

1. SUMMARY OF PROPOSED RESEARCH

This research addresses the contribution of prosodic features to the interpretation of utterances. While there has been much progress in describing phonetic and phonological aspects of prosody, there has been less research into semantic and pragmatic interpretations. There have been even fewer corpus studies of meaning and prosody in natural, interactive speech (Hirschberg 2004). Our goal is to conduct a series of mid-scale corpus studies of North American English, investigating how the prosodic system expresses meanings relating to sentence types and information structure. We will annotate a corpus of cooperative telephone conversations and a corpus of argumentative television discussions for phonology and meaning and examine the relation between these levels of analysis. We will also supplement the corpus studies with a series of laboratory perception experiments to further test some of our conclusions.

Sentence type involves the distinction between declarative and interrogative sentences, and these can be positive or negative in polarity. Different sentence types are prototypically used to express certain speech acts (e.g., assertions, questions, and contradictions) and have prototypical prosodic shapes. In actual discourse, however, all of these sentence types can have a final rise or fall in pitch, as well as having certain pre-final prosodic configurations. Through a series of corpus and experimental studies, we will evaluate theories that rely on linguistic intuitions in explaining the differences in meaning between the different intonation patterns and different declarative and interrogative forms. Due to its precision, we will adopt interactive dynamic semantics as our semantic/pragmatic model. The theory is dynamic in that it explicates the way that a given utterance is interpreted relative to the already established context and then serves to change that context. It is interactive in that the changing information states of both speaker and hearer are modeled.

Information structure involves aspects of meaning such as topic, focus, contrast and givenness. We will integrate existing theories of these components of information structure and their prosodic correlates, and evaluate resulting hypotheses about the information structure of different declarative and interrogative forms against natural speech corpora. According to our theory, the topic links an utterance to its context, and the prosodically prominent comment or ‘focus’ expresses the information about the topic that is conveyed by the utterance. Contrastive and less given expressions are also associated with a greater degree of prosodic prominence. Particular pitch patterns as well as accent placements will be examined in disentangling the prosodic correlates of different information structure categories.

For the phonological/phonetic coding, we will use an enriched version of the standard, autosegmental-metrical system for transcribing the prosody of North American English, ToBI (Beckman & Ayers-Elam 1997).

Through previous pilot studies, we have identified four hypotheses to test further, both in extended corpus studies and perception experiments. These hypotheses are that high-rise yes-no questions ($H^*H-H\%$) are associated with a ‘reminder’ meaning, as opposed to the possibility of an ‘out-of-the-blue’ meaning associated with unmarked low-rise ($L^*H-H\%$) questions; that falling yes-no questions are perceived as ‘non-genuine’ questions; that a distinctive low-rise contour ($L^*L-H\%$) marks contradictions; and that a particular rising pitch accent ($L+H^*$) marks ‘contrast’ instead of ‘topic’. However, our corpus studies will not be limited to the study of these four hypotheses.

Our investigation of the prosody of sentence types and information structure has significant potential for improving computational speech generation and analysis technology, primarily because we are exploring the meaning of prosody in real speech. It will also have applications in English as a Second Language instruction. The contribution of our research in both domains will be to clarify the relevance of prosody in discourse, by defining the role of accent placement, phrasing, and type of intonation contour in the meaning and appropriateness of utterances in spoken communication.

2. DETAILED DESCRIPTION

Program of Research

Objectives

The goal of the proposed research is to test, using natural spoken data, conclusions about the meaning of prosodic patterns that have been previously arrived at primarily using intuitive judgments. This will be accomplished through both a series of mid-size corpus studies and experimental studies on the meaning and prosody of sentence types and information structure in North American English. The specific objectives of this research are as follows:

- Evaluate and extend existing theoretical proposals about prosody and meaning of sentence type and information structure against corpora of natural interactive discourse;
- Test the corpus study conclusions concerning the correlations between prosody and meaning in the laboratory by conducting a series of perception experiments.
- Begin to develop a research proposal extending this research to investigate similarities and differences between English and Spanish.

Context

Theoretical background/framework

The meaning of prosody has long been studied with regard to the intonation patterns that distinguish question speech acts from statement speech acts. Prototypically, questions are articulated using rising interrogative syntactic forms, and statements are articulated using falling declarative forms. However, all three levels of analysis need to be distinguished since, for example, a sentence expressed with a rising declarative form is often used to perform the speech act of asking a question. The prosody overlaid on the syntactic sentence type thus serves to disambiguate speaker intent with regard to speech act type. Prosody is also used to mark parts of a sentence as new or focused information in relation to other parts of the sentence that express information that is background.

