



Published in final edited form as:

Exp Aging Res. 2009 ; 35(1): 129–151. doi:10.1080/03610730802565091.

COMPONENTS OF SPEECH PROSODY AND THEIR USE IN DETECTION OF SYNTACTIC STRUCTURE BY OLDER ADULTS

Ken J. Hoyte,

Volen National Center for Complex Systems, Brandeis University, Waltham, Massachusetts, USA

Hiram Brownell, and

Volen National Center for Complex Systems, Brandeis University, Waltham, Massachusetts, USA;
and Department of Psychology, Boston College, Boston, Massachusetts, USA

Arthur Wingfield

Volen National Center for Complex Systems, Brandeis University, Waltham, Massachusetts, USA

Abstract

Young and older adults heard sentences in which one character was describing another character (“The doctor said the nurse is thirsty”), where the character being described could be determined only by the prosodic pattern in which the sentence was heard. Using computer editing, the authors generated sentences that were heard with either one (Experiment 1) or two (Experiment 2) of three ordinarily co-occurring prosodic features reduced (pitch variation, amplitude variation, timing variation). For both age groups, timing variation was the most valuable of the three prosodic features. These results add to our understanding of the effective preservation of spoken language comprehension in normal aging.

A striking feature of normal aging is the significant sparing of spoken language comprehension in spite of its heavy demands on processing speed, hearing acuity, and working memory capacity, all of which decline to greater or lesser degrees in normal aging (Kausler, 1994; Salthouse, 1994, 1996; Schneider & Pichora-Fuller, 2000). In large part, this is because of older adults’ ability to use spared linguistic knowledge and procedural rules for its use to mitigate what might otherwise be far more severe consequences of these processing limitations (Kemper, 1992; Wingfield & Stine-Morrow, 2000).

Less well studied in aging is listeners’ use of prosody to facilitate speech comprehension. *Prosody* is a generic term that includes the full array of suprasegmental acoustic features that accompany natural speech. These include the intonation pattern (pitch contour) of a sentence that is carried by the fundamental frequency (F_0) of the voice and word stress (a complex subjective variable based on loudness [amplitude]), as well as pitch and syllabic duration. Prosody also includes timing variations, such as the pauses that sometimes occur between major syntactic elements of a sentence. Especially notable is the lengthening of clause-final words, a common marker in English that a clause boundary is being reached (Shattuck-Hufnagel & Turk, 1996). Whether intended for the benefit of the listener or a consequence of production planning (cf., Kraljic & Brennan, 2005; Snedeker & Trueswell, 2003), the prosodic marking that speakers spontaneously produce in their natural utterances can thus aid the listener in syntactic analysis of speech, an essential aspect of language comprehension.

Studies conducted since the pioneering work of Cohen and Faulkner (1986) have shown that older listeners are no less sensitive to prosody than their younger counterparts. Speakers' awareness of this preserved sensitivity may help explain so-called "elderspeak," a simplification of syntax and exaggeration of speech prosody often observed in individuals speaking to older adults (Kemper, 1994; Kemper & Harden, 1999). At the word level, older adults have been shown to make as good use of syllabic stress in spoken word identification as young adults (Wingfield, Lindfield, & Goodglass, 2000) and, at the sentence level, older adults have been shown to use prosody to identify syntactic boundaries that in turn tell a listener how to "chunk" incoming speech into its meaningful elements (Kjelgaard, Titone, & Wingfield, 1999; Wingfield, Lahar, & Stine, 1989; Wingfield, Wayland, & Stine, 1992).

To show that older adults can make effective use of prosody, however, does not tell us that they necessarily use each of the major prosodic features (variations in amplitude, pitch, and timing) in the same way as do young adults. Past work that has considered prosody as a whole may have simply collapsed over interesting differences in how normal aging affects use of the individual components of prosody and their combination to support sentence comprehension. This is the question we ask in this present paper.

