

## **Production, perception, and communicative goals of American newscaster speech**

EMILY GASSER, BYRON AHN,  
DONNA JO NAPOLI AND Z.L. ZHOU

*Swarthmore College, USA*

*Princeton University, USA*

*Swarthmore College, USA*

*University of California Los Angeles, USA*

### ABSTRACT

Listeners often have the intuition that the speech of broadcast news reporters somehow ‘sounds different’; previous literature supports this observation and has described some distinctive aspects of newscaster register. This article presents two studies further describing the characteristic properties and functions of American English newscaster speech, focusing specifically on prosody. In the first, we investigate the production of newscaster speech. We describe the measurable differences in pitch, speed, intensity, and melodic features between newscaster and conversational speech, and connect those traits to perceptions of authority, credibility, charisma, and related characteristics. In the second, we investigate the perception of newscaster speech. Our experiments demonstrate that listeners can distinguish newscaster from conversational speech given only prosodic information, and that they use a subset of the newscasters’ distinguishing features to do so. (News, prosody, discourse registers, speech perception, credibility, authority)\*

### INTRODUCTION: WHAT MAKES NEWSCASTERS DIFFERENT

Prosodic, lexical, syntactic, and other features—linguistic and paralinguistic—combine to produce certain effects in speech. Studies show that prosodic features can convey authority, credibility, competence, and likability in the speaker (Zuckerman & Miyake 1993; Chattopadhyay, Dahl, Ritchie, & Shahin 2003; Rosenberg & Hirschberg 2005, 2009; Arrabito 2009; Rodero, Larrea, & Vázquez 2010; Elbert & Dijkstra 2014; inter alia). Projecting these qualities is especially critical in news reading. Cotter (1993:90) points out that ‘[p]rosody is KEY to defining the broadcast news register’. ‘Newspeak’ provides a prime opportunity to explore the use and implications of particular prosodic structures, relevant in this age of ‘fake news’ and questions about reporters’ trustworthiness. In this study, we focus on American newscasters’ adoption of a distinctive set of prosodic features that differentiate

on-air speech from that of ordinary conversational settings, the communicative purposes served by these features, and listeners' detection of and reaction to them. We ask four interconnected questions.

- Q1: What prosodic features distinguish on-air speech of newscasters from conversational speech?
- Q2: Can listeners identify newscaster speech based solely on these features?
- Q3: Which features employed by newscasters do listeners use to identify them, and which go unnoticed?
- Q4: Do features employed by newscasters serve identifiable purposes, such as establishing authority, and how does this align with their communicative goals?

We conducted two experiments to answer these questions. In a perfectly efficient communication system, the set of features that newscasters use differently (Experiment 1: Production) and the set by which listeners identify them (Experiment 2: Perception) would be identical. In actual practice, this is not the case. In a production experiment, we compared American English radio news broadcasts with recordings of non-newscasters reading the same material to identify their measurable differences.<sup>1</sup> We identified a set of prosodic features that reliably distinguish newscaster speech, though these do not neatly align with previous findings on prosodic features related to credibility, authority, or other traits desirable to a journalist (e.g. Cotter 2010).

In a perception experiment, participants heard recordings manipulated to preserve prosodic features but remove lexical information, and judged whether they were spoken by newscasters. The rate at which they did so is referred to here as a clip's 'newscaster-ness'. Participants successfully identified newscaster speech, confirming that newscasters employ a distinct speech style and listeners perceive this (Bell 1984, 1991). To do so, they exploited some true identifiers, some false positives, and some false negatives.

These results raise questions about speech features, their communicative functions, and their perceived relationship to personal traits. In a follow-up survey, we asked radio newscasters about their priorities and intentions in crafting their on-air voice. Their responses contradicted our findings, indicating that prosodic features characteristic of newscaster speech are instantiated on a subconscious level, despite public awareness of the effects of intonation (c.f. Brooke & Ng 1986; Helfrich & Wallbott 1986; Johnson & Hunter 2009).

To our knowledge, our perception experiment is the first to ask listeners to distinguish between news and non-news recordings for American English, though Escudero, González, Gutiérrez, & Rodero (2017) conducted a similar experiment for Spanish.<sup>2</sup> We are also the first to experimentally test the contribution of prosodic factors disassociated from lexical scaffolding. We test for a wider range of potentially distinguishing features than previous studies, including speech rate, pitch level and variation, and a number of pitch accents, breaks, and boundary tones.

## THE EXISTENCE OF A NEWSCASTER REGISTER

Prosody differs across discourse types, including conversation, acting, and read speech (e.g. Johns-Lewis 1986). A characteristic newscaster intonation may reinforce speaker credibility and encourage listener confidence, the ‘basic mandate of the profession’ (Raymond 2000:355). Under this hypothesis, newscasters use prosody to present themselves as trustworthy sources disseminating accurate information. Moreover, their speech occurs in a specific discourse context: reading to an absent listener. Together, these contextual factors create the need for a particular and distinctive set of prosodic behaviors. Listeners report recognizing news versus other genres when listening to audio carrying through a wall, even if they cannot identify specific words. This suggests that newscaster speech has prosodic content that listeners distinguish as distinct from other speech types. Hence our two primary hypotheses:

HYPOTHESIS 1 (PRODUCTION): The prosody of American English newscaster speech is measurably different from that of non-newscaster speech.

HYPOTHESIS 2 (PERCEPTION): Listeners can distinguish audio clips of newscaster speech when lexical and segmental information is removed.

Our production experiment tests H1, and our perception experiment tests H2. Both are confirmed.

Despite widespread acknowledgement of a characteristic newscaster intonation, little work has identified the relevant features. Most studies on newscaster speech focus on lexical choices, turn-taking behavior, and the pairing of words and stress (Wheatley 1949; Raymond 2000). Our interest is prosodic features, thus work analyzing lexical content is not applicable here. The existing literature is surveyed below.

*Previous studies of newscaster speech*

Bolinger (1982, 1989) notes that ordinary sentential stress patterns are violated in radio-news speech: emphatic stress appears on words when no semantic emphasis is justified. While Bolinger considers these differences blunders, a designation Cotter (1989) disputes, his observations lay the groundwork for the idea that newscaster speech is partially characterized by prosodic features.

Cotter (1993) aims to define the prosodic properties of the ‘broadcast news register’. She finds that it contrasts with non-news speech in its more frequent pauses at the end of grammatical units, faster pace,<sup>3</sup> and use of intonation to indicate where text would have punctuation. Cotter notes an apparent contradiction: the newscaster voice should convey excitement in order to hold listener attention, but must simultaneously convey impartiality. She suggests that the features newscaster speech

shares with formal discourse modes lend an air of objectivity, while those it shares with conversation encourage connection to the listener. We therefore make no predictions about features correlated with traits other than credibility, authority, and related characteristics, but do measure their frequency and impact on newscaster-ness judgments.

Work on other languages has also found a separate news register. In Spain, broadcast news speech is characterized by distinctive pitch movements utilized at regular intervals (Rodero 2006; Rodero Antón 2013), reproducing Bolinger's English observations. In Brazil, television news speech displays fast pace, lack of pauses, and highly variable intonation (Castro, Serridge, de Moraes, & Freitas 2010).

Escudero and colleagues (2017) use recordings by newscasters and professional voice-over actors to determine the set of pitch accent sequences most characteristic of Iberian Spanish newscaster speech. They impose these melodies on synthesized speech and present subjects with pairs of sentences, one with newscaster intonation and one with a control intonational pattern. Asked which is more likely to appear on a news broadcast, subjects correctly identified news intonation 72% of the time.

Studying Australian radio news, van Leeuwen (1984) again confirms Bolinger's stress findings, and adds that newscaster speech has unnatural prosodic grouping of constituents, a different distribution of pauses, and a faster rate than nonannouncing speech by the same speakers. Price (2008:305) finds that Australian newscaster speech on AM stations uses 'continuative utterance-medial low rises, and utterance-final falling tunes suggesting completeness and finality', while FM radio speech is distinguished by 'its broad pitch range... and the variety of rising tunes it employs'. Price finds similarities between Australian and American broadcast styles, but does not compare these to measurements of conversational speech, leaving it unclear which properties are distinctive.

