Research Article

On Older Listeners' Ability to Perceive Dynamic Pitch

Jing Shen,[a] Richard Wright,[b] and Pamela E. Souza[a]

Purpose: Natural speech comes with variation in pitch, which serves as an important cue for speech recognition. The present study investigated older listeners' dynamic pitch perception with a focus on interindividual variability. In particular, we asked whether some of the older listeners' inability to perceive dynamic pitch stems from the higher susceptibility to the interference from formant changes.
Method: A total of 22 older listeners and 21 younger controls with at least near-typical hearing were tested on dynamic pitch identification and discrimination tasks using synthetic monophthong and diphthong vowels.

Results: The older listeners' ability to detect changes in pitch varied substantially, even when musical and linguistic experiences were controlled. The influence of formant patterns on dynamic pitch perception was evident in both groups of listeners. Overall, strong pitch contours (i.e., more dynamic) were perceived better than weak pitch contours (i.e., more monotonic), particularly with rising pitch patterns.
Conclusions: The findings are in accordance with the literature demonstrating some older individuals' difficulty perceiving dynamic pitch cues in speech. Moreover, they suggest that this problem may be prominent when the dynamic pitch is carried by natural speech and when the pitch contour is not strong.

V oice pitch, which is correlated with fundamental frequency (f0) in speech, is one of the most powerful cues in speech communication. As a linguistic cue, voice pitch conveys important information for speech recognition in quiet, such as the voiced/voiceless distinction for a preceding stop consonant in English (Haggard, Ambler, & Callow, 1970) and lexical information in tone languages (McCawley, 1978). As an acoustic cue, voice pitch facilitates separation of one talker from background talkers and improves speech recognition in the presence of noise (Assmann, 1999; Bird & Darwin, 1998; Brokx & Nooteboom, 1981; Summers & Leek, 1998).

Dynamic pitch makes speech easier to follow when compared with monotone speech because of its role as a major prosodic cue (Ladd, 1996; Lehiste, 1976). Described by linguists typically using the term *intonation*, dynamic pitch has a critical role in providing information on emotion (Fairbanks, 1940; Murray & Arnott, 1993; Uldall, 1960). For instance, early work by Uldall (1960) added one of

16 different pitch contours to four neutral test sentences and found that listeners could use intonation in rating the speech on strength of feeling, friendliness, and authority. Using synthetic speech, Murray and Arnott (1993) showed that pitch range and pitch change patterns (e.g., rising/falling, abrupt/smooth) are associated with major dimensions of emotion (e.g., anger, happiness, sadness).

For speech perception, dynamic pitch can flag prominence on word level and aid spoken word recognition (Cutler, 1976; Friedrich, Kotz, Friederici, & Alter, 2004; Fry, 1958). On the sentence level, it is one of the prosodic cues that can be exploited to differentiate meanings in syntactic ambiguity (Kjelgaard & Speer, 1999; Weber, Grice, & Crocker, 2006). For example, an intonation contour preceding a target word can facilitate speedy processing of the initial phoneme (Cutler, 1976). Kjelgaard and Speer (1999) demonstrated that prosodic cues (duration and pitch) could influence the resolution of temporary syntactic ambiguities very early in the parsing process. For speech recognition under adverse conditions, dynamic pitch is found to be even more helpful. When a speech signal is degraded by low-pass filtering, flattened or inverted pitch contour in speech leads to decrements in intelligibility (Hillenbrand, 2003). A number of studies further provided data from younger listeners with typical hearing, consistently showing the effect of natural pitch contour (compared with

---

[a]Northwestern University, Evanston, IL
[b]University of Washington, Seattle

Correspondence to Jing Shen: jing.shen@northwestern.edu

---

flattened/inverted pitch) on speech recognition in noise (Binns & Culling, 2007; Laures & Bunton, 2003; Watson & Schlauch, 2008).

Young listeners with typical hearing, as a group, are able to perceive dynamic pitch, as shown by control group data from several studies on the perception of intonation (Chatterjee & Peng, 2008; Souza, Arehart, Miller, & Muralimanohar, 2011) or emotion (Dupuis & Pichora-Fuller, 2010; Mitchell, Kingston, & Barbosa-Bouças, 2011; Orbelo, Grim, Talbott, & Ross, 2005). We know older listeners, even with clinically typical hearing sensitivity, have degraded ability to process static frequency (Clinard, Tremblay, & Krishnan, 2010; He, Dubno, & Mills, 1998) and dynamic frequency (Clinard & Cotter, 2015; Grose & Mamo, 2012; Harkrider, Plyler, & Hedrick, 2005; He, Mills, & Dubno, 2007; Sheft, Shafiro, Lorenzi, McMullen, & Farrell, 2012). This compromised pitch perception has been attributed to an age-related loss of neural synchrony (Frisina et al., 2001) and is associated with a declined ability to perceive concurrent speech with different pitch levels (Arehart, Souza, Muralimanohar, & Miller, 2011; Lee & Humes, 2012; Summers & Leek, 1998; Vongpaisal & Pichora-Fuller, 2007). One possible consequence is that older listeners may have more difficulty with competing speech because they benefit less from pitch cues for segregating speech streams.

However, the literature mostly consists of psychoacoustic studies and we know little about older listeners' ability to perceive dynamic speech pitch. Data from Souza et al. (2011) first demonstrated large variability in the dynamic pitch perception of older listeners with near-typical hearing. In that study, older listeners with good hearing and young controls were tested on dynamic pitch perception. They were asked to identify the direction of pitch change (i.e., rising or falling) that was carried by one of four synthesized diphthongs. Individual ability to perceive static pitch was also measured. The older listeners as a group had more difficulty identifying pitch contours when compared with the younger controls, even when audibility was controlled. It is important to note that a large variability in dynamic pitch perception was observed even among listeners with relatively good static pitch perception. This finding suggests that intersubject variability in dynamic pitch perception may be due to multiple factors besides pitch processing ability in general (i.e., that is measured by static pitch perception task). Although these results documented older listener's ability to perceive dynamic pitch, Souza et al. (2011) did not address several issues that may have confounded their results.

Is the Perception of Dynamic Pitch Influenced by Formant Patterns?

First, Souza et al. (2011) used four vowels (/ɑʊ/, /ei/, /ɑi/, and /oi/) in their dynamic pitch identification task. With all of them being diphthongs, this set of stimuli inevitably introduced changes in formant frequencies within each syllable. These dynamic patterns in formant frequencies come with changes in both spectral and temporal envelopes

of the signals, which can obscure the cues that listeners need for perceiving changes in pitch. Using vocoded stimuli, Green, Faulkner, and Rosen (2002) tested dynamic pitch perception in younger listeners across two conditions of sawtooth frequency glides and synthetic diphthongal vowel glides. The performance was poorer for diphthongs than for sawtooths, which suggests interference from complex formant patterns on dynamic pitch perception. Taken together with the findings from Souza et al. (2011), it is possible that some of the older listeners' inability to perceive dynamic pitch stems from a higher susceptibility to the interference from formant changes. This effect may have a negative impact on their ability to benefit from dynamic pitch in speech recognition because speech sound naturally carries dynamic changes in both spectral and temporal domains. Therefore, the present study was designed to include both diphthong and monophthong stimuli to control for and investigate the effect of changes in formant frequencies on dynamic pitch perception.

