



Perceptual effects of a flattened fundamental frequency at the sentence level under different listening conditions

Jacqueline S. Laures^{a,*}, Kate Bunton^b

^a*Program in Communication Disorders, Department of Educational Psychology and Special Education, Georgia State University, 30 Pryor Street, Atlanta, GA 30303, USA*

^b*Institute for Neurogenic Communication Disorders,
The University of Arizona, Tucson, AZ, USA*

Received 13 November 2002; received in revised form 2 April 2003; accepted 15 April 2003

Abstract

The purpose of this series of experiments was to examine the effect of a flattened fundamental frequency (F0) contour on the intelligibility of sentence length material in different listening environments. Eight speakers of different genders and ages produced sentences from the Speech Perception in Noise Test (SPIN). Each utterance was subjected to a resynthesis technique that allowed flattening of the fundamental frequency while maintaining the timing and spectral characteristics of the utterances. To avoid learning effects two groups of listeners were chosen to complete word transcription and interval scaling tasks of the unmodified and flattened F0 utterances under different listening conditions (competing white noise or multi-speaker babble) to obtain measures of speech intelligibility. Results were that a flattened fundamental frequency contour negatively influences speech intelligibility regardless of the nature of the competing background noise.

Learning outcomes: (1) To appreciate the role fundamental frequency variation plays in speech intelligibility. (2) To understand the importance of considering environmental noise in clinical speech intelligibility testing.

© 2003 Elsevier Science Inc. All rights reserved.

Keywords: Speech intelligibility; Prosody

* Corresponding author. Tel.: +1-404-651-0119; fax: +1-404-651-4901.

E-mail address: spejsl@langate.gsu.edu (J.S. Laures).

Prosodic variation is important in the recognition of spoken communication (Cutler, Dahan, & Donselaar, 1997; Liss, Spitzer, Caviness, Adler, & Edwards, 1998, 2000). Yorkston, Buekelman, Strand, and Bell (1999) report that prosodic features, such as changes in fundamental frequency, intensity, or rate, aid in signaling important information or content words, distinguishing between a declarative statement and a question, indicating emotion, and helping the listener understand the syntactic structure of the utterance. There is also evidence that prosody is important in speech intelligibility (Cutler & Norris, 1988; Parkhurst & Levitt, 1978; Price & Levitt, 1983) because it aids the listener in filling in missing or distorted pieces of the speech signal (Wingfield, Lombardi, & Sokol, 1984). Improving prosodic characteristics in individuals with dysarthria results in an increase in speech intelligibility (Ramig, Countryman, O'Brien, Hoehn, & Thompson, 1996; Ramig, Sapir, Fox, & Countryman, 2000; Scott & Caird, 1984, 1983, 1981). The fundamental frequency contour (F0), which is the acoustic correlate of pitch (Kent & Read, 1992), is a primary prosodic feature that may aid intelligibility.

In neurologically healthy individuals, there is considerable variation of F0 during spontaneous speech production, typically ranging from 70 to 150 Hz (Kim, 1994; Silverman, 1987). For speakers with dysarthria, a lack of F0 variability, as measured in sentence level material (Bunton, Kent, Kent, & Duffy, 2001; Kent & Rosenbek, 1982), likely corresponds to the perception of monopitch originally discussed by Darley, Aronson, and Brown (1969a, 1969b). Assuming that improvements of prosodic characteristics (specifically F0 variability) result in improved speech intelligibility, prosody has become a primary focus of treatment efforts for several types of dysarthria (Yorkston et al., 1999). Despite this clinical emphasis, very little data has focused on the isolated impact of prosodic features, specifically F0 variation, on speech intelligibility.

Two studies have focused on the effects of flattening F0 contours on speech intelligibility for sentence level material produced by healthy male speakers. Wingfield et al. (1984) used a channel vocoder technique to flatten the F0 across discourse produced by one male speaker. Results suggested that speech production in the absence of F0 variation was less intelligible than speech with normal F0 variation. Similarly, Laures and Weismer (1999) used a linear predictive coding (LPC) resynthesis technique to flatten the F0 across sentences produced by two male native English speakers with no history of speech disorders. Utterances with an unaltered F0 had higher intelligibility ratings and better transcription scores than those utterances with flattened F0 contours. These findings were reported as preliminary in nature due to the small size of the speaker and listener groups.

Several studies investigating hearing-impaired speech support the notion that F0 variability is important for speech intelligibility. Stevens, Nickerson, and Rollins (1983) found that pitch contours deviating from the normal prosodic range of English speech result in reduced intelligibility. Increasing the participants' F0 ranges to near normal values was positively correlated with increased intelligibility. Maassen and Povel (1984) demonstrated that the intelligibility of the

speech of the hearing-impaired improved with computer-implemented correction of both temporal aspects of deaf speech and F0 contour. Although the speech samples used in these studies contained multiple errors (segmental, timing, and F0), the results are consistent with the notion that F0 variation may be an important factor contributing to speech intelligibility.

