which may reduce toward a high central position. It is more difficult to determine the vowel space expansion of the other vowels. As a mid vowel, /ɛ/ may be more limited in the amount it may move away from the center of vowel space; perhaps this allows greater variation between speakers. The nodes for Tucker and Eli are arranged in similar configurations, but their locations are different: *new-attitude* and *new-control* are reversed relative to each other, as are *given-attitude* and *given-control*. The highest node for Tucker’s /ɛ/ is *new-attitude* (Figure 21b), but for Eli, the highest node is *new-control* (Figure 21d). A similar problem arises with the central vowel /ʌ/, common to all speakers except Eli. For Tucker (Figure 21b), Novelty seems to increase toward the upper right-hand corner, Ron’s configuration (Figure 21c) is rotated 180 degrees so that Novelty increases toward the lower left-hand corner, and Pat’s configuration (Figure 21a) orients Novelty toward the lower right-hand corner. Evaluation for Tucker heads toward the upper left, toward the lower right for Ron, and toward low central for Pat. The distances between *new*- and *given-control* tokens are small for Ron and Tucker but not for Pat, suggesting that Novelty does not affect the vowel /ɛ/ in *control* tokens greatly for the former two speakers. Finally, the nuclei of Ron’s /aɪ/ (Figure 21c) suggest a minimal effect for Novelty and a comparatively large effect for Evaluation: the distances between *new* and *given* for both *attitude* and *control* tokens are small, but the distance between *control* and *attitude* tokens is large, with *attitude* being much more expanded than *control* tokens.

3.2.3.3.4 Summary

The general pattern common to most vowels analyzed supports Hypothesis 3: there is an interaction between Novelty and Evaluation. Specifically, Evaluation outweighs Novelty as an effect on hyperarticulation as measured by formant values representing vowel space expansion. Mean distances between combinations of token types increase in the order: Novelty(control), Novelty(attitude), Evaluation(given), Evaluation(new), but only the most extreme values are significantly different. Vowel space plots are difficult to interpret for individual speakers’ vowels and show no consistent or common configuration. This difficulty may be due in part to the central quality of many of the vowels analyzed.