Are Strong and Rising Pitch Contours Perceived Better Than Weak and Falling Pitch Contours?

The pitch perception literature has suggested that strong pitch contours (i.e., a larger amount of pitch change within a certain duration) are more easily perceived when compared with weak pitch contours (Molis, Srinivasan, & Gallun, 2015; Stalinski, Schellenberg, & Trehub, 2008). Moreover, rising tonal sweeps are better identified than falling ones, particularly for high frequency sounds (Gordon & Poeppel, 2002; Molis et al., 2015). This phenomenon has been attributed to the spatial representations of the frequencies along the basilar membrane (Dau, Wegner, Mellert, & Kollmeier, 2000; Shore & Nuttall, 1985). However, we know very little about whether aging may play a role in influencing these response patterns.

Is the Inability to Identify Dynamic Pitch Due to Labeling Difficulty?

As suggested by earlier research, many speech perception phenomena involve dual processing of both psychophysical perception and a categorizing process (Macmillan, Kaplan, & Creelman, 1977; Pollack & Pisoni, 1971). In an identification paradigm, there is a possibility that poor performance may stem from difficulties categorizing or labeling a stimulus rather than an inability to perceive a signal. This raises the question of whether some of the older listeners in Souza et al. (2011) were unable to identify dynamic pitch due to labeling difficulty rather than an inability to perceive the pitch change. The present study examined this question by including a discrimination task in addition to an identification task with the same stimulus set. If the older participants have substantially more difficulty in the identification than the discrimination paradigm, the difficulty is likely to come from labeling or categorizing problems rather than the perceptual process.

Is the Interindividual Variability in Dynamic Pitch Perception Related to Musical Experience?

Musical training has been shown to enhance pitch processing at subcortical and cortical levels (Besson, Schön, Moreno, Santos, & Magne, 2007; Kraus & Chandrasekaran, 2010; Wong, Skoe, Russo, Dees, & Kraus, 2007). Therefore, individual differences in musical experience could potentially contribute to variability in dynamic pitch processing. In particular, physiological studies have suggested the involvement of the auditory cortex (predominantly the right hemisphere) in the ability to determine the direction of a pitch change in animal models (Syka, Rybalko, Nwabueze-Ogbo, & Suta, 2003; Wetzel, Ohl, Wagner, & Scheich, 1998) and human data (Brechmann & Scheich, 2005; Foxton, Weisz, Bauchet-Lecaignard, Delpuech, & Bertrand, 2009; Pardo & Sams, 1993). Foxton et al. (2009) measured the performance in determining the direction of pure tone pitch glides in a group of young listeners with typical hearing and found considerable differences across individuals. Because the good performers had more musical experience when compared to the poor performers (i.e., 6.4 vs. 1 year of instrument playing), the authors concluded that the difference in pitch glide perception could be attributed to musical training.

Following this rationale, musical experience is likely to contribute to individual differences in dynamic pitch perception. This factor was not controlled for in Souza et al. (2011), and the participants in that study had varying degrees of musical experience. The present study investigated this possibility by asking whether interindividual variability would decrease when compared with the previous dataset when musical experience is controlled. Furthermore, data from the present study would serve as a reference point of older (and younger) listeners' performance on dynamic pitch perception that is not confounded by musical background.

Method

Participants

Two groups of adults participated in the study. The older group consisted of 22 older adults aged 55 to 82 years (mean age = 67.6 years, SD = 7.88 years), and the younger group consisted of 21 younger adults aged 18 to 28 years (mean age = 21.7 years, SD = 3.08 years). There were 16 women and 6 men in the older group, and 13 women and 8 men in the younger group. To control for the factor of musical training and tone language experience, none of the participants had more than 2 years of instrument practice (or vocal training). None of them spoke any tone languages.

All of the participants had at least near-typical hearing, defined as pure tone thresholds of 25 dB HL or better at octave frequencies from 250 Hz to 2000 Hz, and 35 dB or better at 3000 Hz (see Figure 1 for audiograms). Participants were tested monaurally in the ear with better hearing (as defined by a lower pure-tone average threshold). Participants were recruited at Northwestern University for one testing session of 2 hours and were paid for their time. As follow-up work of the earlier paper (Souza et al., 2011),

the present study deliberately controlled gender, age, and hearing of the older group to be comparable with the previous study (see the comparison in Table 1).

Stimuli

Static Pitch Stimuli

The vowel sound of phonetic schwa /ə/ with a 300-ms duration (with 20-ms rise and fall times) were synthesized using Sensimetrics cascade formant software (Klatt, 1980) with a sampling rate of 20 kHz. The first three formant frequencies were modeled on averages for male adults (Peterson & Barney, 1952). The values of the first five formants were as follows: F1 490 Hz, F2 1350 Hz, F3 1690 Hz, F4 3350 Hz, F5 3850 Hz. The f0 values were set to have a baseline f0 of 100 Hz plus an additional f0 increment that ranged from 0 Hz to 130 Hz in 0.1 Hz steps.

Dynamic Pitch Stimuli

Two monophthongs (/ɑ/, /i/) and two diphthongs (/ɑi/, /iɑ/) were included in the stimulus set. These tokens were 620-ms long. They were modeled on a single male talker from the northern cities and synthesized using a Klatt synthesizer in cascade mode. The stimuli had an f0 in the range of 80 Hz to 160 Hz. The ratio of start to end point f0 varied in six equal logarithmic steps from 1:0.5 to 1:2.0. The f0 at the midpoint in time of the stimulus was kept at 113 Hz (see Figure 2 for sample spectrograms and pitch trajectories).

Procedure

All of the signals were sent from a custom MATLAB program (MathWorks, Natick, MA) to a Tucker-Davis Technologies (Alachua, FL) digital signal processor for digital-to-analog conversion. The signals were then routed through a programmable attenuator before being delivered to an ER-2 insert earphone (Etymotic Research, Inc., Elk Grove, IL). The presentation level was set to be 70 dB SPL for all participants.

All of the tasks were implemented using customized MATLAB programs. Each participant was asked to register the responses by clicking buttons on a user interface on a computer screen.