Other works find cross-linguistic variability and cultural influence, sometimes intentional, within newscaster speech. Iivonen, Niemi, & Paananen (1995) study American English, British English, Finnish, and German news, revealing inter-language differences in pause length, speech rate, and pitch range. They report cross-linguistically consistent use of extra-high F0 peaks early in newscasts and low utterance-final pitch targets. This research suggests that some aspects of newscaster prosody are widespread, while others are language- and culture-specific. However, they did not compare newscaster prosody with other speech genres.

Newsreaders at the Japanese national radio station train to use 'proper and clear' Japanese at a prescribed pace (Krauss 2000). Newscasters on Danmarks Radio are required to use 'clear precise language', avoid reduced syllables, and speak at a slower pace (Schüppert, Hilton, Gooskens, & Heuven 2012). Thus, if there is a communicative basis for a given prosodic indicator, cultural differences may interfere. While some hallmarks, such as unnatural stress placement, unify many of these characterizations, enough features vary across languages that a study of English-language American news in particular is warranted.

PREVIOUS RESEARCH ON COMMUNICATIVE  
USES OF PROSODY

Much research, largely in business and advertising, addresses the relationship of prosodic features to perceptions of speaker credibility, persuasiveness, likeability, charisma, authority, and other traits relevant to the image a newscaster may wish to project. We predict that newscaster speech aims to win listener trust, convey credibility and authority, and keep listener interest without incurring negative judgments of the newscaster, and that prosodic features associated with these traits will be more common in newscaster speech. The subhypotheses presented here further specify precisely how we predict H1 (production) and H2 (perception) to be realized.

*Pitch and sex/gender*

Voices are generally identifiable for gender based on prosody (primarily pitch and formant characteristics; Bachorowski & Owren 1999), and speakers' identified gender can affect persuasiveness (Edworthy, Hellier, & Rivers 2003; Arrabito 2009; Jones, Feinberg, DeBruine, Little, & Vukovic 2010; Grable & Britt 2011; though see Rodero et al. 2010). When gender is a factor in persuasiveness, context-sensitive effects surface. Whipple & McManamon (2002) found that male and female presenters were judged equally persuasive in television commercials about products not marketed to a particular gender, but not so for gender-specific products. Given the targeting of news reports to a general audience, we hypothesize:

HYPOTHESIS 2.1 (PERCEPTION): Speaker gender should not impact how often a clip is judged to come from a news broadcast.

Cross-linguistically, lower-pitched voices have been judged more credible, truthful, pleasant, and attractive, whereas higher-pitched voices have been judged weak, cold, nervous, immature, unpersuasive, and less credible (Cohler 1985; Zuckerman & Miyake 1993).<sup>4</sup> Ohala (1984, 1994) claims that lower pitch indicates authority or aggression, while higher pitch signals submission or appeasement. Rodero and colleagues (2010), looking at Spanish, showed that lower-pitched voices are more effective in advertising. Chattopadhyay and colleagues (2003) reproduced this finding for English, though their sample differed only for high speech rate; their subjects also rated a brand spokesperson more attractive when their voice was digitally manipulated to have a lower mean F<sub>0</sub>.

Rosenberg & Hirschberg (2005) show that vocal pitch is relevant in the judgment of 'charisma', which they define to include persuasiveness.<sup>5</sup> Asking experimental participants to rate American political speech, they found that higher values for mean and maximum F<sub>0</sub> correlated with greater perceived charisma in recordings of male speakers, as did wider pitch range, as measured by standard deviation of F<sub>0</sub>. Tokens realized higher in a speaker's pitch range were rated more charismatic. However, these findings show no connection to credibility or trustworthiness.

Since newscasters want to promote an objective and dispassionate stance, we expect them to aim for credibility and trustworthiness (Bell 1991; Cotter 1993, 1999, 2010), hence we expect lower pitch (pace Escudero et al. 2017). Further, they might eschew pitch characteristics associated with charisma that do not correlate positively to credibility and trustworthiness. We therefore expect newscasters to avoid frequent use of pitches at the higher end of their range, and use a narrower pitch range overall.

HYPOTHESIS 1.1 (PRODUCTION): Newscasters will have a lower pitch (mean, minimum, and maximum F0) than non-newscasters.

HYPOTHESIS 1.2 (PRODUCTION): Newscasters will have a smaller pitch range than non-newscasters.

HYPOTHESIS 1.3 (PRODUCTION): Newscasters will make less use of the higher end of their pitch range than non-newscasters.

### *Pitch variability and intonational markers*

In many types of speech interaction (face-to-face or via telephone) invariable intonation (monotony) is unpersuasive, while highly variable intonation increases persuasiveness (Pittam 1990; Rosenberg & Hirschberg 2005; van der Vaart, Ongena, Hoogendoorn, & Dijkstra 2006; Chebat, El Hedhli, G  linas-Chebat, & Boivin 2007; Biadisy, Rosenberg, Carlson, Hirschberg, & Strangert 2008). Variable intonation and a wide pitch range are related but distinct matters.

The listener's cognitive involvement in a message can influence the effects of prosodic features. For highly involved listeners, both intensity (loudness) and intonational variability influence the degree of credibility they attribute to the speaker, with greater intensity increasing credibility and high pitch variability decreasing it. However, for less-involved listeners, only intensity matters and it increases credibility (G  linas-Chebat, Chebat, & Vaninsky 1996; Elbert & Dijkstra 2014), presumably because loudness captures attention.<sup>6</sup>

Because highly variable pitch correlates with high persuasiveness but low credibility, we expect newscasters to find a middle ground and use moderate-to-low variability in their intonation, skewing towards credibility over persuasiveness. Since no lexical information is available in our perception experiment, listeners cannot be involved in the message; they should therefore fail to make use of pitch variability in their classifications.

HYPOTHESIS 1.4 (PRODUCTION): Newscasters will use moderately variable intonation.

HYPOTHESIS 2.2 (PERCEPTION): Pitch variability will not impact newscaster-ness judgments.

In their studies on political speech, Biadisy and colleagues (2008) and Rosenberg & Hirschberg (2009) look beyond variability to specific prosodic contours.<sup>7</sup> They find that the more downstepped high pitch accents (!H\*) in a given utterance, the more the speaker is judged charismatic. Conversely, greater proportions of

(nondownstepped) H\*, L\*, and L\*+H pitch accents communicates lower charisma. In particular, L\*+H has been linked to incredulity or uncertainty (Pierrehumbert & Hirshberg 1990), explaining its negative correlation with charisma.

As it is unclear whether the newscasters in our sample would aim to project charisma—as Cotter (1993) points out, they need to maintain listener interest while also delivering ostensibly impartial reports—we have no specific expectations as to their use of particular pitch accents. However, we do expect them to avoid intonational patterns that suggest uncertainty, such as L\*+H.

HYPOTHESIS 1.5 (PRODUCTION): Newscasters will make only minimal (if any) use of the L\*+H pitch accent.

Vermillion (2004, 2006) looks at prosodic cues used by New Zealand English speakers to convey emotions and discourse content, including authority, in scripted speech. Speakers used lower L% boundary tones in phrases with H\* L-L% contours when conveying authority; there was no significant difference in the height of the preceding H\*. Vermillion suggests that low L%s indicate social dominance or the completeness of the utterance.

This accords with Gussenhoven's (2004:88) Effort Code hypothesis: wide pitch excursions indicate authority, enthusiasm, or helpfulness, while smaller ones show lack of commitment or interest in the material. By further enlarging the pitch difference between the high and low targets in the H\*L-L% contour, as Vermillion's participants did, speaker authority is reinforced. We therefore expect to find this pattern more often in newscaster speech.

HYPOTHESIS 1.6 (PRODUCTION): Newscasters will use wider pitch excursions and lower L% targets to convey authority.

### *Speed and duration*

Greater speed correlates with greater perceived credibility and persuasiveness (Miller, Maruyama, Beaber, & Valone 1976; Rodero, Mas, & Blanco 2014) and greater perceived charisma in English (Rosenberg & Hirschberg 2005, 2009; Biadys et al. 2008), though too-fast speech impedes listener comprehension (Goldstein 1940; Rodero 2015). Longer utterances, with more words, are judged more charismatic (Rosenberg & Hirschberg 2005). These longer utterances are characterizable by the number and types of prosodic phrases that comprise them. In particular, the more intermediate phrases within an utterance, the higher the association of charisma with the speaker.<sup>8</sup> We therefore have the following expectations.

HYPOTHESIS 1.7 (PRODUCTION): Newscasters will have a faster speech rate than non-newscasters.