Previous studies investigating the effects of modifying the F0 contours of normal speech (Laures & Weismer, 1999; Wingfield et al., 1984) required listeners to judge speech intelligibility in a quiet listening environment or with a background of white noise. A noise background was used in the Laures and Weismer (1999) study to prevent ceiling effects for highly intelligible speakers. Ferguson and Kewley-Port (2002) used background noise for similar reasons in a study of vowel intelligibility in clear and conversational speech; however, they chose a 12-talker babble. There is evidence in the literature that speech identification is greatly influenced by the environment in which it is heard (Wilson, Zizz, Shanks, & Causey, 1990). A majority of studies have focused on use of hearing aids in difficult listening environments rather than changes in speech intelligibility per se. Several studies, however, have reported that a background of multi-speaker babble has a greater negative effect on the identification of speech than when speech is heard in quiet or with a flat spectrum noise competitor (Danhauer & Leppler, 1979; Papsó & Blood, 1989; Sperry, Wiley, & Chial, 1997).

Noises are described in terms of both short- and long-term spectral characteristics. If the spectral characteristics of a noise closely match that of the intended speech target, the task of separating the target from the background noise becomes more difficult (Carhart, Johnson, & Goodman, 1975; Schum, 1996). This is seen as a decrease in rates of speech identification. White noise has a flat, continuous spectrum and provides stimulation of the peripheral auditory system that is different from speech; therefore, any effects of the background noise should be related to audibility of the speech target over the noise (i.e., signal to noise ratio) rather than signal interference (Schum, 1996). Recordings of speech babble from a few or many speakers, on the other hand, have been shown to have similar spectral characteristics to speech produced by a single talker (Carhart, Johnson, & Goodman, 1975; Miller, 1947; Schum, 1996). In the long-term spectrum of babble from multiple talkers there are less amplitude fluctuations, but still there is not perfectly constant amplitude over time. Therefore, these signals interfere with relevant speech cues similar to babble with few talkers. Schum, Matthews, and Lee (1991) showed that the long-term spectrum of competing 12-speaker babble from the Speech Perception in Noise Test (SPIN) (Kalikow, Stevens, & Elliot, 1977) corresponds quite closely to long-term spectrum of speech from any given talker. Recordings of speech babble are often used as background noise during audiologic testing in an attempt to recreate real-life situations.

Based on these studies, it appears that in quiet and white noise environments, speech intelligibility is limited by the audibility of the speech material. In contrast, speech presented with a multi-speaker babble background is affected by the competing acoustic spectrum as well as the audibility of the speech

material. The results of these studies led to the current hypothesis that a lack of F0 variation may have a deleterious effect on speech intelligibility in noise when competing background is speech-like (e.g., multi-speaker babble) as well as when the competing background is of a continuous nature (e.g., white noise). This hypothesis was tested using neurologically normal speakers in order to ensure that articulatory and suprasegmental deficits other than the altered F0 contours would not contaminate speech intelligibility scores.

1. Methods

1.1. *Speech samples*

In order to represent speakers of various ages and genders, speech samples were collected from four older speakers (two male, two female) and four younger speakers (two male, two female). Four different speakers representing the different gender and age categories were used for each listening experiment described below. During the first listening experiment the speaker ages were (22, 25, 71, and 73 years) and during the second listening experiment the speaker ages were (23, 29, 67, and 71 years). The speakers were nonsmokers, had no history of major health problems, and were native English speakers. All speakers had self-reported normal hearing and were judged to be 100% intelligible by the investigators. Each speaker produced a set of 10 low predictability utterances selected from the Speech Perception in Noise Test (Kalikow, Stevens, & Eliot, 1977) which were recorded direct-to-disk using a Realistic microphone. Samples were digitized in real-time using CSpeech at a sampling rate of 22.05 kHz and high pass filtered at 9.8 kHz (Milenkovic, 1994). The sentences selected for the experiments are shown in Appendix A. Two sets of 10 sentences were used in the experiment; female speakers spoke one set and male speakers the second set. Different sets were chosen for each gender to control listener-learning effects. The choice of low predictability utterances was thought to be important to minimize any contextual effects during the transcription; high predictability sentences might artificially inflate transcription scores by allowing listeners to make use of context rather than the acoustic signal itself. The use of low predictability sentences also helps ensure that the sentence sets were comparable across gender. Each speaker produced the sentences first with a habitual prosody and then repeated the same utterance with a somewhat exaggerated prosody. The latter sample was used during the resynthesis and listening portions of the experiment. The exaggerated utterances were utilized to ensure clear perceptual differences between the range of fundamental frequencies covered in the unaltered and flattened utterances ($\Delta = 136$ Hz across the exaggerated utterances and $\Delta = 64$ Hz across the habitual sentences). These values are consistent with typical F0 ranges for simple declarative utterances reported in the literature (70–150 Hz) (Kim, 1994; Silverman, 1987).