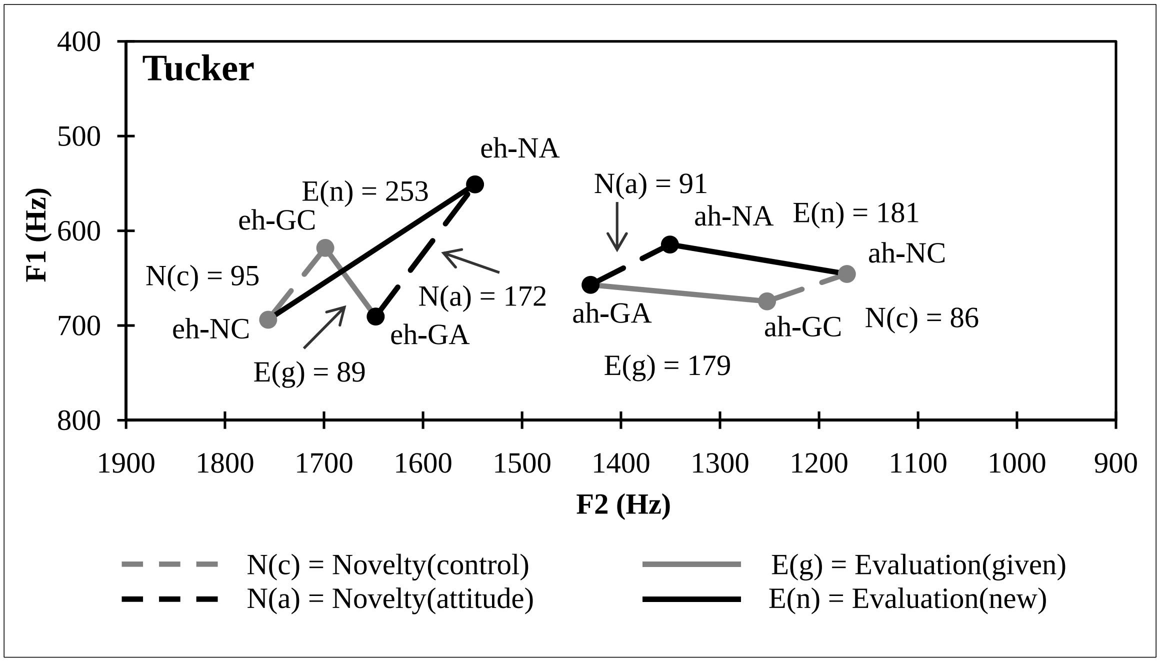


(a) Vowel space plot: Pat

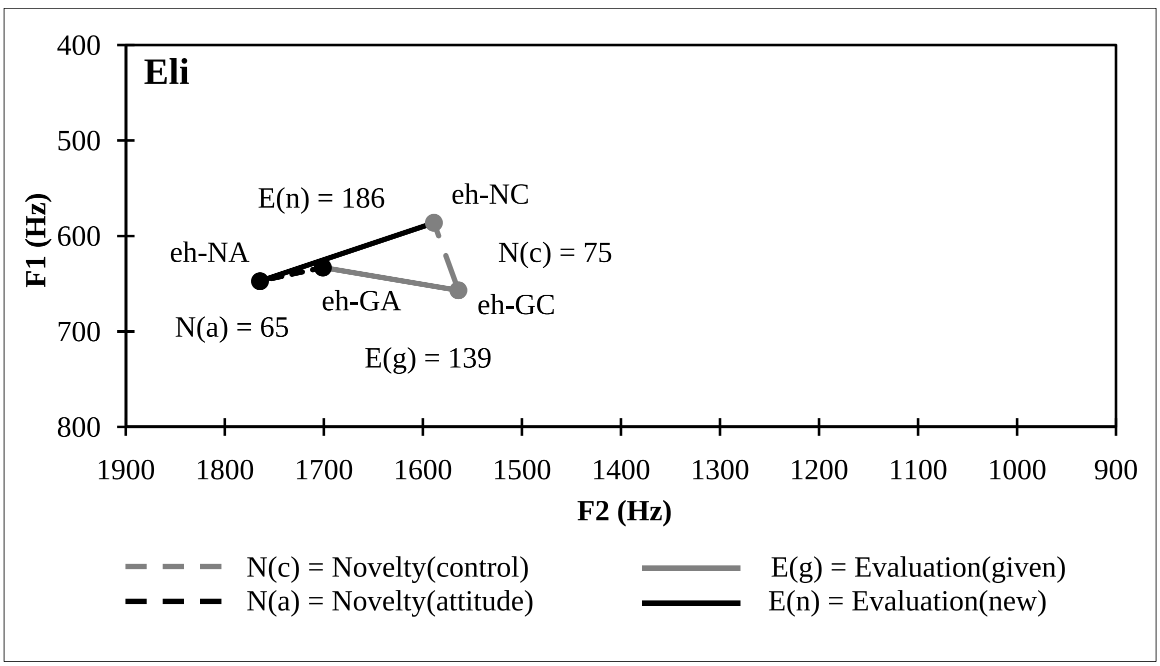


(c) Vowel space plot: Ron

Figure 21. Mean formant values of combinations of token types (*new/given* and *attitude/control*) and effects of Novelty and Evaluation on them as measured by the distances between them. F1 is shown on the vertical axis, F2 on the horizontal axis, both measured in Hertz. Nodes represent token types: N = *new*, G = *given*, A = *attitude*, C = *control*. Lines represent effects on types of tokens: Evaluation(new) = the effect of Evaluation on *new* tokens, Evaluation(given) = the effect of Evaluation on *given* tokens, Novelty(attitude) = the effect of Novelty on *attitude* tokens, Novelty(control) = the effect of Novelty on *control* tokens. Values for these lines are their lengths (Euclidean distance) in Hertz.



(b) Vowel space plot: Tucker



(d) Vowel space plot: Eli

Chapter 4: Discussion

This chapter first discusses the results of this study in section 4.1 and then states the conclusions that can be drawn regarding the effects of Novelty and Evaluation on hyperarticulation in section 4.2. Remarks on the design of the project are outlined in section 4.3, followed by considerations for future research in section 4.4. The paper concludes with a brief discussion of the work's contributions and applications in section 4.5 and an overall summary in section 4.6.

4.1 Results

This section briefly summarizes the results of all three experiments, first in terms of group patterns and then taking into account individual variation.

4.1.1 Group patterns

Of the four measures of hyperarticulation in this study, speaking rate and vowel space expansion were the most reliable indicators of group behavior. These measures also showed the predicted patterns that support all three hypotheses: there are independent effects for both Novelty and Evaluation, and the two interact, with Evaluation having a greater effect on hyperarticulation than Novelty overall, so that all *attitude* tokens are hyperarticulated compared to all *control* tokens, but within each condition, *new* tokens are hyperarticulated compared to *given*. Stressed vowel duration was only useful for signaling attitude, and pitch produced a null result.