HYPOTHESIS 1.8 (PRODUCTION): Newscasters will use more total prosodic phrases than non-newscasters.



Previous work allows us to make predictions about newscasters' use of a range of prosodic features (H1.1–1.8), but we expect newscaster speech to differ from non-newscaster speech in unanticipated ways as well. We therefore study variation in additional prosodic features, including intensity, duration, distribution of pitch within quartiles of a speaker's range, and several pitch accent types, about which we have no specific hypotheses.

#### EXPERIMENT 1: PRODUCTION

To test our production-based hypotheses, we created a small corpus of twelve target sentences originally recorded by radio newscasters during on-air broadcasts, which we re-recorded with two sets of non-newscaster volunteer readers. These recordings were compared to identify distinguishing features of the newscaster register.

##### *Corpus design*

Twelve target sentences (see the appendix) were identified in news broadcasts archived in the Boston University Radio News Corpus (Ostendorf, Price, & Shattuck-Hufnagel 1996). These originally aired in the early 1990s on WBUR, an NPR affiliate in Boston. The target sentences appeared in six different broadcasts; three were read by male newscasters and three by female newscasters. Each recorded sentence lasted between three and eight seconds (eleven to twenty-four words), deemed long enough to convey sufficient prosodic information. Target sentences were declarative statements; none were transitions to other stories or introductions of other reporters. The pieces in which they originated addressed local/regional politics, science, sports, and human-interest stories.

Each target sentence occurred within a larger text, recorded under three different conditions. In the Original script/Newscaster (N) condition, the target sentences were the original aired recordings, as read by the newscaster. In both the Original-script/Volunteer (OV) and Modified-script/Volunteer (MV) conditions, readers were volunteers without journalistic experience or prior knowledge of the experiment's goals. We re-recorded the same target sentences using volunteers (as in Cotter 1993), rather than recording spontaneous utterances, to control for lexical and segmental content, which otherwise may confound comparisons of prosodic variables. All volunteers were instructed to read 'as naturally and conversationally as possible'; none were told the source of the material. In the OV condition, six volunteers read from abridged transcripts of the original broadcasts. Modified scripts containing the same target sentences in more conversational-sounding contexts were created to minimize any effects on intonation of readers' awareness of genre by obscuring the sentences' origins in a news report. These were read by six additional volunteers, as the MV condition. The OV and MV conditions are combined here as 'non-N(ewscaster) recordings' (see Ahn, Gasser, Napoli, & Zhou (2018) for discussion of differences between them).



This yielded thirty-six recordings (twelve sentences x three conditions). All newscasters in our sample and all volunteer readers were post-college adult native speakers of American English, and readers of the same target sentence were matched in gender. Our volunteer readers currently live in the Philadelphia/Swarthmore area. Nine were white, two black, and one of South Asian heritage.<sup>9</sup> The features under investigation were extracted from these clips and compared between the newscaster (N) and non-newscaster (OV and MV) categories.

We measured speech rate (syllables per second); length of recording; variation in intensity; minimum, maximum, range, mean, and standard deviation of F0; total number of pitch accents; number of H\*, L\*, !H\*, L + H\*, L\*+H, and H+!H\* pitch accents; number of intermediate phrases (iPs) and intonation phrases (IPs); and number of (!)H- and L- boundary tones in each recording.<sup>10</sup> We also calculated what proportion of time each speaker spent in each quartile of their pitch range, as well as in the lowest 10% of their range.

Our findings are reported in the RESULTS section alongside results from our perception study, as these are not entirely independent. In brief, newscasters spoke more slowly than non-newscasters, had less variable intensity, spent more time in the middle 50% and highest quartile of their pitch range, had a lower minimum F0, had more L + H\* and no L\*+H pitch accents, and divided their speech into more IPs. Female newscasters also had a lower maximum F0 and larger standard deviation of pitch than their non-news counterparts.

## EXPERIMENT 2: PERCEPTION

To test listener response to newscaster speech, we designed a perception-based task, in which participants rated whether (filtered) audio recordings came from a news broadcast, as well as their confidence in making that judgment.

### *Experimental design*

*Participants.* We recruited participants with US IP addresses through Amazon Mechanical Turk (MTurk). After discarding those who failed the catch questions, 481 usable responses remained.

Participants self-reported demographic information in a survey at the end of the task. Ages ranged from eighteen to seventy-nine, with a mean of 37.7 years. 221 (45.9%) participants identified as male and 259 (53.8%) as female; one reported gender identity as 'Other (don't want to say)'. 97.1% were native English speakers.


*Recordings and stimuli.* The stimuli were the same recordings used in the production study. To exclude possible effects on newscaster ratings of segmental information (vowel quality, lexical information, syntax), these thirty-six recordings were low-pass filtered using Praat (Boersma & Weenink 2017) so that

individual words and phones were not readily identifiable. Participant judgments thus relied only on prosodic cues (e.g. pitch, intensity, phrasing).

*Method.* Participants listened to sound files and indicated whether they thought each was from a ‘newscaster’ or ‘everyday speech’, as well as their confidence in this judgment.

The task was implemented as an online Qualtrics survey. Participants were encouraged to replay a recording as often as they wished before answering, but were not permitted to change earlier decisions. Figure 1 illustrates what participants saw for each question. A response to the current question was required in order to progress to the next.

Does this recording sound like a newscaster's report, or everyday speech?



Newscaster

Everyday speech

How confident are you?

On a scale of 1-5, with 1 being not confident at all, and 5 being very confident.

1 – Not at all

2

3 – Neutral

4

5 – Very confident

FIGURE 1. Sample question from the perception task.

The thirty-six original test stimuli were divided into two sets of eighteen; each set was rated by approximately half of the participants. Each set of eighteen included six target sentences, three read by men and three by women, from each of the three conditions (N, OV, MV). Participants first answered three practice questions, one from each experimental condition, using clips beyond the thirty-six test stimuli. This allowed participants to acclimate to the muffled, odd-sounding filtered clips; no feedback was given. Next, the eighteen test questions and one catch question were presented in a randomized order. A second catch question was always presented last to subjects, to verify continued attention. Thus, each trial consisted of three practice questions for acclimation, eighteen test questions, and two catch questions (totaling twenty-three).

## RESULTS

481 participants successfully completed the perception experiment. Of these, 246 completed the first stimulus set, 211 completed the second, and twenty-four additional participants completed both sets.

The results of both the production and perception studies are presented together here, organized by feature. In an ideal communicative system, the features used by listeners to identify newscasters would be exactly those that newscasters use differently. However, that is not the case here. Newscaster-ness (the rate at which clips were judged as news speech) was predicted by many of newscasters' distinctive prosodic features; however, some distinctive features of newscaster prosody did not contribute, and some prosodic features that did not distinguish newscasters and non-newscasters did contribute. Thus while the results of these studies are not entirely dependent on one another, they are nonetheless interconnected.

### *Ability of participants to identify newscasters*

Participants correctly determined whether a given clip was read by a newscaster at a rate better than chance (1-sample  $t(35) = 2.93$ ;  $p = 0.002$ ). The distribution of participant abilities was approximately normal (Anderson-Darling;  $p < 0.005$ ) with an average correctness of 57.83% and standard deviation of 11.13%. Across all participants, average recall was 67%, precision was 42%, and F1 score was 51%: listeners were better at positively identifying newscasters than at excluding others from the group. This may be related to the fact that only 33% of clips came from newscasters; participants may have posited a more even distribution. It may also be that positive identification is an easier task.<sup>11</sup>

No information on background or demographics (age, gender, frequency of news listening, education level, number of languages spoken, musical experience) correlated with participants' rate of success. Increased participant confidence in their identification abilities had no significant correlation with increased

identification accuracy ( $p = 0.300$ ). Taken together, these findings indicate: (a) prosodic information directly encodes newscaster-ness, and (b) the ability to identify this genre from prosodic information may be independent of other demographics—it may be part of a fluent speaker’s implicit knowledge of American English and not the result of increased practice or specialized interest.

In the following sections, see the tables for statistical details for each measured indicator. To test for effect on newscaster-ness ratings we ran a logistic regression. The resulting model has an adjusted R-squared of 0.94 ( $F(6,30) = 113.1$ ;  $p < 0.001$ ).

### *Speaker and clip features*

Speaker gender did not have a significant effect on participants’ perceived newscaster-ness of audio clips.