1.2. Resynthesis procedures

The resynthesis of the speech samples was completed using an LPC-based algorithm in CSpeech (Milenkovic, 1997). This algorithm allowed for sentence level F0 modifications, but did not affect the temporal or spectral characteristics of the utterances. For each utterance, the voiced portions were identified and hand marked using cursors placed at the first glottal pulse and final glottal pulses based on a spectrogram generated in CSpeech. The LPC resynthesis routine was then followed prior to modifying the F0 contour. The mean F0 for each utterance was computed using a pitch extraction routine in CSpeech (Milenkovic, 1997). The F0 contour for each voiced segment was flattened by manually setting all pitch periods equal to the mean of the utterance and rerunning the LPC routine. An example of a glottal waveform, spectrogram and pitch contour for an original unmodified utterance, a resynthesized unmodified utterance, and a resynthesized utterance with flattened F0 are shown in Fig. 1. In a previous experiment utilizing the same technique, comparisons of the temporal, spectral, and intelligibility ratings of the original and resynthesized utterances showed no differences (Laures & Weismer, 1999).

1.3. Listening experiment 1

Twenty-one listeners (10 women, 11 men; mean age 22.19 years) participated in this portion of the experiment. Listeners had no self-reported history of hearing loss, no clinical or research experience in communication disorders and were native English speakers. Each listener was seated in a sound booth and utterances were presented through a loudspeaker placed 1 m in front of the listener. A constant level of white noise was mixed with the speech signals to generate a noisy listening environment. The noise level was set 5 dB below the frequency peaks of the speech samples. Sentences from four of the eight speakers were used for listening experiment 1. The listening task was divided into four comprised of 22 randomized utterances. Each set included 10 utterances produced by an older speaker and 10 by a younger speaker and two randomly selected reliability items. Within each set, the older and younger speakers were the same gender and utterances within a single set had the same F0 pattern (i.e., unmodified F0 or flattened F0). Each listener heard only two of the four sets, one produced by male speakers and one produced by female speakers. The selection of which two sets each listener heard and the order of presentation were randomized. For each set, listeners were first asked to orthographically transcribe what they heard as accurately as possible. The same set of utterances was then presented a second time and listeners were asked to assign a scaled value of intelligibility using a seven-point equal-interval scale, with 7 labeled as most intelligible (100%) and 1 as least intelligible (0%). Intelligibility for the scaling task was defined as “the ease with which you understand the sentences” (Laures & Weismer, 1999).

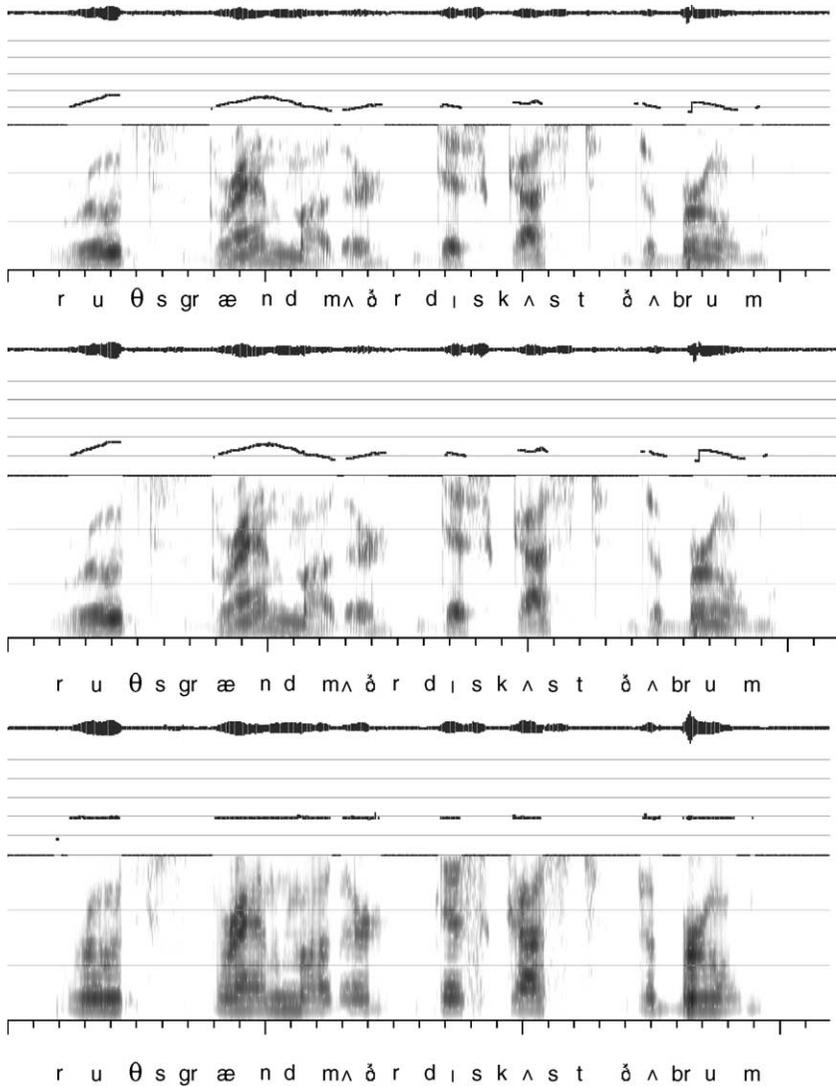


Fig. 1. The top panel shows the glottal waveform, pitch trace and spectrogram for the original recording produced by the older male in experiment 1. The middle panel is the resynthesized, unmodified utterance, and the lower panel shows the utterance with a flattened F0. The sentence shown is “Ruth’s grandmother discussed the broom.”