Static Pitch Perception

Consistent with many studies that measure frequency (or pitch) difference limen (e.g., Arehart, 1994; Moore, 1973; Schodder & David, 1960), a two-interval forced-choice (2IFC), three-down one-up procedure (Levitt, 1971) was used to estimate the fundamental frequency difference limen (f0DL). In each trial, the participant heard two stimuli, with one being a standard stimulus and the other a comparison stimulus. The order of the two stimuli was randomized in each trial. The task was to choose the vowel with higher pitch. The f0 of the standard stimulus (f0S) was randomly selected over a 20-Hz range from 100 Hz to 120 Hz, and different f0S values were used across trials. The f0 of the

Figure 1. Audiograms of the two groups (dotted line: individual better ear; solid line: group average).
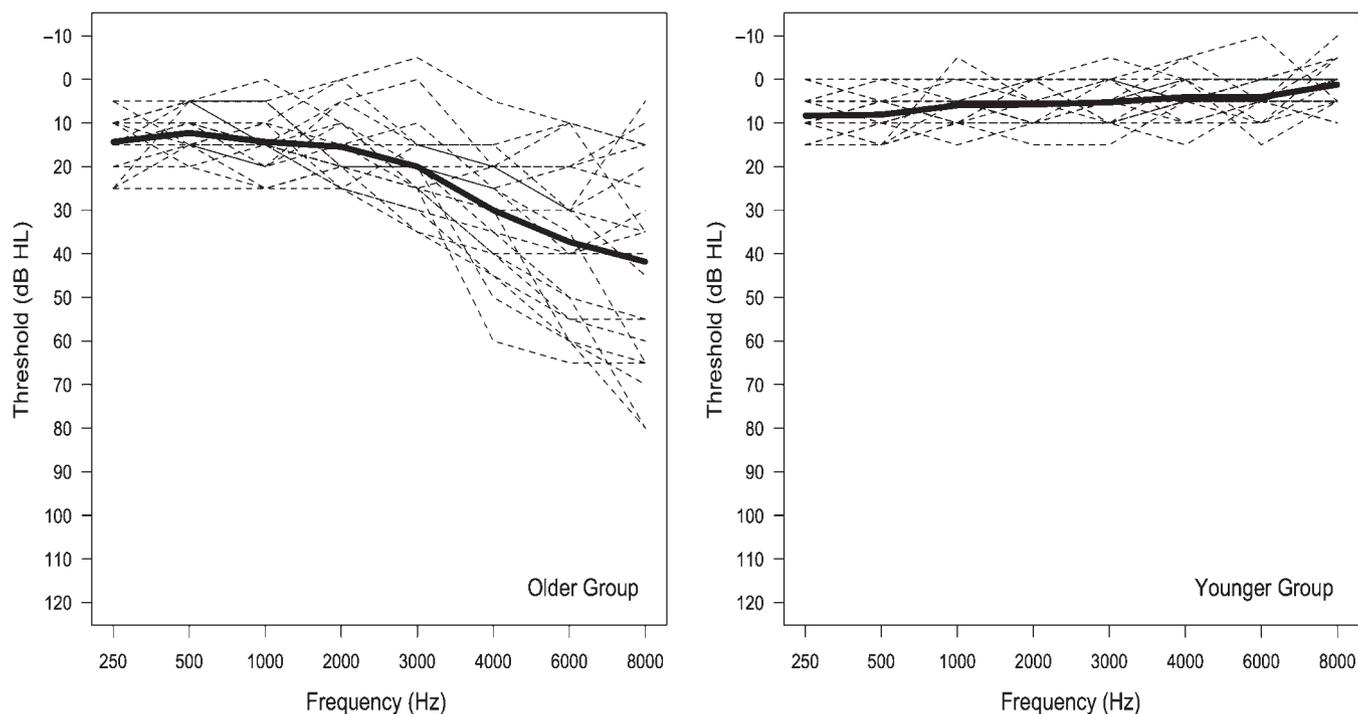
comparison was f0S + Δf0. Δf0 had an initial value of 25 Hz. The Δf0 value was increased or decreased by a factor of 2 for the first two reversals and by a factor of 1.26 for the last 10 reversals during the adaptive procedure. To limit the use of intensity cues, the level of the vowels was also roved across ±1.5 dB in each trial.

Prior to testing, each participant had one practice block with feedback to be familiarized with the stimuli and paradigm. The practice block was nonadaptive with 20 trials and a fixed Δf0 of 25 Hz between the standard and the comparison stimuli. Afterward, each participant had three adaptive testing blocks without feedback. The final f0DL measure was based on the geometric mean of the f0DL values across three blocks for each participant.

Dynamic Pitch Identification

A two-alternative forced-choice identification paradigm was used to measure a participant's ability to identify dynamic pitch with rising or falling glides. In each trial, a stimulus was presented and the participant was asked to select the button that corresponded to the direction of pitch change ("rise" or "fall"). A testing block consisted of 168 trials (7 repetitions × 6 pitch contours × 4 vowels) and a practice block consisted of 24 trials (one presentation of each token). The order of the stimuli was randomized across participants. Each participant had a practice block with feedback prior to two testing blocks. No feedback was given in the testing blocks.

Dynamic Pitch Discrimination

An AAX same–different discrimination paradigm was used to measure a participant's ability to discriminate dynamic pitch with rising or falling glides. In each trial, the participant was presented with three stimuli. The first two tokens were always the same stimulus. The third token was the same vowel as the first two but had either the same

Table 1. Comparison of the participants' characteristics in the two studies.

Study	Gender	Age	Hearing	Musical background
The present study	16 women and 6 men	55 to 82 years (mean age = 67.6 years)	Pure tone thresholds of 25 dB HL or better at octave frequencies from 250 Hz to 2000 Hz and 35 dB or better at 3000 Hz	Participants had less than 2 years of musical experience.
Souza et al. (2011)	16 women and 5 men	66 to 82 years (mean age = 70 years)	Pure tone thresholds of 25 dB HL or better at octave frequencies from 250 Hz to 2000 Hz and 30 dB HL or better at 3000 Hz	Participants had a range of musical experience, including some with more than 2 years of musical training.

