HDSP #396132, VOL 46, ISS 6 (September 29, 2009)

Prosodic Markers of Saliency in Humorous Narratives

Lucy Pickering, Marcella Corduas, Jodi Eisterhold, Brenna Seifried, Alyson Eggleston, and Salvatore Attardo

QUERY SHEET

This page lists questions we have about your paper. The numbers displayed at the left can be found in the text of the paper for reference. In addition, please review your paper as a whole for correctness.

Q1. Au: For author Corduas, please check my English translation of University and Affiliation.

Q2. Au: Please confirm or correct correspondence address.

Q3. Au: Spelled Nakajima in the References section. Fix spelling where necessary.

Q4. Au: I added author Waletzky to this Labov (1967) citation to match that found in the References section. Correct as edited?

Q5. Au: I added author Waletzky to this Labov (1967) citation to match that found in the References section. Correct as edited?

Q6. Au: Can a page number be provided for this quote from Audrieth (1998)?

Q7. Au: For these page numbers (pp. 56–57), are you citing from Brown et al. (1980) or Brown & Yule (1983)?

Q8. Au: Is Martin et al. still in press? Provide update here if now published.

(*Queries continued on next page*)

TABLE OF CONTENTS LISTING

The table of contents for the journal will list your paper exactly as it appears below:

Q9. Au: Cohen (1960) is not found in the References section. Add full entry there.

Q10. Au: Provide name of manufacturer and city/state location for this CSL model.

Q11. Au: Provide a year of reference here for the Kolmogorov–Smirnov test, and include a full, corresponding entry in the References section for this test and its authors as well.

Q12. Au: Provide a year of reference here for the Wilcoxon signed-ranks test, and include a full, corresponding entry in the References section for this test and its author as well.

Q13. Au: Burns & Ward (1982) is not found in the References section. Add full, corresponding entry there.

Q14. Au: Attardo & Pickering (in press) is not found in the References section. Add full entry there, and provide an update here if this has now been published.

Q15. Au: Provide a year of reference here for the Mann–Whitney test, and include a full, corresponding entry in the References section for this test and its author as well.

Q16. Au: For this Pickering & Attardo citation, provide the year of the draft you read instead of "in preparation".

Q17. Au: Audrieth (1998): Provide publisher information.

Q18. Au: Grosz & Hirschberg (1992): Provide editor(s) names for these proceedings, and clarify publisher information (name of publisher and city/state or country locations.

Q19. Au: Johns-Lewis (1986): Provide article/chapter page range.

Q20. Au: Klewitz & Couper-Kuhlen (1999): Can you provide the article page range?

Q21. Au: Provide name(s) of editor(s) of this larger work, and verify publisher information as edited.

(*Queries continued on next page*)

Q22. Au: Marasek (2006): Provide publisher information.

Q23. Au: Is Martin et al. still in press? Provide update here if now published.

Q24. Au: Nakajima & Allen (1993): Provide editor(s) names for these proceedings, and clarify publisher information (name of publisher and city/state or country locations.

Q25. Au: Nakatani (1993): I am unsure as to what type of entry this is—Proceedings, a paper presented at the workshop? Provide editor(s) name(s) if applicable, spell out ESCA, and provide publisher information.

Q26. Au: Pickering & Attardo: Provide the year of the draft you read instead of "in preparation," unless this has now been published—if so, update here and in the text. Provide publisher information regardless.

Q27. Au: Pickering & Levis (2002): Spell out TESOL, and provide the month this paper was presented.

Q28. Au: Quene (2006): Provide journal volume number and article page range.

Q29. Au: Robson & MackenzieBeck (1999): Provide name(s) of editor(s) of these proceedings, and provide name of publisher.

Q30. Au: Urios-Aparisi & Wagner (2007): Provide the month this paper was presented.

Discourse Processes, 46:1–24, 2009

ISSN: 0163-853X print/1532-6950 online
DOI: 10.1080/01638530902959604

Routledge
Taylor & Francis Group

Prosodic Markers of Saliency in Humorous Narratives

Lucy Pickering
Department of Applied Linguistics
Georgia State University

Marcella Corduas
Department of Statistical Science Q1
University of Napoli Federico II, Italy

Jodi Eisterhold and Brenna Seifried
Department of Applied Linguistics
Georgia State University

Alyson Eggleston
Department of Linguistics
Purdue University

Salvatore Attardo
Department of Literature and Languages
Texas A&M University

Much of what we think we know about the performance of humor relies on our intuitions about prosody (e.g., "it's all about timing"); however, this has never been empirically tested. Thus, the central question addressed in this article is whether speakers mark punch lines in jokes prosodically and, if so, how. To answer this question, this article unites both the recently emerged research agenda

---

Correspondence concerning this article should be addressed to Lucy Pickering, Department of Applied Linguistics, Georgia State University, P.O. Box 4099, Atlanta, GA 30302. E-mail: esllup@langate.gsu.edu Q2

grounding spoken discourse analysis in the precision and verifiability of acoustic analysis and a research agenda within the field of discourse and humor focused on the “performance” of humorous narratives. This article presents an analysis of a relatively simple form: the joke or short humorous narrative. The starting point of this analysis is the folk theory of joke-telling. Through instrumental measurement of pitch, volume, and speech rate, this study shows that, contrary to the folk theory of joke-delivery, punch lines are not delivered significantly louder than the preceding text, but rather at a significantly lower pitch and slower speech rate than the text preceding the punch line. In addition, punch lines are often, but not necessarily, signaled by a laughing voice or a smiling voice and are not preceded by significant pauses. This article concludes that the folk theory of joke-delivery is largely refuted. This study further investigates whether the saliency of punch lines, which would predict higher volume and pitch, is less significant than their final position in the narrative, which, being associated with final position in a *paratone*, or spoken paragraph, predicts that they will demonstrate lower volume and pitch values. The conclusion is that final positioning trumps the saliency of the punch lines and accounts for the significantly lower pitch and lack of significantly higher volume in punch lines.

The central question addressed in this article is whether speakers mark punch lines in jokes prosodically and, if so, how. A good example of the phenomenon is the old sexist joke, “Take my wife … please!” In this text, the punch line “please!” is preceded by a pause and presumably uttered with emphatic stress. Thus, the humor is said to be signaled by the prosodic features of intonation and pausing. Folk theories of humor performance suggest that this is the case for all jokes or all punch lines, and our research question is designed to test that intuition.

Central to distinguishing information structure in spoken discourse are prosodic features such as stress, intonation, rhythm and pause structure, all of which are used to distinguish between salient (i.e., standing out in relation to its context) and non-salient information (Brazil, 1997; Gumperz, 1982; Nakajama & Allen, 1993; Swerts & Geluykens, 1993, 1994). Experimental studies have indicated that prosodic features, such as pitch and pause structure, are used in the production and processing of local (utterance level) and global (discourse level) information structure (Cutler, Dahan, & van Donselaar, 1997; Grosz & Sidner, 1986). Swerts and Geluykens (1994, p. 38) found that discourse structure can be deduced from prosody and that both pause duration and pitch variation appear to be key perceptual cues. Collectively, this research suggests that speakers employ prosodic structure to organize information and that listeners use prosodic cues to parse in-coming information and predict upcoming discourse structure. Researchers have also found systematic differences in the patterning of cues across discourse genres such as interviews, conversations, and narratives (Chafe, 1994; Couper-Kuhlen & Selting, 1996; Kowal & O’Connell, 1980;

Kowal, Wiese, & O'Connell, 1983), supporting Swerts and Geluykens's (1994) conclusion that "this points to a very sophisticated use of global F0 [fundamental frequency] features by the speaker, and shows that we should look beyond the local level when studying the discourse function of F0 variation" (p. 31).