Thus, prosody works in conjunction with syntactic sentence type to indicate the speech act being performed by an utterance, while also indicating different aspects of information structure expressed by different components of the sentence being uttered. Our primary goal is to elucidate both of these aspects of the meaning of prosody and to examine the interaction between them. In the following subsections we outline the specific existing theoretical frameworks guiding our proposed research.

Sentence Type. The proposed research will seek to study the prosody and meaning of sentence types within a theoretical framework that dynamically models the changing information states of speaker and hearer and the common ground. The framework adopted in our approach integrates the work of Bartels 1999, Gunlogson 2001, and Steedman 2006. Bartels proposes that the L- phrase accent, prototypically present in declarative statements, alternative questions, and wh-questions, as well as in atypical falling yes-no interrogative questions marks ‘speaker assertiveness’ (Stalnaker 1978), whereas the H- phrase accent present in rising interrogative and declarative questions denotes the absence of this meaning. ‘Assertiveness’ here means that the speaker expresses an instruction to the addressee to commit himself publicly to the proposition expressed. Gunlogson focuses on falling and rising declarative sentences, comparing both with rising yes-no interrogative sentences, and concludes that the declarative form expresses commitment to the propositional content: commitment by the speaker for falling declaratives and by the addressee for rising ones. Steedman further extends the interactional meaning of prosodic

shape to all sentence types, and proposes that H-H% and L-H% rising contours always indicate hearer commitment to the proposition evoked by the utterance, while L-L% falling contours and H-L% level contours indicate commitment by the speaker.

At the prosodic level, all of these authors elucidate the meaning of final nuclear tones, factored compositionally into the meanings of particular pitch accents and boundary tones, following the ToBI system of autosegmental phonology (Beckman & Ayers-Elam 1997). Furthermore, the data analyzed has almost exclusively been analysts' native speaker intuitions about what it would mean to pronounce a given sentence in a given way as opposed to data drawn from corpora of actual speech.

Information Structure. The literature on information structure is vast and apparently conflicting. We propose that the literature can be reduced to three frameworks that are especially relevant to prosody. The central concepts of these frameworks can be defined in terms of two notions: the binary topic vs. focus notion ('information focus' in Gundel & Fretheim 2004), and Rooth's 1985, 1992, 1996 unary, alternative set notion of focus. The three frameworks combine these two notions in different ways. We propose that both notions of focus are linguistically relevant and indispensable to a complete account of the prosodic data.

The primary framework that we adopt is that of Gundel & Fretheim 2004, in which topic is defined as what the utterance is about relative to the focus. This relational topic-focus distinction, we claim, is a less formalized version of Vallduvi's 1992 ground-focus distinction. Vallduvi divides the ground into link and tail, which corresponds to Hedberg & Sosa's 2006 unratified-ratified topic distinction, following Lambrecht & Michaelis 1998. In our approach, we supplement the relational topic-focus distinction with Vallduvi & Vilkuna's 1998 second information-structure parameter, which they call 'kontrast'. We suggest that kontrasts correspond to Rooth's alternative set foci. In the proposed research, we will compare the results we get with the Gundel and the Vallduvi frameworks with the 'theme-rheme' framework of Steedman 2000, 2006. Relational topics, themes and kontrast have all been claimed by different researchers to be marked intonationally by the pitch accent or tune L+H*(L-H%) in English. An evaluation of these claims is crucial to our project. Hedberg 2006 presents a preliminary discussion of the relationship between the three information-structure frameworks and their predictions about L+H* intonation in English and use of the topic-marker in Korean.

In order to analyze natural data, it is necessary to first formulate guidelines for determining the information structure of utterances. We propose to carry through with considering the entire set of three frameworks discussed above, and will consider all of the sentence and polarity types. Gundel's 1985 relational definition of topic (the entity which "the speaker intends to increase the addressee's knowledge about, request information about, or otherwise act with respect to") applies to all sentence types, including questions. Some subset of traditional 'topic tests' can be applied to determine what the possibilities for relational topic are in a given utterance (Gundel 1974, Reinhart 1981, Cohan 2000): the 'question test', 'as for test', 'said about test', and 'pseudocleft test.' Vallduvi & Vilkuna's kontrasts, and Steedman's themes and rhemes can be identified by looking for evocations of alternative sets.