4.1.2 Individual variation

In aggregating the patterns of individual variation displayed in each experiment, it is clear that speakers employ different combinations of measures in their treatment of Novelty and Evaluation. Ron behaves as predicted using nearly every measure, hyperarticulating *new* over *given* and *attitude* over *control*, with Evaluation having a greater effect than Novelty overall. Pat is also fairly consistent in his treatment of Evaluation, but not in the expected direction: he tends to hyperarticulate *control* rather than *attitude* tokens, or perhaps a more accurate conclusion is that he reduces rather than hyperarticulates *attitude* tokens. This does not contradict the hypotheses made in this study; rather, Pat is an example of someone who may vary his articulations habitually so that flattening various features is contrastive. Regarding Novelty, Pat does not make a distinction between *new* and *given* for speech rate or pitch excursion, but there is an interaction with his treatment of Evaluation for the other two measures. With *control* vowels, *new* are longer than *given*, but with *attitude* vowels, *new* are shorter than *given*, another example suggesting Pat uses reduction rather than expansion to contrast important information. For one of his two vowels examined for vowel space expansion, he shows little difference between *attitude* and *control*, while the distance between *new* and *given* is large. (For the other vowel, he shows the expected pattern obtained with the group results.)

The other speakers display mixed patterns. Eli does not vary his speech rate or vowel durations on any dimension, but he has the expected pattern for vowel space, and his treatment of pitch illustrates one predicted possibility: that both *new* and *attitude* tokens are hyperarticulated, but he reaches a "ceiling" in his excursion so that *new-attitude* tokens cannot be hyperarticulated more than *new-control* or *given-attitude* tokens. Eugene does not appear to use speech rate to signal anything, but his *control* vowels are longer than his *attitude* vowels, like Pat's. Although he did not have a full data set for the pitch experiment, he showed no effect for Evaluation but a large effect for Novelty, in the opposite direction expected (with *new-attitude* tokens reduced compared to *given-attitude*), suggesting he might behave like Pat for this measure as well. Finally, Tucker, the host, lengthens *new* tokens but patterns like Pat for pitch, contracting his pitch range for *attitude* tokens and displaying no difference between *new* and *given*. One of his vowels shows the expected pattern for vowel space expansion, but for the other, there is little difference between *given-attitude* and *given-control* tokens.

4.1.3 Summary

Speaking rate and vowel space expansion were more reliable indicators of group behavior than vowel duration and pitch excursion, for which the impact of individual variation was also greater. With each speaker employing a different combination of features, individual variation cannot be ignored, particularly with measures such as pitch excursion, for which only two speakers displayed similar patterns.

4.2 Conclusions

After conducting this study, several conclusions can be drawn about the effects of Novelty and Evaluation on hyperarticulation. The support found for each hypothesis is detailed below.

H1: There is a reliable effect for Novelty. Speaking rate and vowel formants reliably demonstrated that new information was hyperarticulated compared to given information.

H2: There is a reliable effect for Evaluation. Speaking rate, vowel duration, and vowel formants reliably showed that speakers hyperarticulated words and phrases about which they express attitudes (*attitude* tokens) when compared to neutral phrases (*control* tokens).

H3: There is an interaction between Novelty and Evaluation. Speaking rate, vowel duration, and vowel formants displayed an additive effect with Evaluation having a greater impact than Novelty, such that all *attitude* tokens were hyperarticulated compared to all of their *control* counterparts. Individual variation played a role in every measure but was particularly strong with pitch excursion, the only measure to show no significant group result.

In short, speakers hyperarticulate new information by lengthening and vowel space expansion, as has been found in previous studies (Aylett 2005; Aylett & Turk 2004). In addition, they lengthen and expand both new and repeated material about which they express attitudes, even more than new information about which they express no attitude. Regarding pitch excursion, individual variation is great: some speakers expand and others contract their pitch ranges to signal new and/or attitudinal material.

4.3 Remarks on project design

This section discusses some practical considerations and limitations of the study's design.