In this article, we investigate the ways in which prosodic variables may coincide with the structural organization of humorous narratives. Concerning prosody and narratives in general, Wennerstrom (2001) suggested that it is "an open field" (p. 219), and this is particularly the case with regard to acoustic analysis. In her own work using Labov and Waletzky's (1967) model of narrative Q4
structure in which she employed instrumental analysis of the acoustic variables, she proposed that there may be correlations between "structural components of narratives" and "characteristic prosodic features" (Wennerstrom, 2001, p. 204), including systematic uses of changes in tempo between the orientation (or setting) and complicating action (introduction of events, states, and people). Transitions between components of the narrative were also marked by pitch changes such that new components were marked by extra-high pitch resets creating *paratones*, or speech paragraphs, which then declined in pitch throughout their realization.

Also beginning from Labov and Waletzky's (1967) model, which was sig- Q5
nificantly modified, Chafe (1994) suggested that the climax (the unexpected event of the story) is usually presented with "bells and whistles" including prosodic effects such as dramatic pausing, heightened amplitude and pitch, and lengthening of initial or final consonants (p. 131). In contrast, he found that the coda, or final part of the narrative, was likely to demonstrate lower volume and pitch values.

Similar to the work on the prosodic features of non-humorous narrative, folk theories of humor performance have tended to focus on the characteristics of the climax or dramatic peak of the joke (i.e., the punch line). Most common, this has concerned matters of timing—specifically, the time delay (or pause length) between the end of the setup of the joke and the punch line:

> Timing can make the difference between a joke that is extremely effective and one that flops. Usually, timing relates to the delivery of the punch line.... Too short a time and the impact is lessened by the abrupt end of the joke. (Audrieth, 1998) Q6

Research on prosody and humorous narratives or jokes has also focused on the rate of delivery of the punch line itself with the suggestion that it has a "more rapid [and] fluid delivery" than the beginning of the text (Norrick, 2001, pp. 260–261). In addition to timing, a few researchers have briefly addressed other prosodic variables including pitch and volume characteristics. Norrick (2001) noted that punch lines may be indicated with a "voice shift" (p. 272). The term is undefined but presumably means some kind of perceived change in pitch. In an analysis of oral narratives in which practical jokes (as opposed to telling jokes directly) were described, Bauman (1986) reported that "dialogic punch lines are rendered in quoted speech which was often set-off by pauses and could sometimes involve altered voices with higher pitch, louder volume, and other paralinguistic features" (p. 66). As other research has suggested that reported speech is routinely marked with higher pitch and volume (Grosz & Hirschberg, 1992; Klewitz & Couper-Kuhlen, 1999; Wennerstrom, 1997, 2001), it is not clear that prosodic change can be directly attributed to the punch line status.

It is clear that this work is in its infancy and statements frequently reflect intuition rather than the observation of instrumentally based acoustic evidence. In this study, we undertook an instrumental computer-aided analysis of prosodic features of the spoken performance of humorous narratives. Using an independently motivated model of topic structure to identify prospective salient and non-salient material and focusing on the particular status of the punch line, we conducted a series of prosodic measures of the setups and punch lines to determine whether the pitch, volume, and pause characteristics typically associated with punch lines could be verified.

SALIENCY AND THE STRUCTURAL ORGANIZATION OF HUMOROUS NARRATIVES

Humor research has established that joke texts essentially consist of two parts. The first part is usually termed the "setup" and presents the characters of the narrative and develops it. The second part is commonly called the "punch line" and comprises a stretch of text that forces a revision of the interpretation of the text up to that point. Considerable discussion has gone into exactly how this "surprising," "sudden," and "unexpected" shift in interpretation and reinterpretation takes place (Attardo, 1994, 2001). It logically follows that the information in the punch line must be new, not previously introduced in the text, and extremely salient. There exists a "semiformal" procedure for locating the punch line of jokes, following Hockett (1973). The procedure consists of stripping away successive phrases progressively from the end of the text moving toward the beginning. When the humorous effect is destroyed, one knows that the last phrase deleted was the punch line. As noted in Attardo, Attardo, Baltes, and Petray (1994), this test is not foolproof insofar as there are exceptions. Selectional restrictions for some verbs, for example, will force the presence of complements in a final position and push the punch line in a pre-final position (e.g., "put the knife *in the drawer*"). Hence, Hockett's method is not limited to single-phrase punch lines, and there exist situations in which grammatical and distributional constraints force the presence of phrases after the punch line. In a situation in which a punch line stretches over two phrases, the larger syntactic unit in which the last phrase to be removed was located is considered as the punch line.

Thus, the test is defined as semiformal in the sense that one cannot apply it mechanically. Nonetheless, it is a very good heuristic, with a success rate of above 90% (Attardo et al., 1994).

Using this method, Attardo et al. (1994) confirmed the relative saliency of the punch line in an investigation of a corpus of 2,000 jokes, which found that the punch lines had a very strong correlation with the *rhematic*[1] position in the text consistent with the expectation that the punch line introduces new, unexpected, and salient material. Ninety-two percent of the jokes in the corpus had a punch line in the last phrase of the last sentence of the text (final punch line). The remaining 8% had a pre-final or non-final punch line. The material occurring after the punch line was found to be semantically empty or at least weak, consisting of repetitions of the punch line, identification of the speaker ("said the bartender"), explanations of the punch line, time and place adverbials, or other punch lines (in the case of jokes with multiple punch lines).

These findings are consistent with Giora's (2003) "saliency" model of jokes (p. 7) in which jokes are seen as a particular type of text that manipulates the ordinary processing of linguistic units—that is, normally, speakers choose the more salient information available; whereas, in jokes, the setup "deliberately" withholds the less salient meaning or interpretation, which will be revealed at the punch line and will require the "reinterpretation" of the text. In other words, within Giora's *graded salience hypothesis*, jokes always go from a salient to a less salient meaning (up to that point in processing); but, of course, in the process of shifting from the first to the second meaning, the second becomes extremely salient (and the first meaning is actually suppressed). This "sudden shift" view of joke processing is further empirically validated by Vaid, Hull, Heredia, Gerkens, and Martinez (2003).

This independently motivated analysis of the structure of the punch line permits the investigation of correlations between prosodic features of the discourse and the structure of the text while avoiding the common methodological problem of a "lack of independent analysis of the structure of the discourse under consideration" (Grosz & Hirschberg, 1992, p. 429). The rhematic, or focal, nature of the punch line allows us to make predictions regarding the prosodic features that speakers may utilize in their performance to cue the punch line.

[1] By *rhematic*, we refer to the functional sentence perspective terminology of the Prague school, which defines the *rheme* as the part of the sentence adding additional meaning. Needless to say, different terminology and conflicting definitions have been proposed for roughly the same concepts in different schools (e.g., Halliday's functional grammar and work by Prince, Chafe, Clark, Givòn, and many others; for discussions and references, see Attardo, 1994, pp. 161–185).