Relation of proposed research to ongoing research

During the past five years, we have integrated the literature and conducted six pilot studies on the prosody of both sentence type and information structure in North American English corpora. Since 2001, Hedberg and Sosa have received four institutional small grants (two from SSHRC), which have enabled us to develop the groundwork for the proposed research.

Sentence Type. In previous research (Hedberg & Sosa, 2002), we found that 73 wh- and yes-no questions in a television discussion corpus could be either rising or falling, and also explored prenuclear accents on the question words in wh-questions and the auxiliary verb in yes-no questions. In Hedberg,

Sosa & Fadden 2004, we conducted a study examining the meaning and prosody of 113 negative and positive question types (wh-, yes-no, declarative) in a telephone corpus. Hedberg, Sosa & Fadden 2006 focused on examining 104 positive yes-no questions in the same telephone corpus, and concluded that low-rise (L*H-H%) is the unmarked question contour used for ‘out-of-the-blue’ questions in American English, and that high rise (H*H-H%) is used for ‘reminder’ questions that relate back to previous discourse. We also found that falling contours (e.g. H*L-L%) correlate with ‘non-genuine’ yes-no questions such as requests for action as opposed to information.

In Hedberg & Sosa 2003, we examined 106 negative utterances in the television corpus and found mostly falling contours, but found quite a few rises (L-H%) on negative declaratives compared to positive declaratives. Based on this work, we hypothesized that the greater number of rises on negatives is due to the addressee commitment to the negated proposition in the case of contradictions, following Steedman 2000. To further test this hypothesis, we explored the meaning of different prosodic types of 40 contradictions on a combined television and telephone corpus (Hedberg, Sosa & Fadden 2003), finding that rising contours were indeed often found on contradictions.

In order to examine the relationship of prosody to utterances that presuppose various types of epistemic commitments by speaker and hearer, we propose to further explore declarative and interrogative sentences with negative as well as positive polarity (c.f. Romero & Han 2004), and in addition to explore tag questions (c.f. Ladd 1981), which typically involve components expressing both positive and negative polarities. At the level of speech acts, contradictions, which are intended to change the addressee’s information state by removing a proposition from the common ground, will be distinguished from negative statements that simply add a negative proposition to the common ground.

In our proposed research, we will extend the pilot studies, so that we have enough data to be able to test our results statistically. We will also conduct a series of laboratory perception experiments designed to test the findings of the corpus studies in a controlled manner.

Information Structure. In previous work on information structure (Hedberg & Sosa 2002), we examined prosodic topic-focus marking on questions in American English from the television corpus, and found that the wh-word in wh-questions and the inverted auxiliary in negative yes-no questions had a tendency to be marked with the L+H* pitch accent. These findings support the proposal of Steedman 2000 that these elements constitute a ‘theme’ because they evoke but do not restrict a set of alternatives, but equally well supported is the hypothesis that L+H* marks contrast. However, in Hedberg, Sosa and Fadden 2004, we tested this hypothesis on the telephone corpus and failed to find confirmation, perhaps due to the less emphatic nature of the questions in the telephone corpus.

In our examination of negative utterances and contradictions in Hedberg & Sosa 2003 and Hedberg, Sosa and Fadden 2003, we found both to be quite often marked with the ‘contradiction contour’ of Liberman & Sag 1974 (L*L-H%) or else L+H*L-H%. The difference was claimed in Hedberg, Sosa & Fadden 2003 to turn on the difference in meaning between L* and L+H* in that L* is said to mark given information, i.e., ‘activated’ information in the terms of Gundel, Hedberg & Zacharski 1993; whereas L+H* is said to mark asserted position on a scale or contrast (Pierrehumbert & Hirschberg 1990). However, Steedman (2000, 2006) claims that L* marks a ‘controversial rheme’, and L+H* marks theme. More data is required in order to assess these different information structure hypotheses.

In relation to this topic, we have recently published a pilot study of the television corpus (Hedberg & Sosa 2006), reporting on 210 phrases aimed at evaluating the claim that relational topics are marked L+H*, and have concluded that L+H* can equally well mark foci. With regard to this data, an interesting controversy has ensued with Steedman 2006 on the proper analysis of nuclear L+H*L-H% in negative contradictions. For him this contour marks an ‘isolated theme’, where the rheme is merely

implicated—i.e. the contour marks a contrastive topic; but for us this contour marks a contrastive focus (Hedberg & Sosa 2006, Hedberg 2006). We will address this issue in a proposed larger study.