1.4. Listening experiment 2

The second listening experiment included a novel group of 20 listeners (10 females, 10 males; mean age 23.32 years) for intelligibility scaling and transcription.

A novel group of listeners was chosen to avoid potential listener-learning effects, which may influence transcription scores and the rating of intelligibility. Listeners met the same criteria as in experiment 1. The design of the experiment and tasks were similar to those used in listening experiment 1 with only the background noise changed. Four speakers, different from listening experiment 1 were used in experiment 2. In this experiment, sentences were presented with the 12-voice babble from the SPIN (Kalikow et al., 1977) used as background noise. The level of babble was set to 5 dB below the frequency peaks of the speech sample, similar to the level used in experiment 1. Listeners completed both the transcription and rating tasks as in experiment 1.

1.5. Data analysis

Transcription scores for each speaker were calculated by counting the number of words correctly transcribed for each utterance and then averaging across listeners. A word was scored as correct when the transcription matched the original or intended word exactly. Scaled intelligibility values for each utterance were also averaged across listeners. Mann–Whitney tests were used to compare listener performance on both the transcription and scaling tasks within and across background noise conditions. The Mann–Whitney test was also used to look at listener performance on the transcription versus scaling task within a background noise condition.

1.6. Listener reliability

Two randomly selected utterances from each F0 condition (unmodified and flattened F0) and speaker were repeated within each set of stimuli. This was done to obtain a measure of intra-listener reliability. Measures of inter-listener reliability were derived from median correlations across listeners and conditions.

2. Results

2.1. Acoustic characteristics of the speakers

The mean F0 variability for each speaker is presented in Table 1. Values ranging from 53 to 100 Hz are consistent with data reported in the literature for conversational speech tasks (Kim, 1994; Silverman, 1987). Mean speaking rates were also computed for each speaker as differing rates of speech may have an impact on listeners' perception of the resynthesized utterances with a flattened F0 (Table 2). The male speakers in listening experiment 1 had speaking rates slower than the other speakers, but all speaking rates were within normal limits (Adams, 1990; Miller, Grosjean, & Lomanto, 1984).

Table 1
Mean F0 variability for each speaker across listening experiments

Speaker	Experiment 1	Experiment 2
Older male	68.4	60.1
Younger male	78.2	87.9
Older female	76.8	53.4
Younger female	98.4	100.3

Values are expressed in Hz.

Table 2
Average speech rate and standard deviation are shown

Speaker	Experiment 1	Experiment 2
Older male	125 (6.8)	302 (86.4)
Younger male	136 (11.3)	206 (54.8)
Older female	189 (6.7)	220 (90.6)
Younger female	241 (18.9)	216 (96.8)

Speaking rate is expressed as syllables per second.

2.2. Listening experiment 1: competing background of white noise

2.2.1. Transcription

Table 3 includes the percent of correctly transcribed words for the unmodified and flattened F0 utterances for each speaker. One utterance for the male speakers

Table 3
The percentage of correctly transcribed words and mean scaled intelligibility scores for the unmodified and flattened F0 utterances

	Transcription		Scaling	
	White noise	Multi-speaker babble	White noise	Multi-speaker babble
Older male				
Flat	79.0	86.8	4.67 (1.46)	3.89 (1.29)
Orig	89.2	95.3	5.93 (1.14)	6.38 (1.09)
Young male				
Flat	89.2	77.0	5.47 (1.18)	3.97 (1.42)
Orig	91.1	99.1	6.22 (.85)	6.44 (1.04)
Older female				
Flat	79.9	90.9	4.84 (1.58)	4.89 (1.35)
Orig	88.5	98.2	5.81 (0.96)	6.53 (0.63)
Young female				
Flat	85.7	93.1	5.43 (1.29)	5.39 (1.41)
Orig	88.7	97.7	6.30 (0.72)	6.61 (0.58)

Values in parenthesis represent standard deviations.

was discarded because the original production was incorrect; therefore, males had a total of 54 words possible in the transcription task compared to 62 for female speakers. The average number of words correctly transcribed was similar for all four speakers on the unmodified F0 utterances. Accuracy scores were between 88 and 91% in the unmodified F0 condition and 82 and 88% in the flattened F0 condition. All utterances with a flattened F0 contour showed a decrease in intelligibility compared to the unmodified F0 condition, with the largest difference noted for the older female speaker (54.89/62 unmodified utterance and 49.58/62 for the flattened F0 condition). Differences between the conditions were statistically significant ($W = 151,623$, $P < 0.0126$). To explore contextual influence on the transcription scores the word position most frequently erred was analyzed. Accuracy rates for the initial word were 97 and 92% for male and female speakers, respectively, in the unmodified F0 condition and 93 and 91% in the flattened F0 condition. For the sentence final word, accuracy rates were 89 and 92% in the unmodified F0 condition and 77 and 68% in the flattened F0 condition for male and female speakers, respectively. A statistically significant difference was found in the accuracy of initial versus final word transcription across F0 condition (unmodified $W = 121,771$, $P < 0.005$; flattened F0 $W = 191,352$, $P < 0.003$). This is consistent with expectations for low predictability utterances (Kalikow et al., 1977).