TABLE 1. *Speaker and clip features.*

Variable	Different between N and non-N		Effect on newscaster-ness
speaker gender	—	(matched)	$t(34) = 0.51$ ; $p = 0.610$
length	$t(34) = 1.46$	$p = 0.15$	positive, $p < 0.001$
syl/sec	$t(34) = -2.83$	$p = 0.007$	$p > 0.05$
intensity range	$t(32.1) = -3.10$	$p = 0.003$	$p > 0.05$

Though N clips were not significantly longer in duration than non-N clips, clip length was significantly positively correlated with perceived newscaster-ness. Conversely, although N clips had significantly slower speech rates than non-N, speech rate had no significant correlation with perceived newscaster-ness.

Measuring intensity in Praat at 10ms intervals, we found that N clips had significantly less variation than non-N clips. Variability of intensity correlated negatively with perceived newscaster-ness when considered individually, but this effect disappeared when other variables were also considered.<sup>12</sup>

### *Pitch*

Pitch was measured at 10 ms intervals using Praat.<sup>13</sup> We grouped the data by speaker gender for some traits, as preliminary visual analysis showed two distinct groupings of data. Gender-specific conclusions are discussed immediately below, followed by general findings.

*Female speakers.* For female speakers, N clips had significantly lower minimum F0 than non-N clips and higher minimum pitch contributed to lower perceived

newscaster-ness (see Table 2 below). N clips had significantly lower maximum F0 and a (borderline significant) larger standard deviation of pitch than non-N clips. There was no significant difference in mean F0 or the size of pitch range between N and non-N clips. Considered individually, there was a significant negative correlation between mean pitch and perceived newscaster-ness and a positive one between standard deviation of pitch and newscaster-ness, but these effects disappeared when other variables were also considered.

TABLE 2. *F0 of female speakers.*

Variable	Different between N and non-N		Effect on newscaster-ness
minimum	t(16) = -3.22	p = 0.005	negative, p < 0.01
maximum	t(13.7) = -2.50	p = 0.025	p > 0.05
range	t(16) = 0.86	p = 0.40	p > 0.05
mean	t(14.7) = -1.57	p = 0.13	p > 0.05
standard deviation	t(16) = 2.55	p = 0.021	p > 0.05

These findings may be influenced by one of the female volunteer readers, LA, who had an especially high speaking voice. Across the two target sentences she read, LA's mean F0 was 248Hz. By contrast, the remaining five female non-newscaster volunteers had a mean F0 of 185Hz, while the mean F0 for female newscasters was 182Hz, making LA a notable outlier. The increased variation within the female speakers as compared to male speakers, attributable largely to LA, accounts in part for the differing results between genders.

*Male speakers.* For male speakers, N clips had significantly lower minimum pitch than non-N clips and higher minimum pitch contributed to lower perceived newscaster-ness.

TABLE 3. *F0 of male speakers.*

Variable	Different between N and non-N		Effect on newscaster-ness
minimum	t(16) = -3.76	p = 0.001	negative, p < 0.01
maximum	t(16) = -1.05	p = 0.30	p > 0.05
range	t(16) = -0.20	p = 0.83	p > 0.05
mean	t(16) = -0.55	p = 0.58	p > 0.05
standard deviation	t(16) = -0.49	p = 0.62	p > 0.05

There was no significant difference between N and non-N clips in maximum pitch, pitch range, mean pitch, or standard deviation of pitch. Only minimum F0 had a significant (positive) effect on perceived newscaster-ness.

*Pitch quartiles.* In order to compare clips irrespective of speaker gender, we calculated the percentage of time in each clip that the speaker spent in each quartile (Q1 through Q4) of their own pitch range (defined independently for each recording), as well as the lowest 10%. Pitch was measured at 10ms intervals; we determined how many points fell in each quartile, and divided by the number of pitched points.<sup>14</sup>

TABLE 4. *Distribution over pitch quartiles.*

Variable	Different between N and non-N		Effect on newscaster-ness
% time spent in Q1	t(34) = -3.06	p = 0.004	positive, p < 0.001
% time spent in Q2-3	t(34) = 2.45	p = 0.019	positive, p < 0.001
% time spent in Q4	t(34) = 3.13	p = 0.003	p > 0.05
% time spent in lowest 10%	t(33.8) = -3.25	p = 0.002	p > 0.05

N clips spent significantly less time in both the lowest 10% and the full lowest quartile of their pitch range than non-N clips. N clips spent significantly more time in the middle 50% and highest quartile of their pitch range than non-N clips. There was a significant positive relationship between time spent in the lowest quartile and newscaster-ness ratings, and a significant positive relationship between time spent in the middle 50% of the pitch range (Q2 + 3) and newscaster-ness.

*Standard deviation of pitch.* Previous studies (Gélinas-Chebat et al. 1996; Rosenberg & Hirschberg 2005; Castro et al. 2010; Elbert & Dijkstra 2014) have used standard deviation as a measurement of pitch variability, so we followed suit to enable direct comparisons to their work.<sup>15</sup> Standard deviations between the N and non-N groups (with genders pooled) did not differ significantly nor did they affect perceived newscaster-ness.

#### *Pitch accents*

There was no significant difference in the total number of pitch accents between N and non-N clips and no correlation between number of pitch accents and newscaster-ness. This finding contrasts with earlier unpublished work by Jennifer Price, Stefanie Shattuck-Hufnagel, and Mari Ostendorf, where they impressionistically found that texts read by a radio announcer in what that announcer considered a 'news style' had far more pitch accents than the same texts read by the announcer in a 'non-news style' (Stefanie Shattuck-Hufnagel, p.c.).

TABLE 5. *Pitch accents.*

Variable	Different between N and non-N		Effect on newscaster-ness
# of pitch accents	t(34) = 0.78	p = 0.43	p > 0.05
# of H*	t(34) = -0.28	p = 0.77	negative, p < 0.05
# of L*	t(34) = 1.22	p = 0.23	p > 0.05
# of !H*	t(34) = -0.41	p = 0.68	p > 0.05
# of L + H*	t(34) = 2.46	p = 0.018	p > 0.05
# of L*+H	t(23) = -3.39	p = 0.002	p > 0.05
# of H+!H*	t(34) = -0.55	p = 0.58	p > 0.05
% L + H*	t(34) = 1.57	p = 0.12	p > 0.05

We found no significant differences in use of simplex pitch accents between the N and non-N clips; nor were there any effects on perceived newscaster-ness. For complex pitch accents, we found no difference in number of H+!H\* pitch accents. There were significantly more L\*+H pitch accents in non-N clips; in fact, our sample of N clips contained none at all. N clips had significantly more L + H\* pitch accents than non-N clips, echoing Escudero and colleagues’ (2017) finding for Spanish. A positive correlation between number of L + H\* pitch accents and perceived newscaster-ness held only when interactions with other variables were not considered. While there was no significant difference in the use of any other individual pitch accent between groups, lower use of H\* targets correlated with higher newscaster-ness ratings.

*Phrase breaks and boundary tones*

N clips had significantly more intonation phrases than non-N clips, but the positive relationship between number of IPs and perceived newscaster-ness disappeared when other variables were also considered. There was no significant difference between the boundary tones used in N and non-N clips for IPs.

TABLE 6. *Phrase breaks and boundary tones.*

Variable	Different between N and non-N		Effect on newscaster-ness
# of intermediate phrases	t(34) = 0.48	p = 0.62	p > 0.05
# of intonation phrases	t(34) = 3.33	p = 0.002	p > 0.05
# of (!)H-	t(34) = 1.18	p = 0.24	p > 0.05
# of L-	t(34) = -0.64	p = 0.52	p > 0.05

There was no significant difference in number of intermediate phrases between N and non-N clips. Of the associated iP tones, we investigated (!)H- and L-;<sup>16</sup> there



was no significant difference in number of either type. The positive correlation between number of (!)H- boundary tones and perceived newscaster-ness disappeared when other variables were also considered.

### *Effects of corpus date*

To address concerns that the differences described here between N and non-N speech are actually due to changes in speech style over the more than twenty years between the two sets of recordings rather than to newscaster origin, we collected twelve additional sentences from WBUR newscasts broadcast in early 2018, matched with the original corpus for gender and report topic. These differed significantly from our experimental N clips only in variability of intensity ( $p = 0.006$ ), male speakers' pitch maximum ( $p = 0.016$ ), and pitch range ( $p = 0.03$ ). None of these factors had a significant effect on newscaster-ness ratings. The original N clips had less variable intensity than non-N clips. The 2018 clips had more variation in intensity than either set, which may reflect a change in speech style or simply be an artifact of recording conditions. The original N clips did not differ from the non-N clips in male pitch maximum or range. The 2018 clips did not differ significantly in male pitch maximum from the non-N clips, but did have a larger pitch range ( $p = 0.042$ ) than non-N clips.