4.3.1 Corpus work

There are a number of advantages and disadvantages to working with a corpus of televised discussion. One great advantage to analyzing pre-recorded speech is the avoidance of any observer effect caused by the researcher. The hosts and guests on the shows knew they were being taped, and were therefore performing for a wider audience than those in the studio, but they were not aware that their speech would later become part of a linguistic study. It is also advantageous to have access to a large amount of data attainable from multiple programs, each with many episodes and a variety of guests, without the expenditure of time necessary to gather it. However, researchers have no control over the content discussed, recording conditions, or speaker demographics. This may be especially problematic for this study because each speaker was from a different dialect area, and so it is impossible to differentiate dialect influences from individual variation. Demographic balance may be difficult to achieve with this corpus in particular: political talk shows might rarely include younger people or working class guests, the hosts (and possibly the guests) are predominantly male, and the ethnicities of participants may be restricted.

Corpus work is also time-consuming. Episodes vary in the number and length of conversational portions suitable for analysis, but in a large study, each must be scanned for usable segments. Segments must be scrutinized multiple times during the coding process, and utterances might not clearly conform to coding requirements and categories. If an item of interest is missing, researchers cannot influence participants in any way in order to obtain it but instead must search through more of the corpus.

4.3.2 Generalizability

A common concern in analyzing television programs is whether the results are generalizable to the population at large. While political talk shows are publicly broadcasted and therefore more performance-oriented than private conversation, their format aims for unscripted, spontaneous discussion. Speakers who regularly appear on television may be formally trained in public speaking, political debate, or acting, but talk show discussions are not held in the scripted, practiced, and carefully-worded styles found in formal debates, which are far from spontaneous “everyday” conversation. The shows also frequently invite guests who are not public figures for their expertise on a subject matter; these guests may not be extensively trained in televised public speaking and therefore may appear more like “regular” people, making the discussion seem more natural and less rehearsed. Because the program is televised, it should be assumed that all participants are giving a performance of some kind, but the resulting speech style is similar to untelevised discussion, with frequent overlapping between speakers, emotional expression of opinion, and attempts to inform and persuade each other. In short, the speech on talk shows such as the one used in this study is very similar to that in many other situations, and therefore generalizability can be reasonably assumed.

An exception to wholesale generalizability may be the behavior of hosts. As discussed in chapter 2 (section 2.4), hosts serve in a moderator role analogous to that held by teachers or people conducting business meetings, but the discourse “rules” for talk show hosts are similar to those in news interviews. As Haddington (2004:120) notes, hosts work toward “maintaining a neutralistic stance” while also “producing hostile/adversarial questions.” They may attribute views to third parties or ask a question that “puts the interviewee in a quandary relative to her institutional background, a publicly taken prior stance, audience expectations, and broader sociocultural beliefs and so on” (Haddington 2007:283). It is possible that the scheme for coding Evaluation used in this study underreports the attitudes expressed by hosts who use such strategies to embed adversarial stances in their questions. The response to this kind of question often appears as an evasion that would not be tolerated in many other situations. Rather than directly answering the question, which is designed “so that answering it poses difficult problems for the interviewee” (Haddington 2007:283) such that a direct answer would be undesirable to listeners or detrimental to the interviewee’s public image, speakers maneuver their remarks in a way that allows them to counteract the stance in the question with their own. Because this is not a common situation in everyday talk (Haddington 2004, 2007), the behavior of hosts may need to be examined more thoroughly in future work.

4.4 Future research

As some of the first work to identify acoustic correlates of attitude expression, this study focused on just a few measures using a limited data set. Future work should improve on both by examining more data and in greater depth. This section provides some suggestions.

4.4.1 Speaker considerations

This study analyzed one randomly-selected episode of one television program, which did not allow for any control over speaker demographics; the speakers were all male but spanned a range of ages and were each from a different dialect region. From corpora of publicly broadcasted discussion, a variety of programs can be analyzed to form a data set which includes a larger sample of speakers representative of various social groupings, such as age, gender, class, ethnicity, and dialect region, each of which may contribute to the ways in which Novelty and Evaluation affect hyperarticulation. For example, one dialect could hyperarticulate using primarily speech rate, while another might favor large pitch excursions. With a large, balanced sample, it would be possible to discover whether the individual variation found in this study is actually reflective of socially-differentiated patterns.

Another way to probe within-speaker variability is to analyze several episodes in which some of the same speakers appear, so that more data is available per speaker, collected on different days and while discussing a variety of topics with different groups of people. This would help identify fluctuation due to situational factors as opposed to intrinsic speaker traits. Having more data per speaker would avoid some of the challenges faced in this study when speakers did not have equal samples for comparison. A variety of vowel qualities could be analyzed, which would be especially helpful when examining vowel space expansion, and portions of particularly productive speakers’ samples could be extracted in order to balance the contributions of each participant.