Figure 2. Spectrograms of the four vowels (left panel) and pitch contour plots (numbers in the plots are pitch values at onset and offset; right panel).
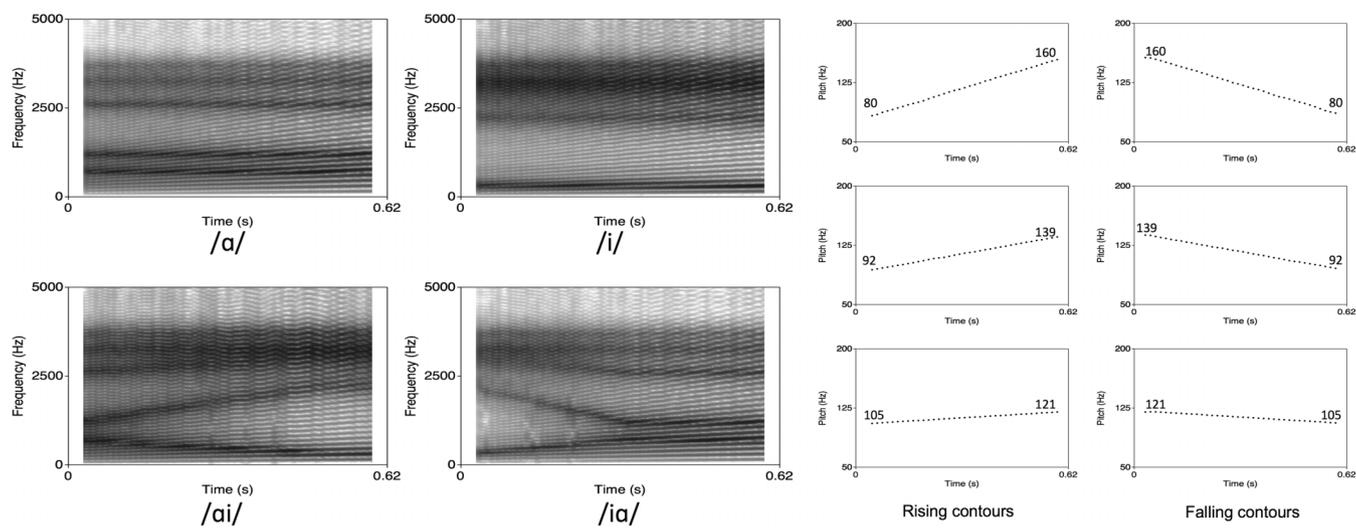
or different pitch change direction when compared with the first two tokens. The slope of pitch change was kept consistent between the first two and the last tokens. The participant was asked to indicate whether the third token was the same or different when compared with the first two tokens. A testing block consisted of 96 trials (two repetitions of each combination) and a practice block consisted of six trials (to familiarize the participant with the paradigm). The order of the stimuli was randomized across participants. Each participant had a practice block with feedback prior to one testing block without feedback.

Results

The dynamic pitch perception data are presented in Figure 3 as individual psychometric functions with the proportion of responses "falling" as a function of the dynamic pitch conditions, as indicated by the ratio between the start and end f0s. To quantify dynamic pitch perception, the f0 ratio was log-transformed with base 10 and logistic regression was fit on each participant's data. The slope of each individual's psychometric function was used for most of the analyses, where a higher value of slope indicates better performance and a lower value indicates worse performance.

First, the dynamic pitch perception performance was plotted as a function of the static pitch perception performance (see Figure 4; note that a linear scale of f0DL is used to compare the present data to the previous studies). Consistent with Souza et al. (2011) and Chatterjee and Peng (2008), the relationship between dynamic and static pitch perception was best fitted with an exponential function, which suggests good pitch processing ability is necessary, but not sufficient for dynamic pitch perception, older group: $b = -0.52$, $t(20) = -3.25$, $p < .001$, Cohen's $f^2 = 2.7$; younger group: $b = -0.08$, $t(19) = -3.17$, $p < .001$, Cohen's $f^2 = .54$.

To minimize the consequences of variability in older listener's hearing loss (particularly in the high-frequency range), we tested the relationship between an individual's dynamic pitch perception performance and the hearing threshold. Dynamic pitch perception was not correlated with either pure tone average (.5, 1, and 2 kHz), $r = -.14$, $t(20) = -0.64$, $p > .1$, or high-frequency pure tone average (4, 6, and 8 kHz), $r = -.19$, $t(20) = -0.87$, $p > .1$.

Effect of Formant Patterns

A repeated measures analysis of variance was used to examine the effects of formant patterns (i.e., monophthong/diphthong) and age group (older/younger). Both groups of listeners performed better with monophthongs than with diphthongs, $F(1,41) = 12.21$, $p < .01$, $\eta^2 = .23$ (see Figure 5). The effect of age group was not significant, $F(1,41) = 1.12$, $p > .1$, $\eta^2 = .03$, and the interaction between formant pattern and age was not significant, $F(1,41) = 2.19$, $p > .1$, $\eta^2 = .05$.

Effect of Pitch Change Strength and Direction

To examine the effect of pitch change strength and direction, accuracy data were transformed to rationalized arcsine unit scores (Studebaker, 1985) and analyzed by repeated measures analysis of variance. The main effect of pitch change strength was significant, $F(2,82) = 61.34$, $p < .01$ (see Figure 6). Pairwise comparison after Bonferroni correction showed that the weakest pitch glide was significantly more difficult than the other two levels ($p < .01$). The performance did not differ across two pitch change directions (i.e., rising and falling), $F(1,41) = 2.2$, $p > .1$, $\eta^2 = .05$. The interaction between strength and direction was significant, $F(2,82) = 28.85$, $p < .01$, $\eta^2 = .41$, which indicates a stronger influence of pitch change strength on the performance when the tokens have rising pitch (see
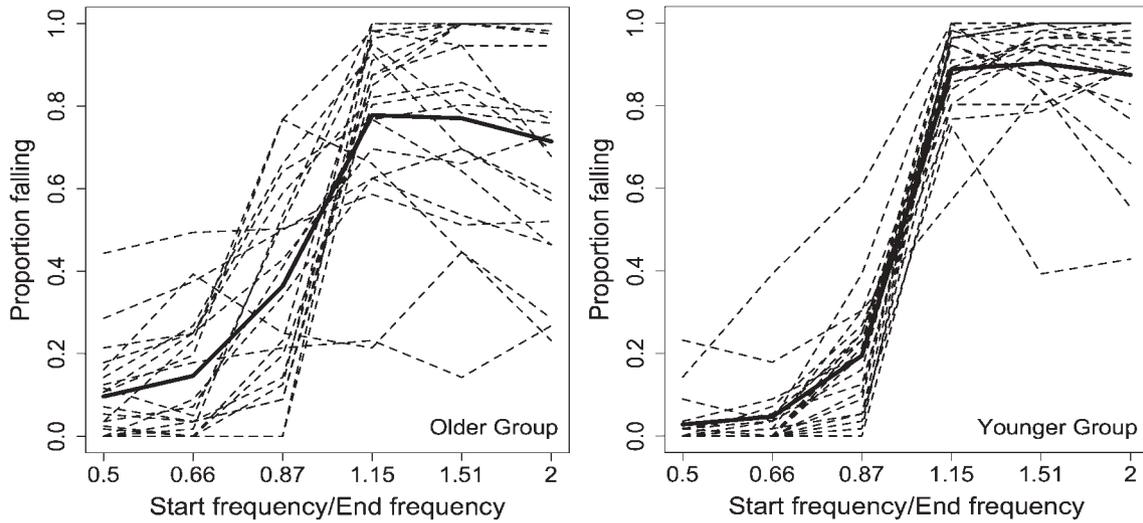
Figure 6). The older listeners had the same response pattern as the younger listeners, suggested by the nonsignificant interactions between age group and pitch change strength, $F(1,41) = .15$, $p > .1$, $\eta^2 = .004$, or direction, $F(1,41) = 0.03$, $p > .1$, $\eta^2 = .001$.

To test the question of whether there was a bias toward rising pitch contour in the participants' responses (Gordon & Poeppel, 2002; Molis et al., 2015), bias scores were calculated as the ratio between the accuracy on rising tokens and the accuracy on falling tokens, that is, p("rise"/rising contour)/p("fall"/falling contour). Overall, the older group had a slightly higher (but nonsignificant) "rising" bias than did the younger group (see Figure 7).