Based on a review of prosodic markers of focus in the literature, the features investigated in this study were pitch, volume, rate of speech, length of pauses, and voice quality.

PROSODIC CUES AND SALIENCY

Pitch

Considerable work has established that "focus" in inner circle varieties of English[2] indicating new information, contrastiveness, emphasis, or saliency[3] is FN2,FN3 marked by a speaker using high pitch (Brown, 1983; Brown & Yule, 1983; Chafe, 1994; Halliday, 1967; Ladd, 1996; Nakatani 1993, 1997; Selting, 1994; Swerts & Geluykens, 1994; Wennerstrom, 2001; Xu & Xu, 2005). Hence, we would anticipate, in line with previous auditory rather than acoustic research, that the punch line will exhibit higher pitch values. An important complication arises, however, when considering the pitch effects that occur in discourse as opposed to isolated utterances. As noted earlier, experimental studies have identified a paratone, or spoken paragraph unit, in discourse that is marked by a high pitch onset (as measured by F0) and a low pitch close (Brown, Currie, & Kenworthy, 1980; Lehiste, 1979; Pickering, 2004; Wennerstrom, 1997; Yule, 1980). Within this unit there is likely to be a "gradual descent" or declination in pitch such that discourse final F0 peaks occur at a lower absolute pitch (F0) than initial peaks (Ladd, 1996; Nakajima & Allen, 1993; Tench, 1996; Wichmann, 2000). This prosodic unit is connected to rhetorical structure and the opening and closing of a speech event.

Although punch lines typically occur at the end of the text and, therefore, may be predicted to typically close a paratone, it is also conceivable that their saliency may result in a pitch reset and the initiation of a new paratone; so, it is unclear how this kind of discourse prosodic organization may affect pitch values in relation to this salient material.

Volume

In addition to pitch perturbations, focus is frequently associated with "loudness peaks" (Wells, 1986, p. 61). Researchers have also suggested that exaggerated volume may co-occur with markers of emphasis or particularly significant elements of a narrative (Brown et al., 1980; Chafe, 1994; Eggins & Slade, 1998;

---

[2]As opposed to outer circle or nativized varieties of English, such as Indian English, which demonstrate different prosodic characteristics (Gumperz, 1982).

[3]We do not claim that new information, contrastiveness, emphasis, and saliency are the same thing; merely that they correlate with the same type of markers, as found in the literature mentioned.

Lehiste, 1979; Selting, 1994; Wennerstrom, 2001; Yule, 1980). Again, however, an underlying discourse paratone structure may also affect increased amplitude as volume tends to decline over the length of the paratone (Ladd, 1996; Nakajima & Allen, 1993; Tench, 1996; Wichmann, 2000).

Speech Rate

Speech rates reported for adults speaking English typically range from 3.47 to 5.70 syllables per second depending on speaking condition (Munro & Derwing, 1994; Pickering & Levis, 2002). Rates for specific genres have been identified across languages. In a meta-analysis of studies focusing on two genres (interviews and story-telling) in five languages (English, Finnish, French, German, and Spanish), Kowal et al. (1983, p. 386) found an average speech rate in story-telling of 3.43 (syllables per sec) and of 4.31 in interviews. In addition to changes in overall speech rate related to genre, experimental findings have suggested that "important or unpredictable material" is spoken at a relatively slower rate than less salient material (Quené, 2006, p. 3). Koopmans-van Beinum and van Donzel (1996, p. 2) confirmed this finding for spontaneous speech in Dutch, finding that "important information" is delivered at a slower rate than other, less important, parts.

In relation to narrative structure, Wennerstrom (2001, p. 203) reported a slowing down of tempo with the shift from orientation to complicating action, but did not directly address speech rates at the climax of the narrative (presumably roughly equivalent to the punch line). Griffiths (1991, p. 348) suggested that slower speech increases comprehension; therefore, a further prediction may be that a speaker's rate would slow down to maximize comprehensibility for the listener.

Pause Structure

Pausological research has typically distinguished between silent and filled pauses (Ballmer, 1980); and, in what follows, we consider only silent pauses. In a comprehensive review of the field, Griffiths (1991, p. 346) proposed both a "cutoff point" (a shortest measurable pause of 0.1 s or 100 ms) and a "threshold level" of 3 s (a pause so long that one can assume the speaker has stopped speaking.) Research has primarily been concerned with examining whether the length of silent pauses can be systematically tied in any way to syntactic and, by extension, topic structure. Experimental studies related to this question have suggested that longer pauses (greater than 0.8 s) mark new topics and occur between topics (Brown et al., 1980; Swerts & Geluykens, 1994; Zellner, 1994). Brown et al. and Brown and Yule (1983) developed a model of pause-defined units in discourse in which they identified three major groups: "topic pauses" of

0.8 s or longer, which “clearly coincide with major semantic breaks”; “substantial pauses” of between 0.6 and 0.8 s and pauses of between 0.4 and 0.6 s, which coincided with shorter units; and “very short pauses” (between 0.2 and 0.4 s), which co-occurred with incomplete syntactic structures (pp. 56–57). Although Brown et al. did not directly address humorous narratives, their work was with spontaneous speech, and this model is taken as a first approximation for pause analysis in the data. As noted earlier, timing in relation to the pause length between the end of the setup of the joke and the punch line is an area that has received some considerable attention in traditional discussions of humor performance (here intended in the sense of stage performance, rather than the linguistic sense of performance as opposed to competence). The consensus of this folk theory of humor performance has been that a substantial pause, which we have conceptualized as between 0.6 and 0.8 s (Brown et al., 1980, p. 56), should precede the punch line.

Laughter

As is well-known, laughter is a significant cue for humor (Attardo, 1994; Glenn, 2003; Jefferson, 1979; Norrick, 1993). In addition to—or in place of—laughter, speakers may employ a smiling voice (SV) or a laughing voice (LV). This is a commonly used (e.g., Jaffe & Walton, 2000; Wooffitt, 2001) impressionistic label describing some effects of the latitudinal adjustment of the vocal tract in smiling, which has clear effects on the acoustics of speech (Marasek, 2006; Robson & MackenzieBeck, 1999; Tartter, 1980; Zacher & Niemitz, 2003). It is often difficult to differentiate between a SV and a LV, as a SV may be regarded as weakened laughter. In this analysis, we treated all cases of actual laughter as a LV (either within-utterance laughter [WUL] or post–utterance laughter [PUL]) and all voice quality modifications as a SV, and investigated the ways in which this paralinguistic variable interacted with the prosodic variables that we identified.

In light of the aforementioned literature review, we take as a central hypothesis that two parts are identifiable in a joke or humorous narrative (i.e., the setup and the punch line), and that prosodic acoustic cues may function as contextualization cues for these structural elements. Given the focal or rhematic nature of the punch line, we make the following predictions, based on the literature:

1. Punch lines will demonstrate higher pitch peaks and higher loudness peaks than those found in the non-focal part or setup of the narrative. Given the possible competition with a discourse level paratone structure, this hypothesis has a weak degree of expectation.
2. Punch lines will demonstrate a rate of speech significantly slower or faster than the rest of the text.