Caution should be exercised in comparing the 2018 N clips to others. All other clips were matched for lexical and phonological content (which can each influence prosody) so as to allow a direct comparison in prosodic factors. As these clips had different content, it is unclear which prosodic differences (if any) are attributable to changes in newscaster style versus differences in content.

## SURVEY OF NEWSCASTERS

To test our assumptions about newscaster intentions, we surveyed radio newscasters about the traits they try to convey in on-air speech.<sup>17</sup> Respondents were asked, 'When delivering the news on air, how important is it to you to sound...', and were presented with eight descriptors related to characteristics discussed in the prosody literature (*trustworthy, charismatic, objective, persuasive, authoritative, enthusiastic, likable, engaging*) to rate on a scale of one to five.<sup>18</sup> Respondents were then asked to list three characteristics they most tried to convey in on-air speech, and to describe how they used their voice to convey these traits, how their broadcast voice differed from their everyday-conversation voice, and, if those differences were intentional, why. Of the twelve radio broadcasters who answered the trait-rating questions, ten also responded to the short-answer and demographic questions. All ten were male; five worked at public radio stations in the Philadelphia, Boston, and New York areas, and five at commercial stations in Pennsylvania and New England.

The traits rated most highly were trustworthiness (mean score 4.83) and engagement (mean score 4.75); the lowest rated was persuasiveness (mean score 2.83). Mean scores for the remaining five traits ranged from 3.83 to 4.33.

In their short-answer responses, respondents listed likeability/reliability/friendliness, trustworthiness/honesty, intimacy/empathy, and authority/knowledgeability as traits they tried to convey, consistent with findings by Cotter (2010). All respondents emphasized their desire to sound relaxed, conversational, and natural and to avoid affectation. Several claimed no difference between their on-air and conversational voices. A typical response stated, 'I work very hard in not sounding like I'm reading the news.... I imagine telling my story to a friend or family member in a conversational and colloquial manner'.

While newscasters describe trying to sound conversational, our data indicate otherwise, suggesting that newscasters adjust their speech style unconsciously. Newscasters correctly claim they are 'just talking naturally': what constitutes natural talk varies based on context, and news broadcasts have different prosodic requirements from normal conversation. Newscasters unconsciously adjust to the conventions of the discourse genre, replicating a familiar (prosodic) standard of on-air speech. This is substantiated by reports of several of our OV (original script/volunteer) readers finding themselves inadvertently falling into a 'newscaster-y' voice based on the script content.

Those respondents who acknowledged a difference in register described speaking more slowly, enunciating more clearly, avoiding fillers, 'chang[ing] vocal inflection' upon reaching the story's crux, accentuating key words (names, numbers), and aiming for a balance between monotone and 'sing-songy' intonation. This accords with our measurements of speech rate, pitch range, and variability. No survey respondent mentioned the exaggerated low targets and shifts in range utilization that most strongly characterized newscaster speech in our sample (see below).

## DISCUSSION

### *Identifying newscaster prosody*

Our results corroborate previous findings that American English newscaster speech is measurably different from conversational speech on several prosodic dimensions, confirming H1. When the same target sentences were spoken by newscasters on-air and by volunteer readers, the newscasters spoke more slowly. They spent significantly more time in the middle 50% of their pitch range than non-newscaster volunteers, as well as less time in the lowest quartile of their range, and more time in the highest quartile. Newscasters had significantly lower minimum pitch than non-newscaster volunteers, and within the female speakers, newscasters also had a lower maximum pitch. Newscasters used more L + H\* pitch accents overall than non-newscasters, and more IPs. Newscaster recordings showed

significantly less variation in intensity than the non-newscaster recordings, though this may be an artifact of recording conditions rather than production.

Perhaps the most striking difference between the two sets of recordings was the utilization of different regions of speaker pitch ranges, visualized in Figure 2. Non-newscasters spent an average of 47% of their time in the lowest quartile of their pitch range, 34% of their time in the second quartile, 12% in the third, and 7% in the highest. In contrast, on average only 25% of the newscaster speech was in Q1, 39% was in Q2, 22% in Q3, and 14% in Q4. Non-newscaster speech is heavily skewed towards the lower end of the speaker's utilized pitch range, suggesting the use of only short excursions into a higher pitch range. Newscasters' speech is more symmetrically distributed across their full range, with more time spent in the middle, and the addition of short excursions into a lower pitch range. Note that two groups make use of essentially the same amount of pitch—newscasters of both genders represented here had lower pitch minima, but their mean and maximum F0 was no different. It is crucially the proportional utilization of those pitches that distinguishes the groups.

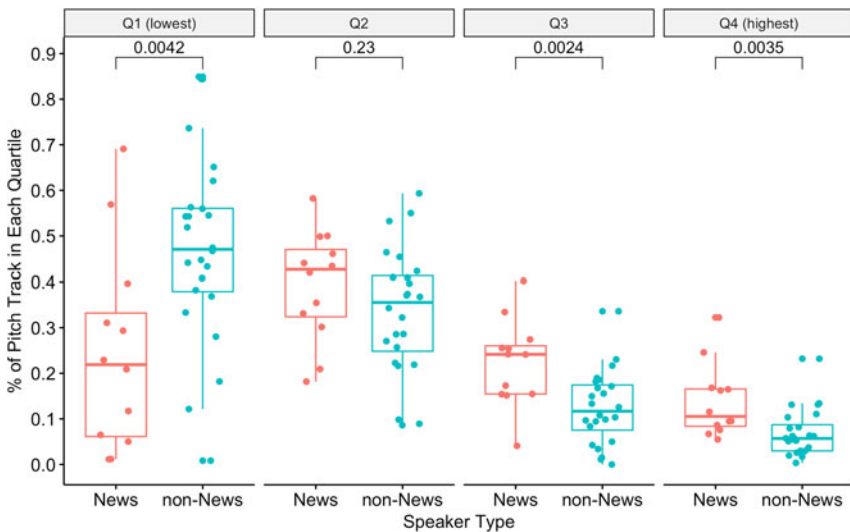


FIGURE 2. Proportion of speaking time spent in each quartile of pitch range by newscaster and non-newscaster readers.

Taken together with the observation that newscasters have a lower minimum F0, this suggests that a key feature of newscaster speech is the presence of a small number of exaggerated low targets, which speakers hit quickly and then move away from. This is supported by the fact that newscasters only briefly occupy the

lowest portion of their range, where these extra-low targets occur. Newscasters spent only 7% of their time in their lowest decile, while non-newscasters produced 16% of their speech there. These low targets shift the quartile boundaries downwards, such that the actual pitch ranges where newscasters' speech is concentrated are the same as that of non-newscasters, but their relative placement within their range is not.

While listeners were able to differentiate between newscaster and non-newscaster sound clips at a rate better than chance, confirming H2, they showed only a 58% accuracy rate. Lexical content and discourse structure are surely the most powerful cues in identifying newscaster speech in real-world scenarios, and while our participants did not have access to that information, it would not have been diagnostic here, as all three experimental conditions covered the same set of target sentences. News readers on NPR stations are generally less extreme in their presentation than those from commercial stations (Cotter 1993:94, 2010). The use of NPR style in this study may have contributed to listeners' low success rate; further work is needed to show whether anchors from commercial stations would be more readily identifiable.

Our two hypotheses regarding perception of newscaster speech are:

HYPOTHESIS 2.1: Speaker gender should not impact how often a clip is judged to come from a news broadcast.

HYPOTHESIS 2.2: Pitch variability will not impact newscaster-ness judgments.

Not all differences between newscaster-and non-newscaster speech noted above affected how the clips were classified, and some differences listeners appeared to use in their classifications were not, in fact, diagnostic of newscaster origin. Clips with lower minimum pitch and more time in the middle 50% of the speaker's pitch range were correctly classified significantly more frequently as spoken by newscasters. Participants failed, however, to use informative traits including slower speech rate, lower maximum pitch (for female speakers), and more time spent in the highest quartile of the speaker's pitch range. Clips with more time in the lowest 25% of the speaker's pitch range were wrongly classified more frequently as spoken by newscasters. Uninformative traits that nonetheless (incorrectly) influenced listener classifications were length of utterance and use of H\* pitch accents. Pitch variability failed to influence listeners' perceived newscaster-ness, confirming H2.2.