4.4.2 Attitude types

This study made no attempt to label types of attitudes, but now that evidence has been found that speakers articulate differently when expressing attitudes, future investigation can center on their treatment of subdivisions of attitudes, e.g., positive/negative, strong/weak, or more specific labels, such as hateful, praising, disdainful, cynical, etc. Another factor to consider is whether hyperarticulation might signal an attitude about something other than the hyperarticulated concept. For example, speakers might alter their speech in order to express an attitude about other speakers. Soukup's (2009) study of Austrian talk shows found that selective employment of features of another dialect can be used to transfer attitudes associated with the dialect to other speakers. If components of hyperarticulation are favored by different social groups, it could be used as a dialect feature to express attitudes about other speakers, or this may happen without dialect associations. For example, Speaker A could hyperarticulate a phrase to express his positive attitude about the concept, and then Speaker B might hyperarticulate the same phrase, not to show distain for the concept itself but for Speaker A – maybe the two disagreed on a previous issue or A offended B in some way, and B wants to discredit A by implying his opinion is invalid.

4.4.3 Useful measures

Speaking rate was a useful metric in this study, and it is easy to measure, so it should be used in future work. Stressed vowel duration was somewhat less helpful, but it is also easy to measure, so it need not be excluded. Vowel space expansion produced some positive results, but it is much more complex: each vowel quality must be examined separately for each speaker in order to calculate the Euclidean distances compared in the final analysis of the measure. Future work could simplify this measure by identifying the aspects most sensitive to attitude-expression. Similar problems are present with pitch excursion in that individual variation was so high in this study that group results were null. Within-speaker differences were also not significant, which suggests that more speakers and more data per speaker are required to determine a larger pattern for pitch and/or pitch is not used to signal Novelty or Evaluation. Pitch excursion is easier to measure than vowel formants and has been studied in a variety of contexts, so future work might benefit from investigating it further.

Finally, additional measures of hyperarticulation should be considered. This study focused mainly on stressed vowels, but consonant closures are also manipulated along various acoustic dimensions, including components of hyperarticulation such as duration and place of articulation. In addition, this study used only measures taken at vowel midpoint, but this may be excluding useful information found throughout the vowel. The *Praat* script also recorded the pitch and formant values at onset, offset, 20% and 80% of each vowel's duration – future work should include this data.

4.4.4 Related perception experiment

This study identified acoustic correlates of attitude-expression in production, but it is unknown which, if any, are perceptible or meaningful to listeners, either in isolation or in conjunction with other acoustic features. As a follow-up, a perception experiment could be conducted to determine whether hyperarticulated words are perceived by listeners as expressing attitude. The measures this study found to be reliable correlates of attitude-expression – speech rate, stressed vowel duration, and vowel space expansion – would be manipulated independently and in combinations to discover the degrees to which each affects the perception of attitude-expression. Each measure would be varied in steps using synthesized speech to ensure that any features which could be confounding factors would remain constant. A variety of stressed vowels and surrounding phonetic environments would be used to help illuminate the role of vowel quality in vowel space expansion. Subjects would hear two variants of the same word (perhaps embedded in a carrier phrase) in succession and then indicate which they perceive as expressing the stronger attitude. Subjects' choices and response times would then be correlated against the measures of hyperarticulation. Response times are expected to increase as the differences in the measures of hyperarticulation decrease between members in a pair of stimuli. For example, response times should be greater for a pair of stimuli that differ by only 100 ms in duration as compared to a pair differing by 500 ms.

4.5 Contributions and applications

This study establishes that attitude expression has measurable acoustic correlates. Reliable metrics included three measures of hyperarticulation: speech rate, vowel duration, and vowel formant values. Pitch excursion was not a reliable indicator of attitude. This knowledge might be of interest to a variety of disciplines. For example, discourse- and conversation analysts performing corpus studies of stancetaking and evaluation (i.e., attitude-expression) might find that acoustic measurement can be used in concert with content analysis to support their interpretations, minimize difficulties with inter-rater subjectivity, identify segments of speech likely to contain attitudes, label specific types of attitudes, or perhaps to illuminate previously unidentified cues to attitudes.

Phonologists studying stress and intonation patterns might be able to identify combinations of measures of hyperarticulation used to signal specific types of attitudes, which could vary between languages. Such variation would be useful in studies of dialect contact, second-language learning, and code-switching, especially if differences in the expression of attitude contribute to cultural misunderstandings.