3. Punch lines will be preceded or followed by substantial or topic pauses (i.e., 0.6 s or longer).
4. Punch lines will be marked by a LV, a SV, or both. Laughter may occur during or after the utterance.

METHOD

Participants and Data

The data used in this study were derived from a set of video recordings made of undergraduate students attending classes at Georgia State University, telling two jokes. For the first joke, the students were given, in writing, a joke to perform (referred to as "the engineer joke," a representative version of which is shown in the Appendix). In this case, the students recalled the text from memory, rather than acted or read out loud, discourse modes that can have can have somewhat different acoustic properties (Brazil, 1997; Johns-Lewis, 1986). As this joke involved a "recreation" of the text, from partial memory cues, there was naturally some variation in how different the versions of the text were, and these issues are discussed later. The time elapsed between the distribution of the engineer joke and the taping varied between a couple of days to 1 week, due to scheduling issues. No instances of students having forgotten the assignment arose during collection. Immediately after they performed the engineer joke, the students were asked, without prior warning, to perform an additional joke of their choice. No student failed to perform on this task. As sound quality was the primary determinant for the data used for this project to effectively use instrumental analysis, recordings that exhibited high levels of ambient noise or problematic voice quality, such as a creaky voice or vocal fry, were discarded prior to the conceptualization of the research questions by the authors. Thus, our data comprised a set of 10 students, each performing the engineer joke and a spontaneous joke of their choice, resulting in a data set of 20 jokes. No demographic data were collected on these participants beyond their undergraduate status, and our sample consisted of 2 male speakers and 8 female speakers.

Although joke-telling is a commonplace activity in which most speakers engage with regularity, this was clearly an experimental situation that could potentially have resulted in unnatural or "bad" joke-telling.[4] Thus, we enrolled FN4

[4]Studies have found that the presence of a camera inhibits laughter, but not smiling (Martin, Sadler, Barrett, & Beaven, in press). We believed that the presence of the interviewer (a graduate student who was working with them) would, however, counterbalance any inhibitory effect the camera may have had because it has been established that the presence of others leads to increased smiling (Dale, Hudak, & Wasikowski, 1991; Fridlund, 1991) and laughing (Freedman & Perlick, 1979). Q8

TABLE 1
Ratings of the Performance of the Joke-Tellers

Sample No.	*Judge 1*	*Judge 2*
1	3	3
2	3	3
3	4	3
4	4	3
5	2	3
6	3	3
7	2	2
8	3	2
9	2	3
10	3	3
11	4	4
12	4	4
13	3	3
14	3	3
15	3	3
16	3	4
17	3	3
18	2	2
19	2	2
20	3	3

two independent judges, both humor researchers, who were unaware of our work, to rate the performance of the joke-tellers on a Likert scale, ranging from 1 (*very bad*) to 5 (*very good*), where 3 (*average*) was the median. Scores are shown in Table 1. Mean scores were 2.95, indicating that that the performances, as a group, were rated as average. The standard deviation was .68 and .60 for each judge, respectively, revealing very little variation: In fact, there were no “very bad” or “very good” marks. We calculated a Cohen (1960) kappa statistics of interrater reliability, which showed a “substantial” (Landis & Koch 1977) agreement (.62).[5]

Procedures

For each humorous narrative, we took a series of measurements of pitch, volume, speech rate, and pause, and noted voice quality characteristics. These procedures were designed to illuminate any systematic prosodic differences between the

[5]We thank three anonymous reviewers and the participants of the 10th International Pragmatics Conference for a number of very helpful suggestions, including this one. We also thank Victor Raskin and Christian Hempelmann for their help.

punch line and the setup of the joke. Video recordings were transferred to a Kay Pentax Computerized Speech Laboratory (CSL) Model 4300b to make Q10 instrumental measures. Each text was divided into intonational phrases (Brazil, 1997; Brown et al., 1980; Wennerstrom, 2001). For each intonational phrase, the following measures were made:

1. Pitch maxima and minima: Calculated by selecting the intonational phrase from the soundwave, performing a pitch analysis using the pitch contour analysis function of the CSL, and then computing result statistics within that specified time domain. These hertz (Hz) values were confirmed by manually measuring the pitch contour using cursors.
2. Volume maxima and minima: Calculated by selecting the intonational phrase from the soundwave, performing an energy analysis using the energy contour analysis function of the CSL, and then computing result statistics within that specified time domain. These decibel (dB) values were confirmed by manually measuring the energy contour using cursors.
3. Pause length: Calculated by selecting the area of the soundwave surrounding the pause, performing a spectrographic analysis, and manually measuring the length of the pause using cursors.
4. Speech rate: Calculated by dividing the overall time taken to produce the sample, manually measured using cursors, by the number of syllables in the sample.

The punch line for the joke was distinguished from the setup using the procedure identified earlier (Hockett, 1973). For the purposes of the acoustic analysis, we considered the entire intonational phrase in which the punch line occurred to be the punch line—that is, for the engineer joke specifically, if the word *frog* and the expression “that’s cool” occurred in separate intonational phrases, we considered the two intonational phrases to be the punch line, to guard against potential difficulties in identifying the single punch line phrase. In the engineer jokes, punch lines occurred in the final intonational phrase in all but four cases in which the punch line was potentially split over the last two intonational phrases:

/but a frog that talks/

(0.12)

/now that’s cool/

There were no cases of punch lines split over two intonational phrases in the spontaneous jokes. The jokes were also coded for SV, WUL, and PUL.

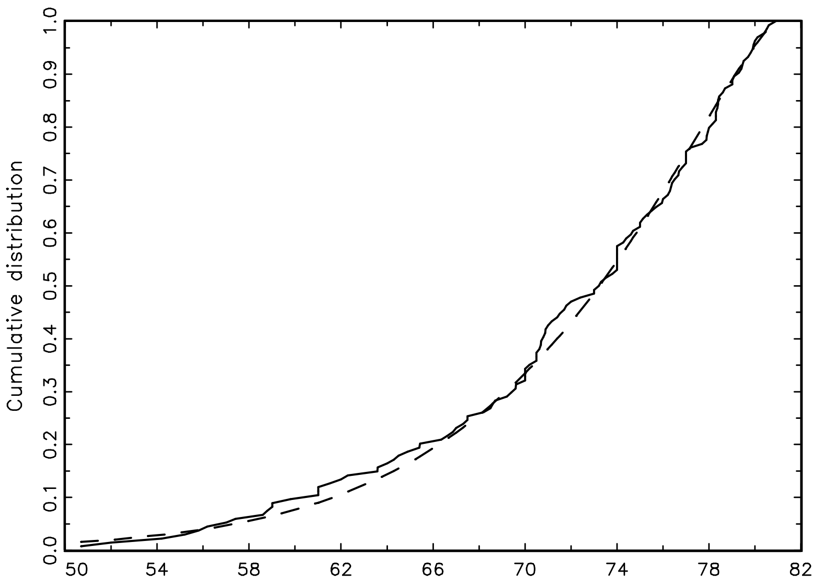

FIGURE 1 Comparison of the Weibull Max Distribution (dotted line) and current data for volume (continuous line).