These relationships suggest that listeners' mental model of newspeak involves slower speech that exploits lower pitch differently. The pitch-related variables that our participants used all involved lower F0; newscasters' tendency to spend more time in the highest portion of their range was ignored. For speech rate, listeners did not recognize the tendency of the N clips to have fewer syllables per second than the non-N clips. However, the removal of segmental information by low-pass filtering may have made syllable boundaries less distinct, leading listeners to use

clip length as a proxy. Whatever other expectations listeners were drawing on in their classifications, speaker gender did not play a significant role, confirming H2.1.

### *Communicative functions of newscaster prosody*

Our eight hypotheses regarding distinctive characteristics of newscaster speech are repeated here.

HYPOTHESIS 1.1: Newscasters will have a lower pitch (mean, minimum, and maximum F0) than non-newscasters.

HYPOTHESIS 1.2: Newscasters will have a smaller pitch range than non-newscasters.

HYPOTHESIS 1.3: Newscasters will make less use of the higher end of their pitch range than non-newscasters.

HYPOTHESIS 1.4: Newscasters will use moderately variable intonation.

HYPOTHESIS 1.5: Newscasters will make only minimal use (if any) of the L\*+H pitch accent.

HYPOTHESIS 1.6: Newscasters will use wider pitch excursions and lower L% targets to convey authority.

HYPOTHESIS 1.7: Newscasters will have a faster speech rate than non-newscasters.

HYPOTHESIS 1.8: Newscasters will use more total prosodic phrases than non-newscasters.

Few of these predictions from the literature are straightforwardly borne out. H1.1 is partially supported: newscasters had lower pitch minima (and female newscasters had lower pitch maxima). However, previous investigations connecting lower pitch with positive traits (Chattopadhyay et al. 2003; Rodero et al. 2010) look primarily at mean F0, which does not differ significantly in our sample. Furthermore, newscasters spend significantly more time in the highest quartile of their range, contradicting H1.3.

This contrasts with previous findings that lower-pitched voices are considered more credible, pleasant, persuasive, and professional than higher-pitched ones (Cohler 1985; Zuckerman & Miyake 1993; Chattopadhyay et al. 2003; Rodero et al. 2010). Furthermore, the newscasters in our survey reported high interest in sounding trustworthy, pleasant, and professional, though not persuasive.

The brief, exaggerated low targets in newscaster speech, discussed above, confirm H1.6, and may serve to project authority. Gussenhoven (2004) claims that wide pitch excursions such as these indicate authority, while Vermillion (2004, 2006) finds that speakers use significantly lower L% boundary targets in certain contexts when conveying authority. While authority was not the highest-rated goal for our newscaster survey respondents (mean rating 3.92/5), four of them mentioned it.

Support for H1.4 is mixed: overall pitch range did not differ significantly from that of non-newscaster readers (this also contradicts H1.2). Again, this only measures the outer boundaries of the range, not utilization of different parts of its

span. If anything, newscasters made more use of the full extent of their pitch ranges than non-newscasters, as shown in the quartile measurements above.

Previous researchers measured variability using standard deviation of pitch rather than pitch range and quartile distribution (Gélinas-Chebat et al. 1996; Rosenberg & Hirschberg 2005; Castro et al. 2010; Elbert & Dijkstra 2014). Looking at standard deviation, we find no difference between newscaster and non-newscaster intonation overall, though female newscasters averaged a larger standard deviation in their pitch measurements than female non-newscasters. This is consistent with our findings on pitch range, which capture more information about the pitch contour by measuring distribution across pitch quartiles.

Newscasters' pitch was more evenly distributed across quartiles, aligning with other studies that find variable intonation to be more persuasive (Pittam 1990; Rosenberg & Hirschberg 2005, 2009; van der Vaart et al. 2006; Chebat et al. 2007; Biadisy et al. 2008; pace Gélinas-Chebat et al. 1996; Elbert & Dijkstra 2014). Furthermore, newscasters' speech exhibits larger pitch variability and greater use of the upper end of the pitch range, traits positively correlated with charisma (Rosenberg & Hirschberg 2005; Biadisy et al. 2008). This conflicts with the assumption that news presenters seek to convey credibility and impartiality rather than charismatic personality. Indeed, it is also the newscaster's job to hold listeners' attention, encouraging more charismatic speech. Wynn (1974:177) instructs newscasters to 'communicate quiet vitality, warmth, ease and authority', and our survey respondents reported trying to sound both credible and engaging.

While Biadisy and colleagues and Rosenberg & Hirschberg also find that higher mean and maximum F0 correlate with higher charisma in male speakers, we find no significant differences between N and non-N recordings on these dimensions, suggesting that the widely reported negative connotations of higher pitch may prevent this aspect of charismatic speech from being realized. Higher numbers of (!)H\* pitch accents are also found to be more charismatic by Rosenberg & Hirschberg (2009), but the newscasters in our sample do not use this strategy. Overall they appear to be using some charismatic traits but not others.

Rosenberg & Hirschberg note that the L\*+H pattern typically expresses uncertain and incredulous meanings (c.f. Pierrehumbert & Hirschberg 1990), so its failure to appear in our N clips is not surprising. This melody was uncommon but attested in the non-N clips, and surfaced as a significant distinguishing feature, confirming H1.5.

H1.7 is contradicted by our findings: the N clips contained significantly fewer syllables per second. This contrasts with Castro and colleagues' (2010) findings on Brazilian TV news anchors and Miller and colleagues' (1976) finding that greater speed is associated with impressions of increased credibility and persuasiveness, and suggests that the American newscasters in our sample are more interested in clarity than confidence. Concordantly, several survey respondents mentioned making adjustments to their voice to enhance listener comprehension, but none mentioned confidence as a goal.

H1.8 also has mixed support. Newscasters used more intonation phrases on average than non-newscasters; however, there was no significant difference between the two groups in number of intermediate phrases used, suggesting that while each newscaster utterance contains more IPs, each IP contains fewer iPs. As higher numbers of iPs are associated with higher charisma (Rosenberg & Hirschberg 2009), this contributes to our finding that newscasters employ some charismatic traits but not others.

Two identified newscaster hallmarks, a higher number of both L + H\* pitch accents and IPs, have not been correlated in the literature to any specific personal traits. Instead, they are linked to perceptual salience and comprehension. L + H\* accents differ from H\* accents in that the preceding low pitch target makes the following high target more noticeable. Moreover, L + H\* conveys a stronger contrastive meaning than H\* (e.g. Pierrehumbert & Hirschberg 1990), and thus encouraging listener attention. Finally, prosodic phrasing is linked to sentence processing (e.g. Schafer 1997); and increased numbers of phrases may enhance comprehension by chunking up meaning more frequently. Grabbing attention and aiding parsing are important newscaster goals (Cotter 1989, 2010), due to the non-interactive nature of the communicative act.

## CONCLUSION

This study finds that American radio newscaster speech is prosodically distinct, and that listeners notice this, although many of the distinguishing features are sub-phonemic. Taken together with existing research on sociolinguistic and discourse effects of prosody, our findings suggest that newscasters' speech promotes intelligibility, an impression of authority, and listener engagement; however, newscasters do not maximize use of prosody for these purposes, nor does their prosody embody traits associated with higher credibility or persuasiveness. Because of the large number of prosodic variables measured here, each should be independently confirmed in future work to ensure significance.

Our study did not align perfectly with results from newscaster studies in other languages and countries. Since 'newscaster voice' is contextually defined, it is unsurprising that it is not cross-culturally uniform. Because newscasters aim for particular discourse effects, and linguistic and social context influence the discourse effects of prosodic variables, we should expect variation in newscaster prosody. Some variables may overlap across cultures and languages due to iconicity, but additional cross-linguistic work is necessary to separate the culturally specific from the more universal.

We find that listeners identify newscaster speech with statistical reliability: prosody alone allowed subjects to discern newscasters from non-newscasters 58% of the time. The moderate levels of this effect may have several root causes. First, because the content of speech affects its prosody (e.g. questions and assertions differ prosodically), controlling for sentence content may have made the two categories more prosodically similar than is representative of a random



natural sample. That is, the intonational differences between N and non-N speech 'in the wild' may be stronger than reported here, because of differences in content.