Figure 4. Dynamic pitch perception performance as a function of static pitch perception performance.
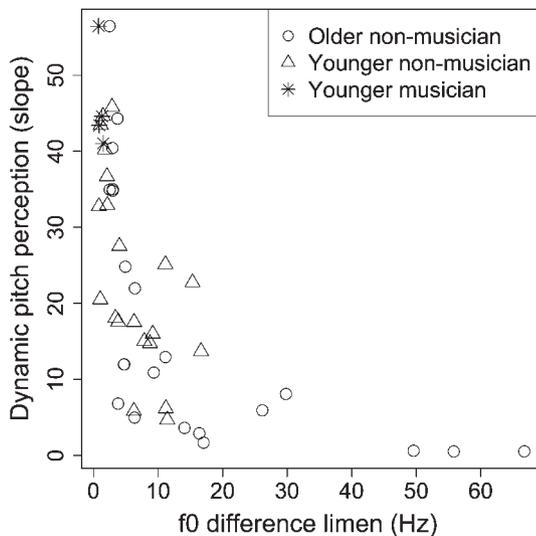
Consistent with the accuracy data, the effect of pitch change strength was significant, $F(2,78) = 4.15$, $p < .05$, $\eta^2 = .1$, which indicates that the listeners tend to respond with "rise" more frequently with stronger pitch contour.

Relationship Between Identification and Discrimination Data

Figure 8 demonstrates the relationship between performance on the identification and discrimination tasks. These data were strongly correlated for the older group, $r = .84$, $t(20) = 6.91$, $p < .01$, but not for the younger group, $r = .33$, $t(19) = 1.51$, $p > .1$, which is likely due to a ceiling effect for most of the younger listeners.

It is worth noting that among all of the listeners, those who performed well in the identification paradigm also performed well in the discrimination paradigm (they

Figure 5. Group means of dynamic pitch perception performance. Error bars indicate ± 1 standard error.
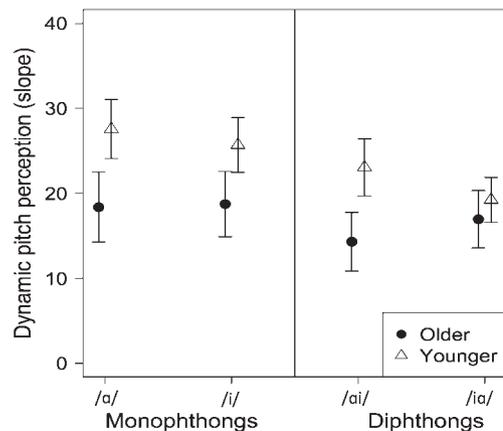
Figure 6. Dynamic pitch perception performance by tokens. Error bars indicate ± 1 standard error.
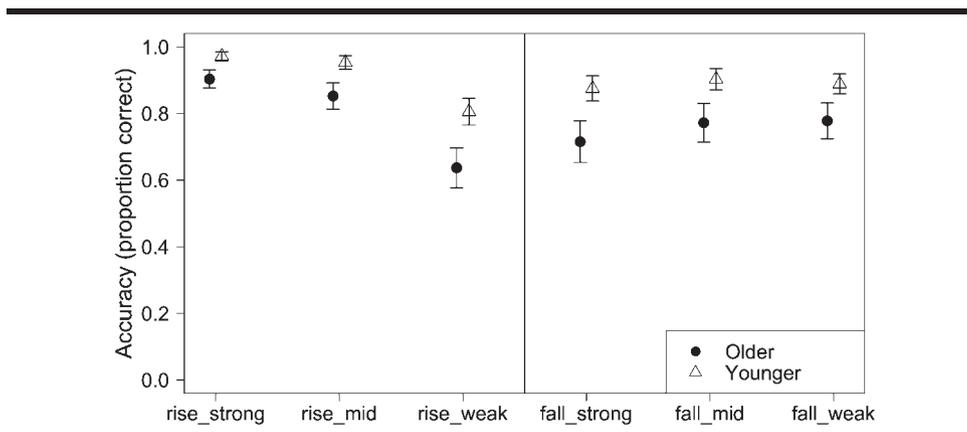
are the data points in the upper-right quadrant in Figure 6), whereas those who performed poorly in the identification paradigm performed poorly in the discrimination paradigm (they are the data points in the lower-left quadrant in Figure 6). In addition, there were a few individuals (four older and two younger) who had relatively high discrimination accuracy but low identification accuracy (they are the data points in the lower-right quadrant in Figure 6). This pattern suggests that the ability to discriminate dynamic pitch is a prerequisite for good performance on dynamic pitch identification.

Interindividual Variability in Dynamic Pitch Perception After Controlling for Musical Experience

One of the goals of the present paper was to test whether controlling for musical experience would reduce

the interindividual variability that was observed in Souza et al. (2011). Although the present study and Souza et al. (2011) used the same paradigm to test dynamic pitch perception, there were different numbers of stimuli in the pitch glide continuum across the two studies (12 in the previous study and six in the present study). Therefore, the slopes of the psychometric functions that indicate dynamic pitch perception ability in the two studies are on different numeric scales.

To quantify the variability in the two datasets, the absolute value of deviation from the mean was calculated. The deviation values were normalized based on the mean of each dataset. Next, the two sets of normalized deviation scores were subjected to a Mann–Whitney U test to test whether they were different from each other. The results showed that the variability was comparable across the datasets (older groups, $U = 212$, $p > .1$; younger groups, $U = 107$, $p > .1$), which suggests that the interindividual variability cannot be attributed to individual differences in musical training alone.

Figure 7. Response bias (values >1 indicate bias toward "rising") in dynamic pitch perception. Error bars indicate ± 1 standard error.
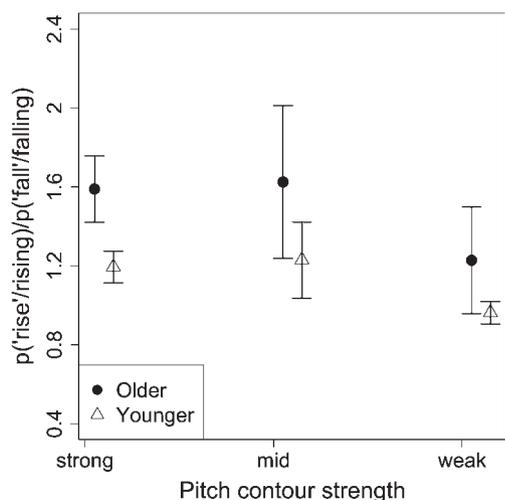
Figure 8. Dynamic pitch identification performance as a function of discrimination performance.
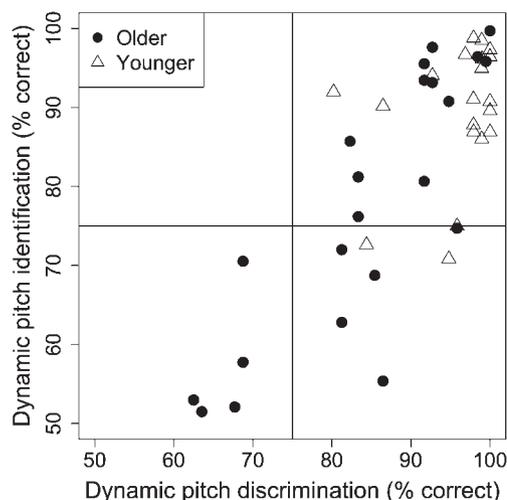
Discussion

The results of the present study are consistent with Souza et al. (2011) in showing large interindividual variability in older listeners' dynamic pitch perception. Although the older listeners as a group performed worse than the younger listeners (see Figure 5), the difference was not statistically significant, which was likely due to the large interindividual variability. The present study further examined four questions that were not addressed by Souza et al. (2011).