RESULTS

Initial investigation of a sample showed that a histogram of the data did not have a bell-shaped distribution; rather, as it often happens in the case of variables representing extremes, the data were found to be closely approximated by the Weibull Max Distribution[6] (Johnson & Kotz, 1970, p. 272). Figure 1 presents FN6,Fig1
the hypothetical cumulative Weibull Max Distribution (the dotted line) and the sample distribution function of decibel measures from our data (the parameters of the fitted Weibull Max Distribution were estimated by maximum likelihood). As the reader should note, the match is extremely close, and the assumed distribution function was accepted as correct according to the Kolmogorov–Smirnov test. Q11

Hence, we used nonparametric tests in investigating the characteristics of our data. Because we were comparing paired data (the setup and the punch line of a text or the same speaker telling different jokes), we selected the Wilcoxon signed-ranks test to test whether the two data distributions differ. This test Q12

[6]The Weibull Max distribution (named after Wallodi Weibull) describes the distribution of maximum values in independent samples. It is widely used in weather modeling and in industrial sampling.

provided a useful nonparametric alternative to the "classical" test of location defined for observations from normal distributions. In the following sections, we examine the findings for each of the prosodic variables under investigation. In particular, we report the significance probability (p value) of the test statistic; the sample is consistent with the stated (null) hypothesis if the significance probability is large and provides evidence against such a hypothesis if the significance probability is small, with respect to a defined value.

Pitch

Table 2 shows the averages of the pitch maxima measured in the 20 texts. Tab2
Setup pitch measures show the average height of pitch maxima found in the intonational phrases that comprised the setup portion of each text. Punch line measures show the pitch maxima of the intonational phrase containing the punch line. Regarding the engineer joke specifically, if the two phrases "the frog" and "that's cool" were in two separate intonational phrases, the measurement of the pitch maxima in each phrase was averaged. Comparison of the pitch measurements in the setup and at the punch line demonstrated that the punch line peaks

TABLE 2
Average Pitch Peaks in the Texts (in Hertz)

Sample	*Setup*	*Punch Line*
1	283.06	253.00
2	299.55	261.00
3	279.72	274.00
4	281.17	296.00
5[a]	150.00	113.00
6[a]	157.50	170.00
7	268.62	278.00
8	288.75	248.00
9	268.85	221.50
10	243.80	219.00
11	331.53	241.50
12	336.50	331.00
13	270.40	254.00
14	299.14	332.00
15	308.00	281.46
16	284.00	274.00
17	312.69	246.00
18	318.35	251.00
19[a]	138.00	109.00
20[a]	140.07	137.00

[a]Male speakers.

were significantly lower in pitch by an average of 23.46 Hz (Wilcoxon matched pairs signed-ranks test, 2-tailed $p < .0037$). Thus, we concluded that punch lines were significantly lower in pitch than the setup of the joke. Consistent with some current analyses of pitch distribution (Couper-Kuhlen & Ford, 2004), we confirmed this finding using semitone data rather than raw F0 measurements. We converted the measurements to semitones using the standard formula in Burns and Ward (1982)—$[12/\log(2)] \times [\log(\max F0/\min F0)]$—and performed Q13 the same statistical test. Again, punch lines were, on average, 1.73 semitones lower than the setup, and this difference was statistically significant (Wilcoxon matched pairs signed-ranks test, 2-tailed $p = .0027$).

Finally, we performed the same test to determine whether there were any significant differences between the engineer joke and the spontaneous joke regarding this measure. For each speaker, we compared the measurements for the engineer joke punch line and the measurements for the spontaneous joke. The data were found to be consistent with the null hypothesis of no differences between the two jokes (Wilcoxon signed-ranks test, 2-tailed $p = .1055$). In other words, the pitch findings do not appear to be a function of a specific grammatical structure. Overall, we conclude that the pitch measurements are consistent with a paratone structure in which pitch peaks show a global pattern of pitch declination (with the possibility of local peaks) throughout the text (Swerts & Geluykens, 1994; Wichmann, 2000).

Volume

Table 3 shows the measurements of the averages of volume maxima in the Tab3 intonational phrases in the 20 texts. Similarly to the pitch measurements, setup averages are given in the first column and punch line values in the second. If the punch line occurred across two intonational phrases, the average of the two phrases' peaks was taken. The results show that punch lines were produced at a slightly higher amplitude, which was, on average, 0.2 dB higher in volume; however, the difference was not statistically significant (Wilcoxon signed-ranks test, 2-tailed $p = .6507$). We conclude that the evidence in the samples does not support the hypothesis that punch lines are delivered at a higher volume than the rest of the body of the narrative.

Again, we considered the possibility that findings would vary for the same speaker between the engineer joke and the spontaneous joke; however, according to the Wilcoxon signed-ranks test, no significant difference between these two groups was found (2-tailed, $p = .0547$); although, on average, the spontaneous joke was 2.2 dB louder. In agreement with the findings regarding pitch structure, these results do not support the statement that loudness peaks will be higher in the punch line due to its rhematic nature.

TABLE 3
Mean Volume Peaks in the Texts (in Decibels)

Sample No.	*Setup*	*Punch Line*
1	77.37	76.50
2	77.22	76.00
3	78.09	81.00
4	79.76	81.00
5	74.76	73.00
6	76.50	75.00
7	79.93	83.00
8	81.00	81.00
9	73.92	71.00
10	71.08	77.00
11	71.69	69.00
12	71.50	70.00
13	57.87	59.00
14	58.28	61.00
15	74.93	76.50
16	77.11	81.00
17	69.53	64.00
18	69.21	72.00
19	70.45	68.00
20	69.64	69.00

Pause Structure

In the literature on humor, the pause before the punch line has essentially been defined as a rhetorical pause, to set-off the punch line. As an initial approximation, we operationalized this category of rhetorical pauses as equivalent to "substantial" pauses (i.e., those between 0.6 and 0.8 s). However, it was also possible that an objectively short pause could be considered substantial if compared to even shorter pauses. In other words, a hypothetical teller could mark a punch line with a pause relatively longer than his or her usual pause length. Thus, we took two pause measures: For each joke by each speaker, we measured the mean length of all the pauses in the telling of the jokes and compared it to the mean length of the pauses in or before the punch line (if the punch line consisted of a single intonational phrase, we included the pause in the punch line; if there were two intonational phrases, we included the pauses before both intonational phrases, plus any other pauses within the intonational phrases). Second, we investigated whether substantial pauses occurred before the punch lines.

Results of mean length of pauses are shown on Table 4. Statistical analysis Tab4
showed that the mean length of pauses in the punch line was shorter by 0.094 s,

TABLE 4
Mean Length of Pauses in the Texts

Sample No.	*Setup*	*Punch Line*
1	0.41	0.17
2	0.34	0.18
3	0.34	0.52
4	0.44	0.45
5	0.49	0.14
6	0.43	0.18
7	0.74	1.00
8	0.48	0.65
9	0.53	0.38
10	0.51	0.37
11	0.41	0.23
12	0.27	0.44
13	0.67	0.59
14	0.63	0.59
15	0.51	0.14
16	0.57	0.23
17	0.70	0.10
18	0.78	1.25
19	0.47	0.47
20	0.55	0.31

and that this result was not statistically significant (Wilcoxon signed-ranks test, 2-tailed $p = .1447$). Hence, we concluded that the hypothesis that punch lines are set-off by relatively longer pauses was falsified.