Second, we only looked at speech by NPR newscasters, who may speak more conversationally than others. Future work should consider a broader sample of newscasters, with different affiliations and distinct points of view, to better understand the variation within newscaster prosody and its relation to particular sociological identities and communicative goals, especially regarding persuasion.

Our questionnaire for newscasters showed that they consciously aim to make on-air speech readily comprehensible, affirming that newscaster prosody reflects the dynamic of speaker-hearer knowledge, with a bias towards lessening of the hearer's perceptual effort at the cost of higher articulatory effort for the (highly motivated) speaker. Our research suggests that newscasters use prosody to project authority just as a doctor/lawyer would, and also to get/keep attention as in motherese. Future research should explore these parallels more deeply.

#### APPENDIX: TARGET SENTENCES

Below are the twelve target sentences used as the text for recordings in all three conditions. The code in parentheses indicates the script number, as found in the Boston University Radio News Corpus (Ostendorf et al. 1996). Sentences 1–6 were recorded by female readers, and 7–12 by male readers.

1. Price was making his third start for Boston since he was signed as a free agent last month. (f1as30p5)
2. The Red Sox beat the first place Baltimore Orioles five to three this afternoon at Fenway Park. (f1as41p6)
3. Grilsh says he's a product of the hearing world and it's frustrating to no longer be able to participate fully. (f2bs30p1)
4. Grilsh hasn't learned sign language because everyone he knows can hear. (f2bs30p1)
5. You've never seen or heard of the victim but you know the punishment is death in the electric chair. (f3asx4p1)
6. Randall Adams spent twelve years in prison before Texas finally overturned his conviction two years ago. (f3asx4p1)
7. Hack is studying the effect these sounds could have on insects which can hear the noises. (m3bs02p4)
8. No one is sure how the insects figure out which trees are withering. (m3bs02p4)
9. And his administration has not exactly welcomed the parking tax proposal either. (m4bs60p6)
10. But the T apparently knows that parking is a lucrative source of income. (m4bs60p6)
11. The legislature authorized a four hundred twenty-million-dollar reduction in Medicaid's account but left it to Weld to decide which services must go. (m4bs62p1)
12. Weld has also warned that he'd veto any changes to local property tax laws which do not allow for a voter referendum. (m4bs62p1)

## NOTES

\*Many thanks to Lynne Steuerle Schofield, Kate Collins, Colleen Cotter, Dan Steele, Shuang Guan, and our twelve volunteer readers for their feedback and assistance with this project.

<sup>1</sup>Compare Cotter's 1993 study, which is similar but more limited in scope.

<sup>2</sup>They presented thirteen listeners with paired utterances; our 481 subjects judged recordings individually.

<sup>3</sup>This finding was contradicted in our study.

<sup>4</sup>Rodero et al. (2014) find that higher pitch correlates with credibility in political speeches, but their study uses L2 speakers and listeners.

<sup>5</sup>We use the term following their sense.

<sup>6</sup>In Elbert & Dijkstra's study, the less-persuasive highly variable condition also had the highest mean pitch.

<sup>7</sup>All intonational markers (e.g. H\* and L-L%) described in this article are reported using the MAE\_ToBI annotation conventions (Beckman, Hirschberg, & Shattuck-Hufnagel 2005). In brief, H and L indicate (relative) high and (relative) low targets. Pitch accents are aligned with respect to stressed syllables, and are annotated with an asterisk \*; phrase accents are aligned with the end of intermediate phrases (iPs) and are annotated with a minus -; boundary tones are aligned with the end of intonation phrases (IPs) and are annotated with a percent sign %. Finally, downstepping is a process by which high targets are lowered; downstepped highs are annotated by a preceding exclamation point !.

<sup>8</sup>Again following MAE\_ToBI conventions (Beckman et al. 2005), two levels of prosodic phrasing are relevant here: intermediate phrases and intonation phrases. All utterances contain minimally one IP, which itself consists of at least one iP.

<sup>9</sup>Race/ethnicity had no effect on any of the production or perception variables measured.

<sup>10</sup>The number of L% boundary tones is highly correlated with the number of IPs, so it was excluded.

<sup>11</sup>Medrado, Ferreira, & Behlau (2005) find the same result for classifying voiceovers.

<sup>12</sup>We can make only perception claims, not production ones, because our non-N clips were recorded in different facilities from the N recordings, with different audio equipment. Moreover, intensity is sensitive to factors such as distance of the speaker from the microphone, a variable we cannot measure for N clips and did not control for.

<sup>13</sup>The pitch range was determined on a clip-by-clip basis; advanced pitch settings were adjusted from the default with the 'very accurate' box selected, octave cost set to 0.04, octave jump cost to 0.65, and voiced/unvoiced cost to 0.16. The time step strategy was set to 'fixed'.

<sup>14</sup>No pitch information was available for intervals where Praat did not detect any voicing.

<sup>15</sup>Of the thirty-six recordings in our sample, only one had pitch measurements that were normally distributed under an Anderson-Darling test.

<sup>16</sup>We collapsed H- and !H- into a single category, as there were not enough tokens of !H- individually.

<sup>17</sup>None of the reporters whose voices were used in the experiment responded to this survey.

<sup>18</sup>This choice of descriptors is supported by e.g. Cotter (1999, 2010) and Wynn (1974).

## REFERENCES

- Ahn, Byron; Emily Gasser; Donna Jo Napoli; & Z.L. Zhou (2018). *Prosodic features of newscaster intonation*. Princeton, NJ: Princeton University, ms.
- Arrabito, G. Robert (2009). Effects of talker sex and voice style of verbal cockpit warnings on performance. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 51 (1):3–20.
- Bachorowski, Jo-Anne, & Michael Owren (1999). Acoustic correlates of talker sex and individual talker identity are present in a short vowel segment produced in running speech. *The Journal of the Acoustical Society of America* 106(2):1054–63.

- Beckman, Mary; Julia Hirschberg; & Stephanie Shattuck-Hufnagel (2005). The original ToBI system and the evolution of the ToBI framework. In Sun-Ah Jun (ed.), *Prosodic typology: The phonology of intonation and phrasing*, vol. 1, 9–54. Oxford: Oxford University Press.
- Bell, Allan (1984). Language style as audience design. *Language in Society* 13(2):145–204.
- (1991). *The language of news media*. Oxford: Blackwell.
- Biadys, Fadi; Andrew Rosenberg; Rolf Carlson; Julia Hirschberg; & Eva Strangert (2008). A cross-cultural comparison of American, Palestinian, and Swedish perception of charismatic speech. *Speech Prosody* 37:579–82.
- Boersma, Paul, & David Weenink (2017). *Praat: Doing phonetics by computer* [Computer program]. Version 6.0.33. Online: <http://www.praat.org/>.
- Bolinger, Dwight (1982). The network tone of voice. *Journal of Broadcasting* 26:726–28.
- (1989). *Intonation and its uses: Melody in grammar and discourse*. Stanford, CA: Stanford University Press.
- Brooke, Mark E., & Sik Hung Ng (1986). Language and social influence in small conversational groups. *Journal of Language and Social Psychology* 5(3):201–210.
- Castro, Luciana; Ben Serridge; João Antônio de Moraes; & Myrian Freitas (2010). The prosody of the TV news speaking style in Brazilian Portuguese. In Antonis Botinis (ed.), *Proceedings of the Third ISCA Tutorial and Research Workshop on Experimental Linguistics: ExLing 2010*, 17–20. Athens: International Speech Communication Association.
- Chattopadhyay, Amitava; Darren W. Dahl; Robin J. B. Ritchie; & Kimary N. Shahin (2003). Hearing voices: The impact of announcer speech characteristics on consumer response to broadcast advertising. *Journal of Consumer Psychology* 13(3):198–204.
- Chebat, Jean-Charles; Kamel El Hedhli; Claire Gélinas-Chebat; & Robert Boivin (2007). Voice and persuasion in a banking telemarketing context. *Perceptual and Motor Skills* 104(2):419–37.
- Cohler, David Keith (1985). *Broadcast journalism: A guide for the presentation of radio and television news*. Englewood Cliffs, NJ: Prentice Hall.
- Cotter, Colleen (1989). *Some pragmatic considerations of broadcast prosody*. Sussex: University of Sussex MA dissertation.
- (1993). Prosodic aspects of broadcast news register. *Annual Meeting of the Berkeley Linguistics Society* 19(1):90–100.
- (1999). Five facts about the Fourth Estate. In Rebecca S. Wheeler (ed.), *The workings of language: From prescriptions to perspective*, 165–79. Westport, CT: Praeger.
- (2010). *News talk: Investigating the language of journalism*. Cambridge: Cambridge University Press.
- Edworthy, J.; E. Hellier; & J. Rivers (2003). The use of male or female voices in warnings systems: A question of acoustics. *Noise and Health* 6:39–50.
- Elbert, Sarah P., & Arie Dijkstra (2014). An experimental test of the relationship between voice intonation and persuasion in the domain of health. *Psychology & Health* 29(9):1014–31.
- Escudero, David; César González; Yurena Gutiérrez; & Emma Rodero (2017). Identifying characteristic prosodic patterns through the analysis of the information of Sp\_ToBI label sequences. *Computer Speech & Language* 45:39–57.
- Gélinas-Chebat, Claire; Jean-Charles Chebat; & Alexander Vaninsky (1996). Voice and advertising: Effects of intonation and intensity of voice on source credibility, attitudes toward the advertised service and the intent to buy. *Perceptual and Motor Skills* 83(1):243–62.
- Goldstein, Harry (1940). Reading and listening comprehension at various controlled rates. New York: Teachers College Columbia University dissertation.
- Grable, John E., & S. L. Britt (2011). A test of the video narration effect on financial risk-tolerance assessment. *Journal of Financial Planning*. Online: <https://www.onefpa.org/journal/Pages/A%20Test%20of%20the%20Video%20Narration%20Effect%20on%20Financial%20Risk-Tolerance%20Assessment.aspx>.