2.2.2. *Scaling*

Table 3 provides the mean values for intelligibility scaling for each speaker and F0 condition. The younger speakers (male and female) had higher scaled intelligibility scores than the older speakers in both the unmodified and flattened F0 condition. Individual speakers had lower scaled intelligibility scores in the flattened F0 condition compared to the unmodified F0 condition. This is similar to findings for the transcription task. Differences between the F0 conditions were statistically significant ($W = 172,296$, $P < 0.000$).

2.3. *Listening experiment 2: competing background of multi-speaker babble*

2.3.1. *Transcription*

Table 3 shows the percentage of correctly transcribed words for the unmodified and flattened F0 conditions for each speaker. For the male speakers the total words possible in the transcription task were 60, for the female speakers the total was 62. As observed in experiment 1, utterances with flattened F0 contours had lower transcription scores than utterances with unmodified F0 contours. Accuracy scores in the unmodified F0 ranged from 95 to 99%, for the flattened F0 condition they were between 77 and 93% accurate. The young male speaker showed the largest decrease in intelligibility (99% for the unmodified and 77% for the flattened F0 condition). Differences between F0 conditions were statistically significant ($W = 179,998$, $P < 0.000$). A word position effect (initial versus final) was found across F0 conditions (unmodified $W = 10,326,162$, $P < 0.0426$;

flattened F0 $W = 15,201,005$ $P < 0.0280$). As in listening experiment 1, the initial word in the utterance was transcribed with greater accuracy than the final word in both unmodified and flattened F0 conditions. Accuracy rates for the initial word were 95% for both male and female speakers in the unmodified F0 condition and 88 and 90% in the flattened F0 condition. The word final accuracy rates were 84 and 89% in the unmodified F0 and 62 and 71% in the flattened F0 condition for male and female speakers, respectively.

2.3.2. *Scaling*

Table 3 shows the mean values for intelligibility scaling for each speaker and F0 condition in experiment 2. As with the transcription scores, scaled intelligibility scores were lower for utterances with a flattened F0 contour compared to utterances with unmodified F0 contours. The mean intelligibility rating across speakers for flattened F0 condition was 4.5 compared to 6.48 for the unmodified F0 utterances. Differences in scaled values for the male speakers between the flattened F0 and unmodified F0 conditions were larger than for the female speakers (2.48 versus 1.43, respectively). Additionally, males had lower scaled intelligibility scores than females in the flattened F0 condition. Younger and older speakers showed the similar decreases in scaled speech intelligibility scores when F0 was flattened compared to their unmodified utterances. Differences between F0 conditions across the group were statistically significant ($W = 219,342$, $P < 0.000$).

2.4. *Listener performance across task*

A comparison of listener performance on the transcription versus scaling task, across type of background noise, showed that listeners were consistent across tasks for a single utterance. In other words, utterances assigned higher scaled intelligibility scores were also the most accurately transcribed. Utterances with lower scaled intelligibility scores were also the least accurately transcribed. Correlations were high for both the white noise listening condition and multi-speaker babble condition (white noise $r = 0.826$, multi-speaker babble $r = 0.805$).

2.5. *Listener reliability*

Two utterances in each F0 condition (unmodified and flattened F0) were remeasured during each listening task to obtain measures of intra-judge reliability. In both experiments there is a moderate to high degree of agreement within listener for transcription (experiment 1, $r = 0.896$, $P < 0.05$; experiment 2, $r = 0.500$, $P < 0.05$) and scaling (experiment 1, $r = 0.937$, $P < 0.05$; experiment 2, $r = 0.671$, $P < 0.05$). Measures of inter-judge reliability derived from median correlations range from 0.456 to 0.841 for transcription across F0 condition and background and 0.105 to 0.417 for scaling across F0 condition and background.

3. Discussion

The present set of experiments was designed to evaluate the isolated impact of a flattened F0 contour on speech intelligibility of neurologically normal speakers in different listening environments. Results of both listening experiments demonstrate that lack of fundamental frequency variation has a significant impact on overall speech intelligibility. Sentence length utterances with a flattened F0 contour had lower transcription and scaled intelligibility scores than utterances with natural F0 contours regardless of the nature of the competing background. These findings are consistent with previous research suggesting that a flattened F0 contour with or without accompanying motor speech problems may be detrimental to speech intelligibility in other listening environments (Bender, Cannito, Buder, Kahane, & Ethington, 2002; Bunton et al., 2001; Laures & Weismer, 1999; Wingfield et al., 1984). These results provide empirical support for a clinical emphasis on prosody to improve speech intelligibility in dysarthric patients.