While our proposed research on information structure will concentrate on the distinction between topic and different types of focus, prosodic cues to global discourse structure (Grosz & Hirschberg 1992, Nakatani 1997, Swerts & Ostendorf 1997, Herman 2000, Venditti & Hirschberg 2003) will need to be taken into consideration. Utterance phrasing will be analyzed, and we will also consider focus projection (Chomsky 1971, Gussenhoven 1983, 1999, Selkirk 1984, 1995, Rochemont 1986, 1998, Cinque 1993, Welby 2003, Féry et al. 2006, German et al. 2006), and deaccenting (Ladd 1980, Schwarzchild 1999, Swerts, et al. 2002) in assigning information structure categories. To the extent that we need to adopt a syntactic framework, we will follow Chomsky & Lasnik 1995, Selkirk 1995 and Zubizarreta 1998.

Phonological Coding. Most of the semantics/pragmatics literature on the meaning of prosody examines the placement of pitch accent and pitch movement in establishing meaning. However, the phonetic/phonological literature concludes that duration and intensity are also factors in the determination of meaning. We will thus explore the role of all three phonetic parameters. In transcribing intonation, we will follow the system of transcription that we developed in our pilot studies, which is closely modeled on the ToBI guidelines (Beckman & Ayers-Elam 1997). Using this approach, we will include the full range of ToBI pitch-accents, including the use of the pitch accent !H* for downstep. We will supplement the ToBI pitch accent system with ;H* to represent ‘upstep’ and ↑H* for ‘increased range’ in specifying high pitch excursions. Increased range is characterized by high peaks and low valleys while upstep is mostly an H* which is higher than a previous H*, reversing any downdrift or declination effect.

The distinction between L+H* and H* is important to the study of information structure, but the distinction is interpreted differently by different authors (e.g. Ladd & Schepman 2003), and ToBI transcribers are often not in agreement in distinguishing them (Herman & McGory 2002). We will distinguish these accents phonetically instead of following the ToBI requirement to simply use the label H* in cases of doubt and to not use L+H* on an accented syllable which is initial in the utterance. L+H* implies a perceptible rise onto the peak as well as a higher peak than a simple H*.

Finally, most previous research on the meaning of intonation within the ToBI framework has concentrated on compositionally analyzing the meaning of the different prosodic parts of the final nuclear tune of an intonation phrase, building upon the pioneering work of Pierrehumbert & Hirschberg 1990. In our research, we will consider whether final tunes might make a contribution to meaning in their entirety instead of compositionally with regard to their ToBI components (c.f. Ladd 1996). In addition, as in Hedberg, Sosa & Fadden 2006, we will examine the role of pre-nuclear intonation in the utterance meaning, and consider the possibility that the contour over the entire sentence needs to be taken into account, expanding on the conclusion of Grabe et al. 1998, for example, who found that the form of the initial boundary tone in Dutch is determined by aspects of the following pitch accent and thus does not make an independent compositional contribution.

One issue that always arises in the study of natural speech is the need to code reliably for all form, tone and meaning categories. To do this we propose to construct coding manuals, which can be consistently consulted and updated, and to have at least two coders for each task, who will be checked statistically for coding reliability. Hedberg, Sosa & Fadden 2006 conducted a three-coder reliability study on the prosodic annotation of 48 positive yes-no questions in American English. We used the transcriber-pair-word technique (Pitrelli et al. 1994) and found reliability ranging from 72% to 85.3%, depending on which aspect of the transcription was measured, and pointed out that this compares about equally with other work on ToBI coding, where reliability has ranged between 68% and 87%, depending on the measure and style of speech (e.g. Syrdal & McGory 2000, Yoon et al. 2004).

Importance, originality, and contribution to knowledge of proposed research.

The primary innovation of this research is that we are looking at natural speech in evaluating and proposing claims about the relation between prosody and meaning. Also innovative is the fact that we will check our main corpus study conclusions about the meaning of particular prosodic patterns through laboratory experiments. Converging evidence from these different methodologies will lend great support to conclusions. The investigation of the prosody of sentence types and information structure has significant potential in the area of speech technology, in both generation and analysis. It also has implications for English as a Second Language instruction.