Is the Perception of Dynamic Pitch Influenced by Formant Patterns?

The results showed that it was more difficult for the listeners to perceive dynamic pitch with diphthongs (i.e., having dynamic formant patterns) than monophthongs. This finding is consistent with those of Green et al. (2002) in suggesting the influence of complex spectral and temporal variation on dynamic pitch perception in speech sounds. This pattern was observed in both older and younger listeners with typical or near-typical hearing, which suggests the involvement of a perceptual interaction between fundamental frequency and formant frequencies in vowels (Kuhl, Williams, & Meltzoff, 1991; Remez, Fellowes, Blumenthal, & Nagel, 2003).

Although both the older and the younger groups showed a similar pattern in perceiving dynamic pitch carried by monophthongs versus by diphthongs, it is worth noting that the consequences may be more detrimental to older listeners due to their poorer overall performance when compared with that of the younger listeners. Because natural speech has constant variations in spectral structure and in temporal envelope, these complex patterns are likely to interfere with the perception and utilization of the dynamic pitch cue in older listeners. Missing the benefit from this cue could potentially increase the older listeners' speech recognition difficulty under adverse conditions.

Are Strong and Rising Pitch Contours Perceived Better Than Weak and Falling Pitch Contours?

The present study yielded two main findings for this question. First, strong pitch contours were better perceived when compared with weak pitch contours. This result aligns with the speech recognition literature that shows stronger pitch contours are associated with higher intelligibility in emotional speech (Dupuis, 2011) and in multitalker and accented speech (McCloy, Wright, & Souza, 2014). It is presumed that stronger dynamic pitch is perceptually more salient and serves better as a prosodic cue for enhancing speech intelligibility.

Although our older group overall performed slightly worse than the younger group on average, the large within-group variability indicates that some of the older individuals have much more difficulty than others with dynamic pitch. Taken together with the finding that older listeners as a group rely heavily on prosodic cues for recognizing speech in quiet (Wingfield, Lahar, & Stine, 1989; Wingfield, Lindfield, &

Goodglass, 2000), it raises new questions that are currently under investigation in our laboratory. First, does an inability to perceive dynamic pitch contribute to some older listeners' speech recognition difficulty under adverse conditions? Second, can stronger dynamic pitch enhance speech intelligibility in background noise, particularly for these older listeners?

The other main finding is that the strength of pitch change affects rising pitch contours more than falling pitch contours. In other words, listeners can identify strong rising pitch contours better than weak rising contours, but this pattern is less clear with falling pitch contours. Although the literature has suggested that rising pitch contour is more easily perceived than falling pitch contour (Gordon & Poeppel, 2002; Molis et al., 2015), a significant interaction between pitch change direction and strength was not shown by these studies. Considering the fact that those studies used nonspeech stimuli and our study used speech stimuli, this finding may be due to cognitive or social mechanisms instead of perceptual mechanisms. Future research is needed to further examine this phenomenon.

It is interesting that the findings from the recognition of emotional speech (e.g., Dupuis, 2011) also suggest strong rising pitch contours, such as those carried by the emotion of pleasant surprise, have stronger saliency for capturing a listener's attention. As a consequence, this type of emotional sound has a boosting effect on speech intelligibility in background noise (as measured by word recognition scores). Although our data support the previous finding in demonstrating the perceptual saliency of a strong rising pitch contour, the question of whether this type of dynamic pitch can also improve intelligibility of continuous speech without strong emotional context remains to be examined. This question may have practical significance. Although pitch is not manipulated by current assistive listening devices or hearing aids, we do have signal-processing technology that can enhance this cue in speech. If we find a benefit in strong dynamic pitch cue in speech recognition, future generations of listening devices can be designed to implement the processing strategy that makes dynamic pitch (particularly rising pitch contour) stronger. Furthermore, if cognitive or social mechanisms of this phenomenon can be identified, future clinical applications may involve a cognitive assessment and auditory rehabilitation regimen.

Is the Inability to Identify Dynamic Pitch Due to Labeling Difficulty?

It has been argued by earlier research that many speech perception phenomena involve dual processing of both psychophysical perception and a categorizing process (Macmillan et al., 1977; Pollack & Pisoni, 1971). Under this assumption, discrimination performance and identification performance should be comparable. On the other hand, when a listener perceives a stimulus only in a psychophysical manner without a categorizing or labeling process, the ability to discriminate should exceed the ability to identify.

Although identification and discrimination data were highly correlated in our study, we observed a combination

of these two types of listener behaviors. The majority of the listeners had comparable discrimination and identification performance. Although six listeners (of 43 listeners in total) were able to discriminate but unable to identify, it should be noted that this cannot be due to aging alone because two younger listeners also fell into this subgroup. Although this result indicates that some listeners may have used a different strategy in perceiving dynamic pitch, it remains a question for future research whether psychophysical perception alone (without the categorizing/labeling) may be sufficient for a listener to use this cue for speech recognition.

It is worth noting that we used an AAX paradigm (instead of 2IFC) to measure dynamic pitch discrimination for several reasons. First, the nature of the dynamic pitch stimuli (more discrete when compared with static pitch stimuli) makes the 2IFC adaptive procedure not the best choice for measuring dynamic pitch discrimination. Furthermore, the AAX discrimination paradigm has also been used by other studies to test pitch glide perception (e.g., Wayland, Herrera, & Kaan, 2010). The purpose of testing dynamic pitch discrimination in our study was to validate the identification data instead of making a direct comparison with the static pitch discrimination data. Therefore, using different paradigms for measuring static and dynamic pitch discrimination should not pose a problem for interpreting the data.

Is the Interindividual Variability in Dynamic Pitch Perception Related to Musical Experience?