Studies of language acquisition might be able to determine at what age and by what means attitude-expression is acquired. Many studies (cf. Burnham, et al. 1998 and references therein) have found hyperarticulation to be an integral aspect of infant-directed speech (IDS): while talking to infants, mothers use higher pitch and greater pitch range, longer durations, and more expanded vowels “to attract and maintain attention; to convey and elicit positive affect; and to teach language” (p.456). Similar functions may also be underlying motivations for hyperarticulation as a means of attitude-expression, although the range of emotions conveyed may be larger, and the focus on teaching may shift to other uses of hyperarticulation, such as clarity, emphasis, or novelty.

Several applications are conceivable for speech recognition and synthesis. Many studies (e.g., Aylett 2005; Tomita 2007, 2008) have sought to improve automated recognition of natural hyperarticulated speech produced by a speaker in response to recognition errors; this study provides information on interactive human expression that could be used to analyze dynamic group conversations in a variety of situations. For example, Wrede & Shriberg (2003) found that certain speech acts are correlated with indicators of speakers’ high involvement in business meetings. The goal was to improve automatic scanning of recorded meetings so that important interactions (e.g., when participants disagree or make decisions) can be found and reviewed more quickly. Attitude-expression may be common during such interactions, and using acoustic correlates such as those in this study could help locate the interactions and define the attitudes expressed. Finally, findings may be useful in speech synthesis for expressing desired attitudes in production.

4.6 Summary

This study analyzed an episode of a televised political talk show, *Tucker*, for evidence that speakers hyperarticulate concepts about which they express attitudes, a use of hyperarticulation that interacts with the discourse function of signaling new information. Using content analysis, the utterances of five speakers were coded on two dimensions: Evaluation (presence or absence of attitude-expression) and Novelty (new or given information in the discourse segment). To compare the resulting groups, four measures of hyperarticulation were used: speaking rate of the phrases, and duration, pitch excursion, and vowel space expansion (first and second formant values) of stressed vowels in the phrases. Group results showed reliable effects for both dimensions and an interaction such that Evaluation has a greater effect than Novelty overall. Attitude-expressing items were hyperarticulated compared to a control group of neutral phrases, and within each group, new information was hyperarticulated compared to given information. Speaking rate and vowel space expansion showed these effects most reliably, followed by vowel duration. Pitch excursion was not a reliable indicator of either dimension. Individual variation contributed to the group results for all four measures, displaying very different patterns between speakers. These findings provide acoustic correlates to attitude-expression, which previous studies have not explicitly investigated and which can be applied in future work on the identification of specific types of attitudes, the perception of attitude-expression, and automatic speech recognition and synthesis.

REFERENCES

Arpabet. (2010). In *Wikipedia, the free encyclopedia.* Retrieved August 1, 2010 from http://en.wikipedia.org/wiki/Arpabet

Aylett, M. (2005). Synthesising hyperarticulation in unit selection TTS. *Proceedings of the European conference on speech communication and technology (Interspeech 2005)* (pp. 2521-2524)

Aylett, M., & Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech*, *47*(1), 31-56.

Biber, D., & Finegan, E. (1989). Styles of stance in English: Lexical and grammatical marking of evidentiality and affect. *Text - Interdisciplinary Journal for the Study of Discourse*, *9*(1), 93-124.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow, England: Longman.

Boersma, P., & Weenink, D. (2008). *Praat, a system for doing phonetics by computer* (version 5.0.34) [computer software]. Retrieved from http://www.praat.org/

Bolinger, D. L. (1963). Length, vowel, juncture. *Linguistics, 1*(1), 5-29.

Burnham, D., Vollmer-Conna, U., Kitamura, C., Francis, E., Averkiou, V., Olley, A., Nguyen, M., & Paterson, C. (1998). Are you my little pussy-cat? Acoustic, phonetic, and affective qualities of infant- and pet-directed speech. *Proceedings of the fifth international conference on spoken language processing* (pp. 453-456). Sydney, Australia.

Caelen-Haumont, G., & Zei Pollermann, B. (2008). Voice and affect in speech communication. In K. Izdebski (Ed.), *Emotions in the human voice, volume 1* (pp. 215-232). San Diego, CA: Plural Publishing.

Chafe, W. L. (1974). Language and consciousness. *Language*, *50*(1), 111-133.