Methodology

During this three-year project, we plan to conduct extended corpus studies, and then select key hypotheses supported by the corpus studies, in order to test them further through laboratory perception experiments. From our pilot corpus studies, we have identified four hypotheses to test in both extended corpus studies and experiments. The four hypotheses are that high-rise questions are associated with a 'reminder' as opposed to an 'out-of-the-blue' meaning, that falling versus low-rise yes-no questions are perceived as 'non-genuine' questions, that contradictions with L+H*L-H% contours as opposed to L*L-H% contours are perceived as 'contrastive', and that L+H* marks 'contrast' instead of 'topic'.

Corpus studies.

During the first two years of the project, we will conduct corpus studies. The natural language corpora to be used and studied in the proposed research are comprised of face-to-face conversations and telephone conversations. For face-to-face conversations, we will use an existing corpus of 75 McLaughlin Group transcripts and videotapes, a half-hour political discussion program aired on the Public Broadcasting Service in the United States. The format is useful since it readily enables modeling as a question-answer dialogue (Carlson 1983, Roberts 1996, 1998, Buring 2003), which aids identification of information structure. For telephone conversations, we will use two electronic corpora from the Linguistic Data Consortium. The CallHome Corpus of American English includes transcripts and speech files of 125 telephone conversations between people who know each other, and the Switchboard Corpus includes transcripts and speech files of 2,400 telephone conversations between people who do not know each other. One advantage of the telephone data is that the two sides of the conversation are recorded on separate tracks, enabling speech analysis free from overlap to be conducted. The telephone data has good sound quality and is more easily searchable electronically, but the television data is more argumentative in nature and is thus a valuable source for contradictions and contrastive elements.

The Principal Investigator and graduate assistant will extract the relevant utterances from the corpora and perform the semantic/pragmatic annotation. Then the Co-Investigator and graduate assistant will determine the phonological annotation. Thus, the semantic and prosodic coding will in part be done independently, helping to avoid the potential for bias. To aid in annotation and to perform measurements, we will use the speech analysis programs PitchWorks and Praat. We will compare the semantic and phonological codings and establish correlations, using statistics such as chi-square.

At every stage of annotation, two coders will code a large part of the data independently, and will compare their results, using the transcriber-pair-word statistic and the kappa statistic (Carletta 1996) to determine reliability. We will also investigate more recent agreement statistics described in Artstein & Poesio 2005. If reliability fails to achieve 80%, we will examine the disagreements and revise our coding scheme accordingly. Coding manuals that articulate the guidelines for annotation will thus be iteratively developed. Any remaining disagreements will be resolved through consensus.

For sentence types, we will first conduct a corpus study of 500 positive yes-no questions from the CallHome corpus, testing the hypotheses that low-rise is the unmarked question type in American English, that high-rise is used to mark questions whose content ties back to immediate or prior context,

and that falling yes-no questions express ‘non-genuine’ questions. Next, we will conduct studies of 500 wh-questions, 250 negative yes-no questions, 250 declarative questions, and 250 tag questions drawn from the CallHome and Switchboard corpora, testing hypotheses derived from the dynamic semantics literature about the epistemic presuppositions of speaker and hearer associated with these sentence types, and examining the role of prosody in modulating those effects. After that, we will study 500 negative statements from the CallHome and Switchboard corpus, and will examine 250 positive and negative contradictions in the television discussions, again relating prosodic contour to aspects of meaning.

For information structure, we will conduct a corpus study with 1,250 phrases in the television discussions, refining the methodology of the pilot study reported in Hedberg & Sosa 2006. The semantic coding team will annotate issue one from 50 McLaughlin transcripts for the categories ratified topic, unrated topic, contrastive topic, contrastive focus, and plain focus, using topic tests supplemented by inspection of the context. We will then select for prosodic coding five exemplars of each category from each transcript, maximally distributed across the transcript, so if there are 20 contrastive topics, we will select every fourth occurrence. The semantic coding team will check the semantic coding for reliability and resolve disputes through consensus. We are confident that we can arrive at a reliable coding for information structure categories since Taboada and Hadic-Zabala 2006 found 83% agreement on identifying the topic in 338 utterances from the CallHome American English corpus using the same topic tests that we will use (e.g. question test, *as for* test, pseudocleft test). The phonological coding team will then annotate each selected utterance for ToBI categories, and once again establish reliability in coding. We will then correlate meaning categories with prosodic categories, and test the resulting distribution for statistical significance using chi-square analyses.

Laboratory perception experiments.