Following the rationale that musical experience improves the central processing of pitch information (Besson et al., 2007; Kraus & Chandrasekaran, 2010; Wong et al., 2007), it is possible that in Souza et al. (2011), the heterogeneity in musical background could have contributed to a large intersubject variability. By including participants that had no or minimum musical experience in the current study, we tested the question of whether homogeneity in musical background would reduce the intersubject variability in the data. Although we did not find evidence for lower variability in the current dataset when compared with the Souza et al. (2011) dataset, further research could shed light on this question by using an identical study setup. Nevertheless, the present study provided a dataset that can serve as a reference point for dynamic pitch perception performance when musical and language experiences are controlled.

It is worth noting that we have additional data (that were not included in the analyses) from four younger participants who had extensive musical experience (three are professional musicians who started musical training by the age of 8 years, the other one is an amateur musician who had 7 years of instrument practice starting at the age of 10 years). The data showed that these musician participants outperformed most of the nonmusical younger participants (see Figure 4). This preliminary data suggests a positive effect of extensive musical training on dynamic pitch perception, which should be further investigated by including a comparison group of musicians.

To summarize, the present study examined dynamic pitch perception in older and younger listeners using synthetic vowels. The older listeners varied substantially across individuals on dynamic pitch perception, even when musical and linguistic experience are controlled. Both groups of listeners performed better in the monophthong condition than in the diphthong condition. Overall, strong pitch contours were perceived better than weak pitch contours, particularly with rising pitch patterns. Taken together with the literature showing that dynamic pitch has a facilitating effect for speech recognition under adverse conditions, our findings raise the question of whether an inability to use this cue may contribute to speech-in-noise difficulty for some individuals, particularly in natural speech and when dynamic pitch cues are relatively weak. These questions should be investigated by future research. The outcome may be beneficial for devising individualized speech-in-noise treatment for older listeners.

References

Arehart, K. H. (1994). Effects of harmonic content on complex-tone fundamental-frequency discrimination in hearing-impaired listeners. *The Journal of the Acoustical Society of America, 95*(6), 3574–3585.

Arehart, K. H., Souza, P. E., Muralimanohar, R. K., & Miller, C. W. (2011). Effects of age on concurrent vowel perception in acoustic and simulated electroacoustic hearing. *Journal of Speech, Language, and Hearing Research, 54,* 190–210.

Assmann, P. F. (1999). Fundamental frequency and the intelligibility of competing voices. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, & A. C. Bailey (Eds.), *Proceedings of the 14th International Congress of Phonetic Sciences* (pp. 179–182). Oakland, CA: University of California Press.

Besson, M., Schön, D., Moreno, S., Santos, A., & Magne, C. (2007). Influence of musical expertise and musical training on pitch processing in music and language. *Restorative Neurology and Neuroscience, 25,* 399–410.

Binns, C., & Culling, J. F. (2007). The role of fundamental frequency contours in the perception of speech against interfering speech. *The Journal of the Acoustical Society of America, 122*(3), 1765–1776.

Bird, J., & Darwin, C. J. (1998). Effects of a difference in fundamental frequency in separating two sentences. In A. R. Palmer, A. Rees, A. Q. Summerfield, & R. Meddis (Eds.), *Psychophysical and Physiological Advances in Hearing* (pp. 263–269). London: Whurr.

Brechmann, A., & Scheich, H. (2005). Hemispheric shifts of sound representation in auditory cortex with conceptual listening. *Cerebral Cortex, 15*(5), 578–587.

Brokx, J. P. L., & Nooteboom, S. G. (1981). Intonation and the perceptual separation of simultaneous voices. *Journal of Phonetics, 10,* 23–26.

Chatterjee, M., & Peng, S. C. (2008). Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition. *Hearing Research, 235*(1), 143–156.

Clinard, C. G., & Cotter, C. M. (2015). Neural representation of dynamic frequency is degraded in older adults. *Hearing Research, 323,* 91–98.

Clinard, C. G., Tremblay, K. L., & Krishnan, A. R. (2010). Aging alters the perception and physiological representation of frequency: Evidence from human frequency-following response recordings. *Hearing Research, 264*(1), 48–55.

Cutler, A. (1976). Phoneme-monitoring reaction time as a function of preceding intonation contour. *Perception & Psychophysics, 20*(1), 55–60.

Dau, T., Wegner, O., Mellert, V., & Kollmeier, B. (2000). Auditory brainstem responses with optimized chirp signals compensating basilar-membrane dispersion. *The Journal of the Acoustical Society of America, 107*(3), 1530–1540.

Dupuis, K. L. (2011). *Emotion in speech: Recognition by younger and older adults and effects on intelligibility* (Unpublished doctoral dissertation). University of Toronto, Toronto.

Dupuis, K., & Pichora-Fuller, M. K. (2010). Use of affective prosody by young and older adults. *Psychology and Aging, 25*(1), 16–29.

Fairbanks, G. (1940). Recent experimental investigations of vocal pitch in speech. *The Journal of the Acoustical Society of America, 11*(4), 457–466.

Foxton, J. M., Weisz, N., Bauchet-Lecaignard, F., Delpuech, C., & Bertrand, O. (2009). The neural bases underlying pitch processing difficulties. *NeuroImage, 45*(4), 1305–1313.

Friedrich, C. K., Kotz, S. A., Friederici, A. D., & Alter, K. (2004). Pitch modulates lexical identification in spoken word recognition: ERP and behavioral evidence. *Cognitive Brain Research, 20*(2), 300–308.

Frisina, D. R., Frisina, R. D., Snell, K. B., Burkard, R., Walton, J. P., & Ison, J. R. (2001). Auditory temporal processing during aging. *Functional Neurobiology of Aging, 39,* 565–579.

Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech, 1*(2), 126–152.

Gordon, M., & Poeppel, D. (2002). Inequality in identification of direction of frequency change (up vs. down) for rapid frequency modulated sweeps. *Acoustics Research Letters Online, 3*(1), 29–34.

Green, T., Faulkner, A., & Rosen, S. (2002). Spectral and temporal cues to pitch in noise-excited vocoder simulations of continuous-interleaved-sampling cochlear implants. *The Journal of the Acoustical Society of America, 112*(5), 2155–2164.

Grose, J. H., & Mamo, S. K. (2012). Frequency modulation detection as a measure of temporal processing: Age-related monaural and binaural effects. *Hearing Research, 294*(1), 49–54.

Haggard, M., Ambler, S., & Callow, M. (1970). Pitch as a voicing cue. *The Journal of the Acoustical Society of America, 47*(2B), 613–617.

Harkrider, A. W., Plyler, P. N., & Hedrick, M. S. (2005). Effects of age and spectral shaping on perception and neural representation of stop consonant stimuli. *Clinical Neurophysiology, 116*(9), 2153–2164.

He, N. J., Dubno, J. R., & Mills, J. H. (1998). Frequency and intensity discrimination measured in a maximum-likelihood procedure from young and aged normal-hearing subjects. *The Journal of the Acoustical Society of America, 103*(1), 553–565.

He, N. J., Mills, J. H., & Dubno, J. R. (2007). Frequency modulation detection: Effects of age, psychophysical method, and

modulation waveform. *The Journal of the Acoustical Society of America, 122*(1), 467–477.