3.1. Influence of background noise

A combined noisy listening environment and a flattened F0 contour contribute to decreased speech intelligibility. Fundamental frequency variation is an important acoustic cue for the listener regardless of whether the noisy listening environment is nonspeech (white noise) or speech-like (multi-speaker babble). Speech intelligibility scores were equally reduced for the flattened F0 condition compared to unmodified utterances for female speakers in both listening environments (younger female 4% reduction, older female 8.5% reduction). For the male speakers, there was a greater reduction in intelligibility measured for the multi-speaker babble condition compared to the white noise condition (younger male 8 versus 3% and older male 11 versus 2% in multi-speaker babble and white noise, respectively). It is difficult to make a general statement about these differences because different speakers were used in the two listening conditions; however, unmodified F0 intelligibility scores for all four male speakers were similar (range 91–95%) in both background noise conditions.

Differences in performance were seen only when F0 variability was removed from the speech signal. The finding of reduced intelligibility is consistent with Bender et al. (2002) who report that that F0 variation was critical to maintaining speech intelligibility in a noisy environment (white noise) compared to a quiet listening environment. A lack of significant difference in the two background conditions may be related to a linguistic effect. Multi-speaker babble makes it difficult to follow meaningful speech of any one talker, thus the overall effect is a background of speech-like sound without significant linguistic meaning. So even though the long-term spectrum of multi-speaker babble is similar to that of a single talker, there is evidence that the presence of linguistically significant and identifiable cues has a greater effect on speech identification rates (Papsó & Blood, 1989; Souza & Turner, 1994; Trammel & Speaks, 1970). This is a question

needing further study. The findings of the present study are important to clinical evaluation and treatment of prosodic disturbance since a considerable portion of everyday communication takes place in the presence of a speech competitor. Many natural contexts include background noise with both speech and nonspeech sound energy (e.g., traffic, cafeteria) whose inherent frequency and amplitude characteristics have the potential to interfere with a target speech signal. Therefore, treatment of prosodic disturbance, specifically monopitch, should include educating the patient about the importance background noise has on the listener's perception of speech intelligibility.

3.2. *Word position*

Low probability utterances were used in both listening experiments to ensure that listeners would not be able to utilize the context of the utterance to predict the final word (also a content word) in the sentence stimuli, thus inflating the transcription scores. Listeners erred more frequently in transcription of the final word compared to the initial word. This finding supports the notion that listeners were unable to utilize contextual cues and indicates that the decreased transcription accuracy for the flattened utterances was directly related to the lack of F0 variability. This result is consistent with Laures and Weismer (1999) who also found higher error rates for final word transcription accuracy in low predictability sentences. Cutler and Foss (1977) suggest that variability in the F0 contour cues the listener to the content words of an utterance, which typically are stressed. Thus, one would predict that removal of this acoustic cue would result in an error pattern as observed in the current study and in Laures and Weismer (1999).

3.3. *Speaker characteristics*

Although the overall findings of this study indicate that a flattened F0 negatively influences speech intelligibility regardless of the nature of the listening environment, individual speakers appeared to be differentially affected by the flattened F0. Descriptively, in the first listening experiment listeners had higher transcribed and scaled intelligibility scores for the younger speakers compared to the older speakers. Additionally, the younger male speaker had the largest decrease in accurately transcribed words during the second listening experiment. Due to the small sample size these findings do not definitively indicate an age or gender effect. However, it does direct further inquiry into the importance of F0 variation, gender and age. Previous data has suggested that gender differences in production of normal prosodic cues of pitch range, pitch slope, and speech rate exist (Fitzsimons, Sheahan, & Staunton, 2001). Because prosodic patterns differ between genders (Fitzsimons et al., 2001; Henton, 1995), F0 variability may contribute more greatly to speech intelligibility in a specific gender. It is also known that there is greater inter-subject variability of speaking fundamental frequency and greater habitual F0 variability in older adult speakers compared to younger adult speakers

(Benjamin, 1981, 1986; Morris & Brown, 1994). Future studies should include a greater number of speakers representing different age and gender groups to explore the possible interactions among these intrinsic subject variables.

3.4. Conclusion

Naturally occurring F0 contours are an important contributor to speech intelligibility. Findings of the present study indicate that sentences with a flattened F0 contour are less intelligible than are sentences with a varying F0 in noisy listening environments, regardless of the type of noise. These findings have implications for speakers with dysarthria who have limited F0 fluctuations with resultant perceptions of monotonic pitch. Further research should investigate possible relations between speaker gender, age and F0 variability as it relates to speech intelligibility.

Acknowledgments

This work was supported in part by NIH R01 DC00319 and T32 DC00042. The authors would like to thank Gary Weismer for his guidance throughout this study. Additionally, we thank Michael Chial for his helpful comments regarding the background noise.