Clark, H. H., & Haviland, S. (1977). Comprehension and the given-new contract. In R. Freedie (Ed.), *Discourse production and comprehension* (pp. 1-40). Hillsdale, N.J: Lawrence Erlbaum Associates.

Conrad, S., & Biber, D. (2000). Adverbial marking of stance in speech and writing. *Evaluation in text: Authorial stance and the construction of discourse* (pp. 56-73). New York: Oxford University Press.

de Jong, K., Beckman, M. E., & Edwards, J. (1993). The interplay between prosodic structure and coarticulation. *Language and Speech, 36*, 197-212.

Du Bois, J. W. (2007). The stance triangle. In R. Englebretson (Ed.), *Stancetaking in discourse: Subjectivity, evaluation, interaction* (pp. 139-184). Amsterdam: John Benjamins Pub.

Eli Pariser. (n.d.). Retrieved August 1, 2010 from http://www.elipariser.com/

Eugene Robinson. (n.d.). Retrieved August 16, 2010 from *The Washington Post Writers Group* website: http://www.postwritersgroup.com/robinson.htm

Fowler, C. A., & Housum, J. (1987). Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language, 26*, 489-504.

Geist, W. (executive producer) (2007, April 10). Tucker [Television broadcast]. New York: MSNBC.

Grice, P. (1975). Logic and conversation. In P. Cole, & J. Morgan (Eds.), *Syntax and semantics, 3: Speech acts*. New York: Academic Press. Reprinted in Geirsson, H. & Losonsky, M. (Eds.) (1996). *Readings in language and mind* (pp. 121-133). Cambridge, M.A.: Blackwell Publishers.

Haddington, P. (2007). Positioning and alignment as activities of stancetaking in news interviews. In R. Englebretson (Ed.), *Stancetaking in discourse: Subjectivity, evaluation, interaction* (pp. 139-184). Amsterdam: John Benjamins Pub.

Haddington, P. (2004). Stance taking in news interviews. *SKY Journal of Linguistics, 17*, 101-142.

Honorable Ronald A. Cass. (n.d.). Retrieved August 1, 2010 from the website *Cass & Associates, PC*: http://cassassociates.net/ronaldcass.html

Hornby, P. A., & Hass, W. A. (1970). Use of contrastive stress by preschool children. *Journal of Speech and Hearing Research, 13*(2), 395-399.

Hunston, S., & Thompson, G. (2000). Evaluation: An introduction. *Evaluation in text: Authorial stance and the construction of discourse* (pp. 1-27). New York: Oxford University Press.

Labov, W. (1972). *Language in the inner city*. Philadelphia: University of Pennsylvania.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle, & A. Marchal (Eds.), *Speech production and speech modeling* (pp. 403-439). Dordrecht: Kluwer.

Linguistic Data Consortium. (2009). Broadcast conversation corpus produced for the DARPA Global Autonomous Language Exploitation (GALE) project. Audio used with permission.

Ogden, R. (2006). Phonetics and social action in agreements and disagreements. *Journal of Pragmatics, 38*(10), 1752-1775.

Pat Buchanan. (2009). In *Encyclopedia of World Biography*. Retrieved August 1, 2010 from http://www.notablebiographies.com/Br-Ca/Buchanan-Pat.html

Prince, E. F. (1981). Toward a taxonomy of given-new information. In P. Cole (Ed.), *Radical pragmatics* (pp. 223-255). New York: Academic Press.

R Development Core Team. (2008). *R: A language and environment for statistical computing* (version 2.9.2) [computer software]. Retrieved from http://www.R-project.org

Rose, P. (1987). Considerations in the normalisation of the fundamental frequency of linguistic tone. *Speech Communication, 6*(4), 343-352.

Soltau, H., & Waibel, A. (2000). Specialized acoustic models for hyperarticulated speech. *Proceedings of the IEEE international conference on acoustics, speech, and signal processing (ICASSP '00)* (pp. 1779-1782).

Soukup, B. (2009). *Dialect use as interaction strategy: A sociolinguistic study of contextualization, speech perception, and language attitudes in Austria*. Wien: Braumüller.