To complete the second pause measure, we calculated pauses appearing before the intonational phrases containing the punch line. Inspection of the results in Table 5 shows that a substantial pause appears only in Samples 7, 8, and 18, Tab5 with Samples 13 and 14 coming very close (within a margin of measurement error). Hence, we conclude that the hypothesis that punch lines are preceded by significant pauses is rejected; in fact, 15 jokes (75% of the texts) did not have a significant pause before the punch line.[7] FN7

Moreover, our measure of substantial pauses overstated the case in favor of the pause-as-marker hypothesis. Some of the pauses have nothing to do with rhetorical marking but, rather, are performance issues. In Sample 7, for example, the speaker appeared confused and hesitant prior to the 1-s pause that marked the

[7] It should be noted that the folk theory of joke-telling makes the prediction that the punch line, in the Hockett (1973) definition, is preceded by a significant pause, whereas we have examined the hypothesis that the intonational phrase in which the punch line occurs is preceded by a significant pause. Both hypotheses end up giving the same results (see Attardo & Pickering, in press). Q14

TABLE 5
Pauses (in Seconds) At and Around Intonational Phrases Containing the Punch Line

No.	*Transcription of the Punch Line Intonational Phrases*
1	(0.16)/but now a frog that talks/(0.18)/That's really cool/
2	(0.18)/I thought you said drinks were on the house/
3	(0.52)/but I do have time for a talking frog that's pretty cool/
4	(0.45)/and so God goes alright do you want that bridge to be two lanes or four/
5	(0.14)/but a talking frog now that's a whole 'nother thing/
6	(0.18)/and the priest says that's a good idea but do you think we have time/
7	(1.0)/now a frogging a talking frog now that's cool/
8	(0.65)/I'm all for it/
9	(0.21)/but a talking frog/(0.55)/That's cool/
10	(0.37)/so the elephant picks it up and wipes its ass with it/
11	(0.23)/a talking frog now that's cool/
12	(0.44)/my fri- my friend is on his knees swearing to help prevent forest fires/
13	(0.59)/but a talking frog now that's cool/
14	(0.59)/shut up, you're next/
15	(0.17)/but a frog that talks/(0.12)/now that's cool/
16	(0.23)/and when he went soaring over the edge he yelled shit[hh]/
17	/a talking frog/(0.10)/now that's interesting/
18	(1.25)/the blond guy looked at them and was like I don't pack my wife's lunch/
19	(0.47)/I like the fact that you're just a talking frog and that's cool/
20	(0.31)/he went to your house (hhh)/

punch line. In fact, he incurred an error ("frogging" for "talking") immediately after the pause, as shown:

/I don't have time for a girlfriend/(1.00)

/Now a frogging, a talking frog, now *that's cool/*

Finally, a comparison of Table 1 and Table 5 shows that the jokes with the longer pauses before the last sentence or before the intonational phrase with the punch line were not those ranked higher in quality of performance by the judges. In fact, those that received the highest quality rankings did not present substantial pauses in the vicinity of the punch line at all. In conclusion, the hypothesis that punch lines are marked by substantive pauses is rejected.

Speech Rate

Table 6 displays the rates of speech in the setups with the rate of speech in the punch lines for the narratives. Speakers tended to deliver the punch line of the joke at a slightly slower rate (on average, about 0.03 s) than the setup in both

TABLE 6
Rates of Speech in the Texts

Sample No.	*Setup*	*Punch Line*
1	4.41	4.34
2	4.21	6.31
3	5.45	5.18
4	3.37	5.00
5	4.17	6.00
6	3.76	5.51
7	3.95	3.66
8	3.44	3.07
9	4.42	2.91
10	3.53	5.18
11	5.14	3.18
12	8.90	5.71
13	3.79	3.33
14	3.59	2.00
15	4.10	2.58
16	3.67	3.81
17	4.09	5.38
18	3.33	3.44
19	3.91	6.11
20	4.07	2.03

types of jokes; however, the difference was not statistically significant. As a tendency was noted, and as the spontaneous jokes tended to be shorter than the engineer joke, this suggested that each speaker may demonstrate a “personal” joke-telling rate. This result provides some initial support for proposals made by Griffiths (1991), Koopmans-van Beinum and van Donzel (1996), and Quené (2006) that utterances containing important or unpredictable material will be produced at a slower speech rate than less salient material.

Laughter

We considered one paralinguistic measure of voice quality: LV and SV. Table 7 shows that 12 of the 20 performances we examined had some type of paralinguistic “humor” marker (i.e., LV or SV). The relative incidence revealed that PUL pulses were the most frequent (6 cases), followed by SV (5 cases) and WUL (2 cases). Two examples, by the same speaker, included both the LV and SV. Regarding this feature, our data were too sparse to perform a statistical analysis of the significance of the presence of paralinguistic markers, but occurrences in 60% of the data suggests at least an indicator of a preference.

TABLE 7
Paralinguistic Markers

Sample No.	*Paralinguistic Markers*
1	SV
2	SV
3	WUL
4	PUL
5	
6	
7	PUL
8	
9	
10	
11	PUL
12	
13	PUL
14	WUL
15	WUL and SV
16	PUL and SV
17	
18	SV
19	
20	PUL

Note. SV = smiling voice; WUL = within-utterance laughter pulses; PUL = post-utterance laughter pulses.

Reported Speech

In one final issue, we addressed the potentially confounding variable of reported speech. Reported speech had previously been connected to the production of a higher pitch and volume and a faster rate in punch lines (see Bauman, 1986), as compared to punch lines that do not contain reported speech. Due to the nature of our data (i.e., the engineer joke has a punch line that occurs in reported speech) and the random nature of the selection of the improvised jokes, we had only two jokes in our corpus in which the punch line did not occur in reported speech. Because most of our samples comprised reported speech punch lines, and we found that the punch lines in our corpus had lower pitch and comparable volume, the hypothesis that punch lines that occur with reported speech are more likely to display higher pitch and volume and a faster rate is not supported by our results.

To investigate this further, we contrasted the two jokes in which the punch line did not occur in reported speech (Samples 10 and 12) with those in which it did. We tested if the distribution of the average volume of the punch lines of the two non-reported speech jokes were equal to the distribution of the average volume of the spontaneous jokes with punch lines in reported speech. The two-tailed Mann–Whitney test provided evidence in favor of this hypothesis ($p = .9474$; Q15 i.e., that there was no significant differences between the two means). A similar conclusion was drawn when testing for pitch differences ($p = .7579$). Finally, we tested whether there was a difference between the rates of the punch lines; but, again, we found that the two-tailed Mann–Whitney test concluded in favor of no difference between the groups ($p = .6737$). Despite these findings, given the very restricted part of the sample to which it is applicable, this conclusion should be taken as tentative.