Appendix A

Sentences produced by male speakers:

1. The old man discussed the dive.
2. They're glad we heard about the track.
3. You're glad they heard about the slave.
4. I should have known about the gum.
5. Mr. Smith knew about the bay.
6. I'm talking about the bench.
7. Tom has not considered the glue.
8. Ruth's grandmother discussed the broom.
9. She has known about the drug.
10. He's thinking about the roar.

Sentences produced by female speakers:

1. We heard you called about the lock.
2. Miss Brown shouldn't discuss the sand.
3. We spoke about the knob.
4. Harry had thought about the logs.
5. She might consider the pool.
6. She wants to speak about the ant.

7. I did not know about the chunks.
8. Sue was interested in the bruise.
9. I hope Paul asked about the mate.
10. Betty had talked about the draft.

Appendix B. Continuing education

1. The perceptual correlate of a flattened fundamental frequency is:
 - a. monoloudness
 - b. monopitch
 - c. bizarreness
 - d. monoproscopy
 - e. none of the above
2. A fundamental frequency contour is one acoustic correlate of:
 - a. intelligibility
 - b. intensity
 - c. prosody
 - d. amplitude
 - e. comprehensibility
3. Based on the findings of this study which of the below are true?
 - a. a flattened F0 contour improves speech intelligibility
 - b. a flattened F0 contour does not affect speech intelligibility
 - c. a flattened F0 contour negatively affects speech intelligibility only when presented with a competing background of white noise
 - d. a flattened F0 contour negatively affects speech intelligibility only when presented with a competing background of multi-speaker babble
 - e. a flattened F0 contour negatively affects speech intelligibility when presented with a competing background of white noise or multi-speaker babble
4. The current results suggest that:
 - a. treatment of dysarthria should emphasize prosody
 - b. speech–language pathologists should educate speakers with dysarthria about the impact of background noise on speech intelligibility
 - c. prosody should not be addressed in treatment
 - d. both a and b above
 - e. none of the above
5. Production of F0 can be affected by:
 - a. age
 - b. gender
 - c. disease process
 - d. hearing status
 - e. all of the above

References

- Adams, S. G. (1990). *Rate and clarity of speech: An X-ray microbeam study*. Unpublished dissertation, University of Wisconsin-Madison.
- Bender, B., Cannito, M., Buder, E., Kahane, J., & Ethington, C. (2002). *Contributions of fundamental frequency to speech intelligibility*. Paper presented at the Annual Convention of the American Speech-Language-Hearing Association, Atlanta, GA.
- Benjamin, B. (1981). Frequency variability in the aged voice. *Journal of Gerontology*, *36*, 722–726.
- Benjamin, B. (1986). Dimensions of the older female voice. *Language and Communication*, *6*, 35–45.
- Bunton, K., Kent, R., Kent, J., & Duffy, J. (2001). The effects of flattening fundamental frequency contours on sentence intelligibility in speakers with dysarthria. *Clinical Linguistics and Phonetics*, *15*, 181–193.
- Carhart, R., Johnson, C., & Goodman, J. (1975). Perceptual masking of spondee by combinations of talkers. *Journal of the Acoustical Society of America*, *58*, 535.
- Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, *40*, 141–201.
- Cutler, A., & Foss, D. (1977). On the role of sentence stress in sentence processing. *Language and Speech*, *20*, 1–10.
- Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology*, *14*, 113–121.
- Danhauer, J., & Leppler, J. (1979). Effects of four noise competitors on the California consonant test. *Journal of Speech and Hearing Disorders*, *44*, 354–362.
- Darley, F., Aronson, A., & Brown, J. (1969a). Differential diagnostic patterns of dysarthria. *Journal of Speech and Hearing Research*, *12*, 246–269.
- Darley, F., Aronson, A., & Brown, J. (1969b). Clusters of deviant speech dimensions in the dysarthrias. *Journal of Speech and Hearing Research*, *12*, 462–496.
- Fitzsimons, M., Sheahan, N., & Staunton, H. (2001). Gender and the integration of acoustic dimensions of prosody: Implications for clinical studies. *Brain and Language*, *78*, 94–108.
- Henton, C. (1995). Pitch dynamism in female and male speech. *Language and Communication*, *15*, 43–61.
- Kalikow, D., Stevens, K., & Elliott, L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *Journal of the Acoustical Society of America*, *61*, 1337–1351.
- Kent, R., & Read, C. (1992). *The acoustic analysis of speech*. San Diego, CA: Singular Publishing Group.
- Kent, R., & Rosenbek, J. (1982). Prosodic disturbance and neurologic lesion. *Brain and Language*, *15*, 259–291.
- Kim, H. H. (1994). *Monotony of speech production in Parkinson's disease: Acoustic entities and their perceptual relations*. Unpublished PhD dissertation, University of Wisconsin-Madison.
- Laures, J., & Weismer, G. (1999). The effects of a flattened fundamental frequency on intelligibility at the sentence level. *Journal of Speech, Language, and Hearing Research*, *42*, 1148–1156.
- Liss, J., Spitzer, S., Caviness, J., Adler, C., & Edwards, B. (1998). Syllabic strength and lexical boundary decisions in the perception of hypokinetic dysarthric speech. *Journal of the Acoustical Society of America*, *104*, 2457–2466.
- Liss, J., Spitzer, S., Caviness, J., Adler, C., & Edwards, B. (2000). Lexical boundary error analysis in hypokinetic and ataxic dysarthria. *Journal of the Acoustical Society of America*, *107*, 3415–3424.
- Maassen, B., & Povel, D. (1984). The effect of correcting fundamental frequency on the intelligibility of deaf speech and its interaction with temporal aspects. *Journal of the Acoustical Society of America*, *76*, 1673–1681.
- Milenkovic, P. (1994). *CSpeech (Computer program)*. Madison, WI: University of Wisconsin.
- Milenkovic, P. (1997). *Dx program (Computer program)*. Madison, WI: University of Wisconsin.