If our hypotheses concerning the meaning of utterances with specific prosodic contours achieve statistical significance at the .05 level in the corpus studies, in the third year of the project we will conduct perception experiments to test these hypotheses in a controlled manner. In these experiments, subjects will be asked to assign interpretations in a forced choice paradigm, to controlled stimuli differing minimally in prosody. We will follow the technique of Hirschberg & Ward 1992, wherein synthetically varied tokens of L*+HL-H% differing in pitch range were presented to subjects in order to establish an ‘incredulity’ or ‘uncertainty’ interpretation. In each experiment, 16 stimulus items recorded by the investigators will be modified using Praat, and a counterbalanced list of 64 stimuli (32 experimental items plus controls) will be presented in random order by the E-Prime experiment operating system. The data will be tested statistically using ANOVA with a .05 level criterion of significance. Twenty-five Simon Fraser University students who are native speakers of North American English and have normal hearing will be used as subjects in each experiment, and will be paid \$10 for their participation in the half-hour studies.

Timeline

Year one: Train graduate student; corpus studies on questions (data extraction, annotation, reliability testing, compilation of results, statistical analysis). Year two: Corpus studies on negation/contradiction and information structure. Year three: Experiments (complete design, conduct experiments, data analysis); outline book sections on English; begin development of research proposal to study Spanish.

Communication of Results

Our long-term goal is to write a book on the meaning of prosody in North American English and Spanish. During the proposed research, we will write several journal articles (e.g. for *Journal of Phonetics*, *Journal of Pragmatics*, *Language and Speech*) and present at several conferences (e.g. LSA, CLA, ICPHS 2007, Speech Prosody 2008). Our website will include a list of current literature, the coding manuals, and our research papers, along with some of the transcripts and speech files.

3. LIST OF REFERENCES

- Artstein, Ron and Massimo Poesio. 2005. $\text{Kappa}^3 = \text{Alpha}$ (or Beta). CS Technical Report CSM-437. Department of Computer Science, University of Essex.
- Bartels, Christine. 1999. *The Intonation of English Statements and Questions: A Compositional Interpretation*. New York: Garland Publishing.
- Beckman, Mary E. and Gayle Ayers Elam. 1997. *Guidelines for ToBI Labeling, Version 3*. Ohio State University.
- Büring, Daniel. 2003. On D-trees, beans, and B-accent. *Linguistics and Philosophy*.
- Carlson, Laurie. 1983. *Dialogue Games: An Approach to Discourse Analysis*. Dordrecht: D. Reidel.
- Carletta, Jean. 1996. Assessing agreement on classification tasks: the kappa statistic. *Computational Linguistics* 22. 249-254.
- Chomsky, Noam. 1971. Deep structure, surface structure and semantic interpretation. In D. Steinberg and L. Jakobovits, eds., *Semantics: an Interdisciplinary Reader in Philosophy, Linguistics, and Psychology*, New York: Cambridge University Press. 183-216.
- Chomsky, Noam and Howard Lasnik. 1995. The theory of principles and parameters. Chapter 1 of Noam Chomsky, *The Minimalist Program*, MIT Press. 13-128.
- Cinque, G. 1993. A null theory of phrase and compound stress. *Linguistic Inquiry* 24. 239-297.
- Cohan, Jocelyn Ballantyne. 2000. *The Realization and Function of Focus in Spoken English*. Ph.D dissertation, The University of Texas at Austin.
- Féry, Caroline and Vieri Samek-Lodovici. 2006. Focus projection and prosodic prominence in nested foci. *Language* 82. 131-150.
- German, James, Janet Pierrehumbert and Stefan Kaufmann. 2006. Evidence for phonological constraints on nuclear accent placement. *Language* 82. 151-168.
- Grabe, Esther, Carlos Gussenhoven, Judith Haan, Erwin Marsi, and Brecht Je Post. 1998. Preaccentual pitch and speaker attitude in Dutch. *Language and Speech* 41. 63-85.
- Grosz, Barbara and Julia Hirschberg. 1992. Some intonational characteristics of discourse structure. *Proceedings of the 7th International Conference on Spoken Language Processing*.
- Gundel, Jeanette K. 1974. *The Role of Topic and Comment in Linguistic Theory*. Ph.D. thesis, University of Texas at Austin. Reprinted 1989, Garland Press.
- Gundel, Jeanette K. 1985. 'Shared knowledge' and topicality. *Journal of Pragmatics* 9:1. 83-97.
- Gundel, Jeanette K. and Thorstein Fretheim. 2004. Topic and focus. In Laurence R. Horn and Gregory Ward, eds., *The Handbook of Pragmatics*. Blackwell. 175-196.
- Gundel, Jeanette K., Nancy Hedberg and Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69. 274-309.
- Gunlogson, Christine. 2001. *True to Form: Rising and Falling Declaratives in English*. Ph.D. dissertation, University of California, Santa Cruz.
- Gussenhoven, Carlos. 1983. Focus, mode and the nucleus. *Journal of Linguistics* 19. 377-417.
- Gussenhoven, Carlos. 1999. On the limits of focus projection in English. In Peter Bosch and Rob van der Sandt, eds. *Focus: Linguistic, Cognitive, and Computational Perspectives*. Cambridge University Press. 43-55.
- Hedberg, Nancy. 2006. Topic-focus controversies. In Molnár, Valeria and Susanne Winkler (eds.) *The Architecture of Focus*. *Studies in Generative Grammar* 82. Berlin, New York: Mouton de Gruyter. 373-397.
- Hedberg, Nancy and Juan M. Sosa. 2006. The prosody of topic and focus in spontaneous English dialogue. In Lee, Chungmin, Matthew Gordon and Daniel Büring (eds.). *Topic and Focus: Cross-linguistic Perspectives on Meaning and Intonation*. *Studies in Linguistics and Philosophy* 82. Dordrecht: Springer. 101-120.