DISCUSSION

In light of the presented results, we return to our initial hypotheses. We find no support for our first prediction that, due to their rhematic nature, punch lines will demonstrate higher pitch peaks than those found in the non-focal part of the setup of the narrative. Not only was this hypothesis disproved, the opposite was found to be true—namely, that punch lines are produced at a significantly lower pitch. The hypothesis also predicted that punch lines should be delivered at higher volumes than the setup. We found that punch lines were produced at a very slightly higher volume, but the difference was not statistically significant, hence leading to the rejection of this hypothesis as well. We suggest that most likely interpretation of these finding is that pitch declination related to the paratone structure of the text takes precedence over the rhematic character of the punch line. To put it differently, in our data, final paragraph level pitch structures trump saliency. The second prediction, that punch lines should have a rate of speech significantly faster or slower than the rest of the text, is not borne out by our data. However, there was a tendency toward a slower rate, which opposes some predictions in the literature on humor (Norrick, 2001) claiming that punch lines will be delivered faster than the body of the joke.

The third hypothesis, that an emphatic or substantial pause should precede the punch line, is also unsupported. We found that only 25% of the jokes analyzed here showed a pause immediately before the punch line. In general, few substantial pauses occurred in the samples. This surprising result cannot be explained by prosodic means. It is possible that the idea that punch lines should be set apart by pauses comes from professional comedians or the folk theory of humor (for a discussion, see Pickering & Attardo, XXXX). If this is the case, data Q16 from professional humor performers may present significantly different results in terms of pause structure. Preliminary data, however, suggest that this is not the case (Urios-Aparisi & Wagner, 2007) because situation comedy data, which are performed by professionals, also do not show pauses before punch lines. We are engaged in further research to investigate whether professional comedians delivering jokes display consistent patterns of pausing or other prosodic features. Finally, we determined that there is a preference for marking punch lines with a LV and a SV. However, these also occur elsewhere in the joke text (e.g., Jefferson, 1979), so this is, at best, a weak correlation. As punch lines also occurred without this paralinguistic marking, we conclude that, although there is a slight preference for marking punch lines with paralinguistic features, it is without predictive power (i.e., the presence or absence of paralinguistic features of humor is not significant either in identifying a punch line or in identifying a non-punch line).

Although our corpus of 10 speakers performing two texts each compares quite favorably with most of the literature investigating prosodic cues and textual organization where the norm is very few speakers and very few texts (e.g., Grosz & Hirschberg, 1992, who used 1 speaker performing 3 texts)[8], it is clear that broader studies using a more varied sample of jokes, more speakers, and focusing on other aspects of the texts will reveal other significant facts, such as differences in story-telling styles, which may demonstrate unique prosodic effects (e.g., Wennerstrom's, 2001, "deadpan story telling style," p. 218). Additional research is also needed to address the generalization issue of our data. These were collected from a limited pool of North American students, and represented only one regional community and one generation, who possibly demonstrate specific stylistic conventions in joke-telling. A significant aspect of this study is that it initiates a research agenda concerning narrative genres and jokes and instrumentally based prosodic analysis; for example, it is possible that anecdotes, which are organized differently from a textual point of view, may show different patterns. Similarly, other texts present punch line-like features (e.g., suspenseful narratives). We plan to go on to compare these kinds of texts to find commonalities and differences in the organization of information and its prosodic correlates. This research is already underway.

REFERENCES

Attardo, S. (1994). *Linguistic theories of humor*. Berlin: de Gruyter.

Attardo, S. (2001). *Humorous texts*. Berlin: de Gruyter.

Attardo, S., Attardo, D., Baltes, P., & Petray, M. (1994). The linear organization of jokes: Analysis of two thousand texts. *HUMOR: International Journal of Humor Research, 7*(1), 27–54.

Audrieth, A. L. (1998). *The art of using humor in public speaking*. Retrieved December 20, 2006, from http://www.squaresail.com/auh.html

Ballmer, T. T. (1980). The role of pauses and suprasegmentals in grammar. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech* (pp. 211–220). Berlin: de Gruyter.

[8] An exception to this is Brown, Currie, and Kenworthy (1980), who utilized over 100 speakers.

Bauman, R. (1986). *Story, performance, and event. Contextual studies of oral narrative*. Cambridge, England: Cambridge University Press.

Brazil, D. (1997). *The communicative value of intonation in English*. Cambridge, England: Cambridge University Press.

Brown, G. (1983). Prosodic structure and the given/new distinction. In A. Cutler & D. R. Ladd (Eds.), *Prosody: Models and measurements* (pp. 67–77). Berlin: Springer.

Brown, G., Currie, K., & Kenworthy, J. (1980). *Questions of intonation*. Baltimore: University Park Press.

Brown, G., & Yule, G. (1983). *Discourse analysis*. Cambridge, England: Cambridge University Press.

Chafe, W. (1994). *Discourse, consciousness, and time. The flow and displacement of conscious experience in speaking and writing*. Chicago: University of Chicago Press.

Couper-Kuhlen, E., & Ford, C. E. (Eds.). (2004). *Sound patterns in interaction.* Amsterdam: Benjamins.

Couper-Kuhlen, E., & Selting, M. (1996). Toward an interactional perspective on prosody and prosodic perspective on interaction. In M. Selting & E. Couper-Kuhlen (Eds.), *Prosody in conversation* (pp. 11–56). Cambridge, England: Cambridge University Press.

Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech, 40*, 141–201.

Dale, J. A., Hudak, M. A., & Wasikowski, P. (1991). Effects of dyadic participation and awareness of being monitored on facal action during exposure to humor. *Perceptual and Motor Skills, 73*, 984–986.

Eggins, S., & Slade, D. (1998). *Analysing casual conversation*. London: Cassell.

Freedman, J. L., & Perlick, D. (1979). Crowding, contagion, and laughter. *Journal of Experimental Social Psychology, 15*, 295–303.

Fridlund, A. J. (1991). The sociality of solitary smiling: Potentiation by an implicit audience. *Journal of Personal and Social Psychology, 60*, 229–240.

Giora, R. (2003). *On our mind: Salience, context, and figurative language.* Oxford, England: Oxford University Press.

Glenn, P. (2003). *Laughter in interaction*. Cambridge, England: Cambridge University Press.

Griffiths, R. (1991). Pausological research in an L2 context: A rationale, and review of selected studies. *Applied Linguistics, 12*, 345–364.

Grosz, B. J., & Hirschberg, J. (1992). Some intonational characteristics of discourse structure. In *Proceedings of the 2nd international conference on Spoken Language Processing* (pp. 429–432). Banff. Q18

Grosz, B. J., & Sidner, C. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics, 12*, 175–204.

Gumperz, J. (1982). *Discourse strategies*. Cambridge, England: Cambridge University Press.

Halliday, M. A. K. (1967). Notes in transitivity and theme in English. Part 2. *Journal of Linguistics, 3*, 199–244.

Hockett, C. F. (1973). Jokes. In M. E. Smith (Ed.), *Studies in linguistics in honor of George L. Trager* (pp. 153–178). The Hague, Netherlands: Mouton.

Jaffe, A., & Walton, S. (2000). The voices people read: Orthography and the representation of non-standard speech. *Journal of Sociolinguistics, 4*, 561–587.

Jefferson, G. (1979). A technique for inviting laughter and its subsequent acceptance declination. In G. Psathas (Ed.), *Everyday language. Studies in ethnomethodology* (pp. 79–96). New York: Irvington.