- Miller, G. A. (1947). Sensitivity to changes in the intensity of white noise and its relation to masking and loudness. *Journal of the Acoustical Society of America*, *19*, 609–619.
- Miller, J. L., Grosjean, F., & Lomanto, C. (1984). Articulation rate and its variability in spontaneous speech: A reanalysis and some implications. *Phonetica*, *41*, 215–225.
- Morris, R., & Brown, W. S. (1994). Age-related differences in speech variability among women. *Journal of Communication Disorders*, *27*, 49–64.
- Papso, C., & Blood, I. (1989). Word recognition skills of children and adults in background noise. *Ear and Hearing*, *10*, 235–236.
- Parkhurst, B., & Levitt, H. (1978). The effect of selected prosodic errors on the intelligibility of deaf speech. *Journal of Communication Disorders*, *11*, 249–256.
- Price, P., & Levitt, A. (1983). The relative roles of syntax and prosody in the perception of the /s/-/c/ distinction. *Language and Speech*, *26*, 291–304.
- Ramig, L. O., Countryman, S., O'Brien, C., Hoehn, M., & Thompson, L. (1996). Intensive speech treatment for patients with Parkinson's disease: Short- and long-term comparison of two techniques. *Neurology*, *47*, 1496–1504.
- Ramig, L. O., Sapir, S., Fox, C., & Countryman, S. (2000). Changes in vocal loudness following intensive voice treatment (LSVT) in individuals with Parkinson's disease: A comparison with untreated patients and normal age-matched controls. *Movement Disorder*, *16*, 79–83.
- Schum, D. J. (1996). Speech understanding in background noise. In: M. Valente (Ed.), *Hearing aids: Standards, options, and limitations* (pp. 368–406). New York: Thieme Medical Publishing.
- Schum, D. J., Matthews, L. J., & Lee, F. (1991). Actual and predicted word-recognition performance of elderly hearing impaired listeners. *Journal of Speech and Hearing Research*, *34*, 636–642.
- Scott, S., & Caird, F. I. (1981). Speech therapy for patients with Parkinson's disease. *British Medical Journal*, *283*, 168.
- Scott, S., & Caird, F. I. (1983). Speech therapy for patients with Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, *46*, 140–144.
- Scott, S., & Caird, F. I. (1984). The response of the apparent perceptive speech disorder of Parkinson's disease to speech therapy. *Journal of Neurology, Neurosurgery, and Psychiatry*, *47*, 302–304.
- Silverman, K. E. A. (1987). *The structure and processing of fundamental frequency contours*. Unpublished PhD dissertation, University of Cambridge.
- Souza, P., & Turner, C. (1994). Masking of speech in young and elderly listeners with hearing loss. *Journal of Speech and Hearing Research*, *37*, 655–661.
- Sperry, J., Wiley, T., & Chial, M. (1997). Word recognition performance in various background competitors. *Journal of the American Academy of Audiology*, *8*, 71–80.
- Stevens, K., Nickerson, R., & Rollins, A. (1983). Suprasegmental and postural aspects of speech production and their effect on articulatory skills and intelligibility. In: I. Hochberg, J. Levitt, & M. Osberger (Eds.), *Speech of the hearing-impaired, research, training and personnel preparation* (pp. 33–351). Baltimore, MD: University Park Press.
- Trammel, J., & Speaks, C. (1970). On distracting properties of competing speech. *Journal of Speech and Hearing Research*, *13*, 438–448.
- Wilson, R., Zizz, C., Shanks, J., & Causey, G. (1990). Normative data in quiet, broadband noise, and competing message for the Northwest University auditory test no. by a female speaker. *Journal of Speech and Hearing Disorders*, *55*, 771–778.
- Wingfield, A., Lombardi, L., & Sokol, S. (1984). Prosodic features and the intelligibility of accelerated speech: Syntactic versus periodic segmentation. *Journal of Speech and Hearing Research*, *27*, 128–134.
- Yorkston, K., Buekelman, D., Strand, E., & Bell, K. (1999). *Management of motor speech disorders in children and adults*. Austin: Pro-ed.