Tomita, K. (2008). Effects of word familiarity in contexts on speaker's vowel articulation. *Bulletin of Yamagata University: Humanities, 16*(3), 55-67. Retrieved from http://www.lib.yamagata-u.ac.jp/kiyou/kiyouh/kiyouh-16-3/image/kiyouh-16-3-055to067.pdf

Tomita, K. (2007). Effects of word frequency values on speakers' vowel articulation. *Bulletin of Yamagata University: Humanities, 16*(2), 65-75. Retrieved from http://repo.lib.yamagata-u.ac.jp/archive/kiyou/kiyouh/kiyouh-16-2-w065to075.pdf

Tucker Carlson. (2010). Retrieved August 16, 2010 from the *NNDB* website: http://www.nndb.com/people/765/000027684/

Uldall, E. (1960). Attitudinal meanings conveyed by intonation contours. *Language and Speech, 3,* 223-234.

Wennerstrom, A. K. (2001). *The music of everyday speech: Prosody and discourse analysis*. New York: Oxford University Press.

Whalen, D. H., Magen, H. S., Pouplier, M., Kang, A. M., & Iskarous, K. (2004). Vowel production and perception: Hyperarticulation without a hyperspace effect. *Language and Speech, 47*(2), 155-174.

Wichmann, A. (2002a). Attitudinal intonation and the inferential process. In B. Bel, & I. Marlien (Eds.), *Proceedings of the speech prosody 2002 conference* (pp. 11-16). Aix-en-Provence, France: Universite de Provence.

Wichmann, A. (2002b). Looking for attitudes in corpora. *Language and Computers, 40,* 247-260.

Wrede, B., & Shriberg, E. (2003). The relationship between dialogue acts and hot spots in meetings. *Proceedings of the automatic speech recognition and understanding workshop*.

Appendix 1: Evaluation Coding Protocol

The following protocol was used for coding concepts as *attitude* or *control*.

Considering what the speaker says about the concept, look for how s/he performs the following actions.

a. Introduction: "I want to say something about this." Speaker works to keep the topic in play.
   - Initiates discussion on the topic
   - Reintroduces or returns to the topic when others have changed focus
   - Overtly changes the subject: "Let's talk about…"
     - Note: Do not include changes of topic performed in a host capacity (which are not considered conversational).
   - Repeats a phrase with the token if interrupted (cf. Labov 1972).
   - Repeats the token in close succession, e.g., in order to add or change modifiers: "A foolish comment, a nasty comment, and ugly comment" (add, change) "a comment, an off-handed comment" (add). **Note: count repetitions used to correct oneself as one repetition when dividing total number of items of support.**
   - Tries to get others' attention: "Look / Listen"
   - Asks / demands to be allowed to speak: "Let me say this, I'll say this, Let me finish."

b. Overt evaluation: Speaker takes a stance (cf. Du Bois 2007).
   - Overtly states opinion: "I think / believe, the way I see it, in my opinion, it's clear to me" (cf. *subjectivity* in Conrad & Biber 2000).
   - Makes a prediction or mentions probability: "Probably, maybe, in all likelihood" (cf. Conrad & Biber 2000, Hunston & Thompson 2000).

c. Evaluative descriptors, modifiers, synonyms/referents: Speaker uses description to evaluate (cf. Hunston & Thompson 2000, Labov 1972).
   - Uses evaluative modifiers with the token (e.g., adjectives, adverbs, intensifiers, comparatives) that show bias, opinion, or are otherwise easily identified as emotional or opinionated (obviously, ridiculous, important, unbelievable, impressive, etc.).
   - Uses evaluative (positive or negative) descriptors when referring to the token, the concept behind it, synonyms or other referents.
   - Makes evaluative comments about the token concept: "it turned my stomach."

d. Credibility: Speaker offers support for a stance (cf. Conrad & Biber 2000).

- Presents stance as fact: “actually, in reality, in fact”
- Cites experts (people, studies, publications) as sources: “Polls show, Most Americans agree, So-and-so says”
- Presents self as expert / authority, as if to say, “I know because I …”
- Expresses certainty: “we all know, that’s just how it is” (Hunston & Thompson 2000, Biber & Finegan 1989).

e. Persuasion and recommendation: Speaker attempts persuasion, makes a suggestion or recommendation (cf. Conrad & Biber 2000).

- Attempts to persuade others to agree: “Think of it this way; You know”
- States what “should” be done regarding the concept

f. Agreement: Speaker agrees or disagrees with another speaker (cf. Du Bois 2007, Conrad & Biber 2000).

- Shows agreement or disagreement with another speaker about the token concept: “I agree/disagree, not at all, absolutely, no, that’s right”

When finished, count the number of items of supporting evidence and divide by the number of repetitions. If the ratio is 2.00 or higher, mark the token *attitude*; otherwise, mark it *control.*