- Gussenhoven, Carlos (2004). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.
- Helfrich, Hede, & Harald G. Wallbott (1986). Contributions of the German 'expression psychology' to nonverbal behavior research, part IV: The voice. *Journal of Nonverbal Behavior* 10:187–204.
- Iivonen, Antti; Tuija Niemi; & Minna Paananen (1995). Comparison of prosodic characteristics in English, Finnish and German radio and TV newscasts. *ICPhS* 95:382–85.
- Johns-Lewis, Catherine (1986). Prosodic differentiation of discourse modes. In Catherine Johns-Lewis (ed.), *Intonation in discourse*, 199–220. San Diego, CA: College-Hill Press.
- Johnson, Brian K., & Marsha Hunter (2009). *The articulate advocate: New techniques of persuasion for trial lawyers*. Prescott, AZ: Crown King Books.
- Jones, Benedict C.; David R. Feinberg; Lisa M. DeBruine; Anthony C. Little; & Jovana Vukovic (2010). A domain-specific opposite-sex bias in human preferences for manipulated voice pitch. *Animal Behaviour* 79:57–62.
- Krauss, Ellis S. (2000). *Broadcasting politics in Japan: NHK and television news*. Ithaca, NY: Cornell University Press.
- Medrado, Reny; Leslie Piccolotto Ferreira; & Mara Behlau (2005). Voice-over: Perceptual and acoustic analysis of vocal features. *Journal of Voice* 19(3):340–49.
- Miller, Norman; Geoffrey Maruyama; Rex J. Beaber; & Keith Valone (1976). Speed of speech and persuasion. *Journal of Personality and Social Psychology* 34(4):615–24.
- Ohala, John J. (1984). An ethological perspective on common cross-language utilization of F0 in voice. *Phonetica* 41:1–16.
- (1994). The frequency code underlies the sound-symbolic use of voice pitch. In Leanne Hinton, Johanna Nichols, & John Ohala (eds.), *Sound symbolisms*, 325–47. Cambridge: Cambridge University Press.
- Ostendorf, Mari; Patti Price; & Stefanie Shattuck-Hufnagel (1996). *Boston University radio speech corpus, LDC96S36*. Philadelphia: Linguistic Data Consortium.
- Pierrehumbert, Janet, & Julia Hirschberg (1990). The meaning of intonational contours in the interpretation of discourse. In Philip R. Cohen, Jerry L. Morgan, & Martha E. Pollack (eds.), *Intentions in communication*, 271–311. Cambridge, MA: MIT Press.
- Pittam, Jeffery (1990). The relationship between perceived persuasiveness of nasality and source characteristics for Australian and American listeners. *The Journal of Social Psychology* 130:81–87.
- Price, Jennifer (2008). New news old news: A sociophonetic study of spoken Australian English in news broadcast speech. *AAA-Arbeiten aus Anglistik und Amerikanistik* 33(2):285–310.
- Raymond, Geoffrey (2000). The voice of authority: The local accomplishment of authoritative discourse in live news broadcasts. *Discourse Studies* 2(3):354–79.
- Rodero, Emma (2006). Analysis of intonation in news presentation on television. In Antonis Botinis (ed.), *Proceedings of ISCA Tutorial and Research Workshop on Experimental Linguistics: ExLing 2006*, 209–12. Athens: International Speech Communication Association.
- (2015). The principle of distinctive and contrastive coherence of prosody in radio news: An analysis of perception and recognition. *Journal of Nonverbal Behavior* 39:79–92.
- ; Olatz Larrea; & Marina Vázquez (2010). Voces masculinas y femeninas en la locución de cuñas publicitarias: Estudio sobre la efectividad y su adecuación al producto. *Icono* 14(4):281–94.
- ; Lluís Mas; & María Blanco (2014). The influence of prosody on politicians' credibility. *Journal of Applied Linguistics and Professional Practice* 11(1):91–113.
- Rodero Antón, Emma (2013). Peculiar styles when narrating the news: The intonation of radio news bulletins. *Estudios sobre el Mensaje Periodístico* 19(1):519–32.
- Rosenberg, Andrew, & Julia Hirschberg (2005). Acoustic/prosodic and lexical correlates of charismatic speech. *Interspeech* 2005:513–16.
- , & ——— (2009). Charisma perception from text and speech. *Speech Communication* 51(7):640–55.

## PRODUCTION, PERCEPTION, AND COMMUNICATIVE GOALS

- Schafer, Amy Jean (1997). *Prosodic parsing: The role of prosody in sentence comprehension*. Amherst: University of Massachusetts Amherst dissertation.
- Schüppert, Anja; Nanna H. Hilton; Charlotte Gooskens; & Vincent J. Heuven (2012). Syllable deletion in contemporary Danish. *Copenhagen Studies in Language* 42:73–99.
- van der Vaart, Wander; Yfke Ongena; Adriaan Hoogendoorn; & Wil Dijkstra (2006). Do interviewers' voice characteristics influence cooperation rates in telephone surveys? *International Journal of Public Opinion Research* 18:488–99.
- van Leeuwen, Theo (1984). Impartial speech: Observations on the intonation of radio newscasters. *Australian Journal of Cultural Studies* 2(1):84–98.
- Vermillion, Patricia (2004). Using prosodic completion tasks to explore the phonetics and phonology of intonation. *Proceedings of the Tenth Australian International Conference on Speech Science and Technology*, 415–19. Online: <http://assta.org/sst/2004/proceedings/papers/sst2004-174.pdf>; accessed 12 July 2017.
- (2006). Aspects of New Zealand English intonation and its meanings: An experimental investigation of forms and contrasts. Online: <http://researcharchive.vuw.ac.nz/handle/10063/791>; accessed 12 July 2017.
- Wheatley, Katherine E. (1949). Anomalies of radio speech. *American Speech* 24(3):213–15.
- Whipple, Thomas W., & Mary K. McManamon (2002). Implications of using male and female voices in commercials: An exploratory study. *Journal of Advertising* 31:79–91.
- Wynn, Earl R. (1974). Performing. In Robert L. Hilliard (ed.), *Radio broadcasting, an introduction to the sound medium*. New York: Hastings House.
- Zuckerman, Miron, & Kunitate Miyake (1993). The attractive voice: What makes it so? *Journal of Non-verbal Behavior* 17(2):119–35.

(Received 8 October 2017; revision received 16 July 2018;  
accepted 1 November 2018; final revision received 17 November 2018)

### Address for correspondence:

Emily Gasser  
Swarthmore College  
500 College Ave Swarthmore,  
PA 19081, USA  
[egasser1@swarthmore.edu](mailto:egasser1@swarthmore.edu)