- Hedberg, Nancy and Juan M. Sosa. 2003. Pitch accents in negative sentences. Proceedings of the 15th Congress of the Phonetic Sciences. Barcelona.
- Hedberg, Nancy and Juan M. Sosa 2002. The prosody of questions in natural discourse. Proceedings of Speech Prosody 2002. 375-378.
- Hedberg, Nancy, Juan M. Sosa and Lorna Fadden. 2006. Tonal constituents and meanings of yes-no questions in American English. Proceedings of Speech Prosody 2006. Dresden, Germany. 386-389.
- Hedberg, Nancy, Juan M. Sosa and Lorna Fadden. 2004. The semantics and configuration of English questions. Proceedings of Speech Prosody 2004. Nara, Japan..
- Hedberg, Nancy, Juan M. Sosa and Lorna Fadden. 2003. The intonation of contradictions in North American English. Paper presented at Prosody and Pragmatics, Lancashire, England, November 2003.
- Herman, Rebecca 2000. Phonetic markers of global discourse structures in English. *Journal of Phonetics* 28. 466-493.
- Herman, Rebecca and Julia Tevis McGory. 2002. The conceptual similarity of intonational tones and its effects on intertranscriber reliability. *Language and Speech* 49. 1-36.
- Hirschberg, Julia. 2004. Pragmatics and intonation. In Laurence R. Horn and Gregory Ward, eds. *The Handbook of Pragmatics*. Blackwell. 515-537.
- Hirschberg, Julia and Gregory Ward. 1992. The influence of pitch range, duration, amplitude and spectral features on the interpretation of the Rise-Fall-Rise intonation contour in English. *Journal of Phonetics* 20:241-251.
- Ladd, D. Robert. 1980. *The Structure of Intonational Meaning: Evidence from English*. Bloomington: Indiana University Press.
- Ladd, D. Robert. 1981. A first look at the semantics and pragmatics of negative questions and tag questions, *CLS* 17, 164-171. Chicago: Chicago Linguistic Society.
- Ladd, D. Robert. 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.
- Ladd, D. Robert and Astrid Schepman. 2003. "Sagging transitions" between high pitch accents in English: Experimental evidence. *Journal of Phonetics* 31. 81-112.
- Lambrecht, Knud and Laura Michaelis. 1998. Sentence accent in information questions: Default and projection. *Linguistics and Philosophy* 21. 477-544.
- Liberman, Mark and Ivan Sag. 1974. Prosodic form and discourse function. *CLS* 10, 416-427. Chicago: Chicago Linguistic Society.
- Nakatani, Christine. 1997. *The Computational Processing of Intonational Prominence: A Functional Prosody Perspective*. Ph. D. dissertation, Harvard University.
- Pierrehumbert, Janet and Julia Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. In P. R. Cohen, J. Morgan, and M. E. Pollack, eds., *Intentions in Communication*. Cambridge, MA: MIT Press. 271-311.
- Pitrelli, J.F.; Mary E. Beckman, and Julia Hirschberg. 1994. Evaluation of prosodic transcription labeling reliability in the ToBI framework. Proceedings of the International Conference on Spoken Language Processing. Yokohama, Japan, 123-126.
- Reinhart, Tanya. 1982. *Pragmatics and linguistics: An analysis of sentence topics*. Indiana University Linguistics Club.
- Roberts, Craige. 1996. Information structure in discourse: Towards an integrated formal theory of pragmatics. In J. Hak Yoon and A. Kathol, eds. *Ohio State University Working Papers in Linguistics*, Volume 49.
- Roberts, Craige, 1998. Focus, information flow, and Universal Grammar. In Peter Culicover and Louise McNally, eds. *The Limits of Syntax*. Academic Press. 109-160.