Hillenbrand, J. M. (2003). Some effects of intonation contour on sentence intelligibility. *The Journal of the Acoustical Society of America, 114,* 2338.

Kjelgaard, M. M., & Speer, S. R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory and Language, 40*(2), 153–194.

Klatt, D. H. (1980). Software for a cascade/parallel formant synthesizer. *The Journal of the Acoustical Society of America, 67*(3), 971–995.

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience, 11*(8), 599–605.

Kuhl, P. K., Williams, K. A., & Meltzoff, A. N. (1991). Cross-modal speech perception in adults and infants using nonspeech auditory stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 17*(3), 829–840.

Ladd, D. R. (1996). *Intonational Phonology*. Cambridge: Cambridge University Press.

Laures, J. S., & Bunton, K. (2003). Perceptual effects of a flattened fundamental frequency at the sentence level under different listening conditions. *Journal of Communication Disorders, 36*(6), 449–464.

Lee, J. H., & Humes, L. E. (2012). Effect of fundamental-frequency and sentence-onset differences on speech-identification performance of young and older adults in a competing-talker background. *The Journal of the Acoustical Society of America, 132*(3), 1700–1717.

Lehiste, I. (1976). Suprasegmental features of speech. In N. Lass (Ed.), *Contemporary Issues in Experimental Phonetics* (pp. 225–239). New York: Academic Press.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America, 49*(2B), 467–477.

Macmillan, N. A., Kaplan, H. L., & Creelman, C. D. (1977). The psychophysics of categorical perception. *Psychological Review, 84*(5), 452–471.

McCawley, J. C. (1978). What is a tone language? In V. Fromkin (Ed.), *Tone: A Linguistic Survey* (pp. 113–131). New York: Academic Press.

McCloy, D. R., Wright, R. A., & Souza, P. E. (2014). Talker versus dialect effects on speech intelligibility: A symmetrical study. *Language and Speech 58*(3), 371–386.

Mitchell, R. L., Kingston, R. A., & Barbosa-Bouças, S. L. (2011). The specificity of age-related decline in interpretation of emotion cues from prosody. *Psychology and Aging, 26*(2), 406–414.

Molis, M. R., Srinivasan, N., & Gallun, F. J. (2015). Effects of hearing impairment on sensitivity to dynamic spectral change. *The Journal of the Acoustical Society of America, 137*(4), 2231.

Moore, B. C. J. (1973). Frequency difference limens for short-duration tones. *The Journal of the Acoustical Society of America, 54*(3), 610–619.

Murray, I. R., & Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *The Journal of the Acoustical Society of America, 93*(2), 1097–1108.

Orbelo, D. M., Grim, M. A., Talbott, R. E., & Ross, E. D. (2005). Impaired comprehension of affective prosody in elderly subjects is not predicted by age-related hearing loss or age-related cognitive decline. *Journal of Geriatric Psychiatry and Neurology, 18*(1), 25–32.

Pardo, P. J., & Sams, M. (1993). Human auditory cortex responses to rising versus falling glides. *Neuroscience Letters, 159*(1), 43–45.

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America, 24*(2), 175–184.

Pollack, I., & Pisoni, D. (1971). On the comparison between identification and discrimination tests in speech perception. *Psychonomic Science, 24*(6), 299–300.

Remez, R. E., Fellowes, J. M., Blumenthal, E. Y., & Nagel, D. S. (2003). Analysis and analogy in the perception of vowels. *Memory & Cognition, 31*(7), 1126–1135.

Schodder, G. R., & David, E. E., Jr. (1960). Pitch discrimination of two-frequency complexes. *The Journal of the Acoustical Society of America, 32*(11), 1426–1435.

Sheft, S., Shafiro, V., Lorenzi, C., McMullen, R., & Farrell, C. (2012). Effects of age and hearing loss on the relationship between discrimination of stochastic frequency modulation and speech perception. *Ear and Hearing, 33*(6), 709–720.

Shore, S. E., & Nuttall, A. L. (1985). High-synchrony cochlear compound action potentials evoked by rising frequency-swept tone bursts. *The Journal of the Acoustical Society of America, 78*(4), 1286–1295.

Souza, P., Arehart, K., Miller, C. W., & Muralimanohar, R. K. (2011). Effects of age on F0- discrimination and intonation perception in simulated electric and electro-acoustic hearing. *Ear and Hearing, 32*(1), 75–83.

Stalinski, S. M., Schellenberg, E. G., & Trehub, S. E. (2008). Developmental changes in the perception of pitch contour: Distinguishing up from down. *The Journal of the Acoustical Society of America, 124*(3), 1759–1763.

Studebaker, G. A. (1985). A rationalized arcsine transform. *Journal of Speech and Hearing Research, 28,* 455–462.

Summers, V., & Leek, M. R. (1998). F0 processing and the separation of competing speech signals by listeners with normal hearing and with hearing loss. *Journal of Speech, Language, and Hearing Research, 41,* 1294–1306.

Syka, J., Rybalko, N., Nwabueze-Ogbo, F. C., & Suta, D. (2003). Influence of auditory cortex lesions on the ability to discriminate between rising and falling frequency-modulated tones in rats. *Association for Research in Otolaryngology Midwinter Meeting Abstracts, 26,* 234.

Uldall, E. (1960). Attitudinal meanings conveyed by intonation contours. *Language and Speech, 3*(4), 223–234.

Vongpaisal, T., & Pichora-Fuller, M. K. (2007). Effect of age on F0 difference limen and concurrent vowel identification. *Journal of Speech, Language, and Hearing Research, 50,* 1139–1156.

Watson, P. J., & Schlauch, R. S. (2008). The effect of fundamental frequency on the intelligibility of speech with flattened intonation contours. *American Journal of Speech-Language Pathology, 17,* 348–355.

Wayland, R., Herrera, E., & Kaan, E. (2010). Effects of musical experience and training on pitch contour perception. *Journal of Phonetics, 38*(4), 654–662.

Weber, A., Grice, M., & Crocker, M. W. (2006). The role of prosody in the interpretation of structural ambiguities: A study of anticipatory eye movements. *Cognition, 99*(2), B63–B72.

Wetzel, W., Ohl, F. W., Wagner, T., & Scheich, H. (1998). Right auditory cortex lesion in Mongolian gerbils impairs discrimination of rising and falling frequency-modulated tones. *Neuroscience Letters, 252*(2), 115–118.

Wingfield, A., Lahar, C. J., & Stine, E. A. (1989). Age and decision strategies in running memory for speech: Effects of prosody and linguistic structure. *Journal of Gerontology, 44*(4), 106–113.

Wingfield, A., Lindfield, K. C., & Goodglass, H. (2000). Effects of age and hearing sensitivity on the use of prosodic information in spoken word recognition. *Journal of Speech, Language, and Hearing Research, 43*(4), 915–925.

Wong, P. C., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience, 10*(4), 420–422.