Johns-Lewis, C. (1986). Prosodic differentiaion of discourse modes. In C. Johns-Lewis (Ed.), *Intonation in discourse* (pp. __–__). London: Croom Helm. Q19

Johnson, N. L., & Kotz, S. (1970). *Continuous univariate distributions—1*. New York: Wiley.

Klewitz, G., & Couper-Kuhlen, E. (1999). Quote–unquote? The role of prosody in the contextualization of reported speech sequences. *Interaction and Linguistic Structures, 12.* Retrieved December 20, 2006, from www.ub.uni-konstanz.de/kops/volltexte/2000/464/pdf/464_1.pdf Q20

Koopmans-van Beinum, F., & van Donzel, M. (1996). Discourse structure and its influence on local speech rate. *Proceedings of the Institute of Phonetic Sciences, 20*, 1–11.

Kowal, S., & O'Connell, D. (1980). Pausological research at Saint Louis University. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech* (pp. 61–66). Berlin: de Gruyter.

Kowal, S., Wiese, R., & O'Connell, D. (1983). The use of time in storytelling. *Language and Speech, 26*, 377–392.

Labov, W., & Waletzky, J. (1967). Narrative analysis: Oral versions of personal experience. In I. J. Helm (Ed.), *Essays on the verbal and visual arts* (pp. 12–44). Seattle: University of Washington Press.

Ladd, R. (1996). *Intonational phonology*. Cambridge, England: Cambridge University Press.

Landis, J. R., & Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159–174.

Lehiste, I. (1979). Sentence boundaries and paragraph boundaries—Perceptual evidence. In *The elements: A parasession on linguistic units and levels* (Vol. 15, pp. 99–109). Chicago: Chicago Linguistics Society. Q21

Marasek, K. (2006). *EGG and voice quality.* Retrieved December 20, 2006, from http://www.ims.uni-stuttgart.de/phonetik/EGG/frmst1.htm Q22

Martin, G. N., Sadler, S. J., Barrett, C. E., & Beaven, A. (in press). Measuring reponses to humor: How the testing context affects individuals' reaction to comedy. *HUMOR: International Journal of Humor Research*. Q23

Munro, M., & Derwing, T. (1994). Evaluation of foreign accent in extemporaneous and read material. *Language Testing, 11*, 253–256.

Nakajima, S., & Allen, J. (1993). Prosody as a cue for discourse structure. In *Proceedings of the 2nd international conference on Spoken Language Processing* (pp. 425–428). Banff. Q24

Nakatani, C. H. (1993). Accenting on pronouns and proper names in spontaneous narrative. In *ESCA workshop on Prosody* (pp. 164–167). Lund: Department of Linguistics. Q25

Nakatani, C. H. (1997). Discourse structural constraints on accent in narrative. In J. P. H. van Santen, R. Sproat, J. Olive, & J. Hirschberg (Eds.), *Progress in speech synthesis* (pp. 139–156). New York: Springer.

Norrick, N. (1993). *Conversational joking: Humor in everyday talk.* Bloomington: Indiana University Press.

Norrick, N. (2001). On the conversational performance of narrative jokes: Toward an account of timing. *HUMOR: International Journal of Humor Research, 14*, 255–274.

Pickering, L. (2004). The structure and function of intonational paragraphs in native and nonnative speaker instructional discourse. *English for Specific Purposes, 23*, 19–43.

Pickering, L., & Attardo, S. (XXXX). *Timing in the performance of jokes.* Q26

Pickering, L., & Levis, J. (2002). *Assessing the intonation patterns of second language learners.* Paper presented at the 36th annual TESOL convention, Salt Lake City, Utah. Q27

Quené, H. (2006). On the just noticeable difference for tempo in speech. *Journal of Phonetics.* Retrieved December 20, 2006, from http://www.let.uu.nl/∼Hugo.Quene/personal/onderzoek/publmss.html Q28

Robson, J., & MackenzieBeck, J. (1999). Hearing smiles—Perceptual, acoustic and production aspects of labial spreading. *Proceedings of the 14th international conference on Phonetic Sciences* (pp. 219–222). San Francisco. Q29

Selting, M. (1994). Emphatic speech style—With special focus on the prosodic signaling of heightened emotive involvement in conversation. *Journal of Pragmatics, 22*, 375–408.

Swerts, M., & Geluykens, R. (1993). On the prosody of information units in spontaneous monologue. *Phonetica, 50*, 189–196.

Swerts, M., & Geluykens, R. (1994). Prosody as a marker of information flow in spoken discourse. *Language and Speech, 37*, 21–43.

Tartter, V. C. (1980). Happy talk: Perceptual and acoustic affects of smiling on speech. *Perception and Psychophysics, 27*, 24–27.

Tench, P. (1996). *The intonation systems of English*. London: Cassell.

Urios-Aparisi, E., & Wagner, M. (2007). *Prosody of humor in Sex and The City.* Paper presented at the 10th international Pragmatics Association conference, Göteborg, Sweden. Q30

Vaid, J., Hull, R., Heredia, R., Gerkens, D., & Martinez, F. (2003). Getting a joke: The time course of meaning activation in verbal humor. *Journal of Pragmatics, 35*, 1431–1449.

Wells, W. H. G. (1986). An experimental approach to the interpretation of focus in spoken English. In C. Johns-Lewis (Ed.), *Intonation in discourse* (pp. 53–75). London: Croom Helm.

Wennerstrom, A. (1997). *Discourse intonation and second language acquisition: Three genre-based studies.* Unpublished doctoral dissertation, University of Washington, Seattle.

Wennerstrom, A. (2001). *The music of everyday speech*. Oxford, England: Oxford University Press.

Wichmann, A. (2000). *Intonation in text and discourse*. London: Longman.

Wooffitt, R. (2001). Raising the dead: Reported speech in medium–sitter interaction. *Discourse Studies, 3*, 351–374.

Xu, Y. & Xu, C.-X. (2005). Phonetic realization of focus in English declarative intonation. *Journal of Phonetics, 33*, 159–197.

Yule, G. (1980). Speakers' topics and major paratones. *Lingua, 52*, 33–47.

Zacher, V., & Niemitz, C. (2003). Why can a smile be heard? A new hypothesis on the evolution of sexual behaviour and voice. *Anthropologie, 41*, 93–100.

Zellner, B. (1994). Pauses and the temporal structure of speech. In E. Keller (Ed.), *Fundamentals of speech synthesis and speech recognition* (pp. 41–62). Chichester, England: Wiley.

APPENDIX

An engineer was crossing a road one day when a frog called out to him and said, "If you kiss me, I'll turn into a beautiful princess."

He bent over, picked up the frog and put it in his pocket. The frog spoke up again and said, "If you kiss me and turn me back into a beautiful princess, I will stay with you for one week."

The engineer took the frog out of his pocket, smiled at it and returned it to the pocket. The frog then cried out, "If you kiss me and turn me back into a princess, I'll stay with you and do ANYTHING you want."

Again the engineer took the frog out, smiled at it and put it back into his pocket. Finally, the frog asked, "What is the matter? I've told you I'm a beautiful princess, that I'll stay with you for a week and do anything you want. Why won't you kiss me?"

The engineer said, "Look I'm an engineer. I don't have time for a girlfriend, but a talking frog, now that's cool."