- Rochemont, Michael. 1986. Focus in Generative Grammar. John Benjamins.
- Rochemont, Michael S. 1998. Phonological and structural focus. In Peter Culicover and Louise McNally, eds. *The Limits of Syntax*. Academic Press. 337-363.
- Romero, Maribel and Chung-hye Han. 2004. On *Yes/No* questions. *Linguistics and Philosophy* 27. 609-658.
- Rooth, Mats. 1985. Association with Focus. Ph.D. dissertation, University of Massachusetts at Amherst.
- Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1. 75-116.
- Rooth, Mats. 1996. Focus. In Shalom Lappin, ed., *The Handbook of Contemporary Semantic Theory*. 271-298.
- Schwarzschild, Roger. 1999. Givenness, AvoidF and other constraints on the placement of accent. *Natural Language Semantics* 7. 141-177.
- Selkirk, Elizabeth O. 1984. *Phonology and Syntax: the Relation between Sound and Structure*. Cambridge, MA: MIT Press.
- Selkirk, Elizabeth. 1995. Sentence prosody: Intonation, stress, and phrasing. In John A Goldsmith, ed., *The Handbook of Contemporary Phonological Theory*, Blackwell. 550-569.
- Stalnaker, Robert. 1978. Assertion. In Peter Cole (ed.) *Syntax and Semantics Vol 9: Pragmatics*. 315-322. New York: Academic Press.
- Steedman, Mark. 2006. Information-structural semantics for English intonation. In Lee, Chungmin, Matthew Gordon and Daniel Büring (eds.). *Topic and Focus: Cross-Linguistic Perspectives on Meaning and Intonation*. *Studies in Linguistics and Philosophy* 82. Dordrecht: Springer. 245-264.
- Steedman, Mark. 2000. Information structure and the syntax-phonology interface. *Linguistic Inquiry* 31. 649-689.
- Swerts, Marc, Emiel Kraemer and Cinzia Avesani. 2002. Prosodic marking of information status in Dutch and Italian: a comparative analysis. *Journal of Phonetics* 30. 629-654.
- Swerts, Marc and Mari Ostendorf. 1997. Prosodic and lexical indications of discourse structure in human-machine interactions. *Speech Communication* 22. 25-41.
- Syrdal A. and A. McGory. 2000. Inter-transcriber reliability of ToBI prosodic labeling. *Proceedings of the International Conference on Spoken Language Processing*. Beijing, China, 235-238.
- Taboada, Maite and Loreley Hadic-Zabala. 2006. What are the units of discourse structure? Deciding on units of analysis for corpus annotation. Unpublished manuscript, Simon Fraser University.
- Vallduvi, Enric. 1992. *The Informational Component*. Ph.D. dissertation, University of Pennsylvania.
- Vallduvi, Enric and Maria Vilkuna. 1998. On rheme and kontrast. In Peter Culicover and Louise McNally, eds. *The Limits of Syntax*. Academic Press. 79-108.
- Venditti, Jennifer and Julia Hirschberg. 2003. Intonation and discourse processing. *Proceedings of the 15th Congress of the Phonetic Sciences*. Barcelona.
- Welby, Pauline. 2003. Effects of pitch accent position, type, and status on focus projection. *Language and Speech* 46. 53-81.
- Yoon, Tae-Jin, Sandra Chavarría, Jennifer Cole, and Mark Hasegawa-Johnson. 2004. Intertranscriber reliability of prosodic labeling on telephone conversation using ToBI. *ICSA International Conference on Spoken Language Processing, Interspeech 2004*, Jeju, Korea, Oct. 2004, 2729-2732.
- Zubizarreta, Maria. 1998. *Prosody, Focus and Word Order*. Cambridge, MA: MIT Press.