Our experiments are motivated by two considerations. One of these is a much disputed and controversial claim that adult aging is accompanied by a differential decline in right hemisphere functions (cf., Cherry & Hellige, 1999; Nebes, 1990; Orbelo, Grim, Talbott, & Ross, 2005; see the review in Kempler, 2005). Such a decline might, for example, result in older adults being less sensitive to the utilization of pitch contour, a feature more effectively processed by the right hemisphere (e.g., Van Lancker & Sidtis, 1992; Zatorre, Evans, Meyer, & Gjedde, 1992). The second consideration relates to age-related audiologic changes that can result in a decline in efficacy of temporal and frequency resolution (Schneider & Pichora-Fuller, 2000). Such changes might raise the question of a selective impairment in sensitivity to timing or pitch variation. Declining efficiency due to either of these factors could be obscured by compensatory use of co-occurring prosodic features.

In natural speech, the three major prosodic features provide cues that typically co-occur as a speaker reaches a clause boundary (Shattuck-Hufnagel & Turk, 1996). Because of this natural redundancy, the absence of any one prosodic feature can be compensated for at least in part by the presence of others (Culler & Darwin, 1981; Streeter, 1978). This inherent redundancy forms a backdrop to our experimental question. When there is redundancy, as in the case of prosodic marking of syntactic boundaries, the loss or removal of one feature may not necessarily produce a total failure in performance. It may result in only a relative deficit, or, under some circumstances, no deficit at all. To anticipate our results, one should not necessarily frame the question in terms of whether young and older adults use the same prosodic features to aid detection of linguistic boundaries in spoken utterances, but whether young and older adults give the same relative weight to each of these features. Because of the ordinary co-occurrence of prosodic distinctions in natural speech, the previously noted finding that older adults can make at least as good use of prosody as young adults for syntactic resolution does not tell us whether or not they may be relying particularly on a single cue to carry their comprehension success.

We took two approaches to this question. The first was to examine young and older adults' accuracy and speed of detecting clause boundaries in potentially ambiguous sentences when computer editing was used to reduce original variations in either amplitude (a major feature of word stress), pitch contour (variation in fundamental frequency of the voice across a sentence), or timing cues (pauses and relative durations of critical words). In this case, when one feature was removed, the other two features were always presented intact as they were in the original normal-prosody sentence (Experiment 1). The second approach was to ask the same question

when only a single feature was presented intact with variation in the other two features reduced (Experiment 2).

The stimuli chosen for this study were sentences in which one character is describing some attribute of another character, as in “The queen, said the king, is hungry.” The listener’s task was to say as quickly as possible who was being described, in this example, the queen or the king. The unifying characteristic of these sentences was that determination of which of the two characters was being described was only possible by using the particular prosodic pattern with which the sentence was heard.

EXPERIMENT 1

Methods

Participants—The older adults were 20 community-dwelling volunteers, 7 men and 13 women, who ranged in age from 63 to 83 ($M = 72.3$, $SD = 5.9$). The group had a mean of 16.8 years of formal education ($SD = 2.5$) and a mean Wechsler Adult Intelligence Scale (WATS-III; Wechsler, 1997) vocabulary score of 53.4 ($SD = 5.7$). The young adult participants were 20 university undergraduate and graduate students, 9 men and 11 women, with ages ranging from 19 to 28 ($M = 21.9$, $SD = 2.6$). The group had a mean of 15.5 years of formal education at time of testing ($SD = 1.8$) and a mean WAIS-III vocabulary score of 52.7 ($SD = 6.2$). Both groups were thus well-educated and had good vocabularies, with the older group having an average of 1.3 more years of formal education, $t(38) = 1.58$, $n.s.$, and, as is common for older adults (e.g., Verhaeghen, 2003), a somewhat higher vocabulary score than the younger group, although not significantly so, $t(38) = 1.28$, $n.s.$ All participants were native speakers of American English, and all reported themselves to be in good health, with no known history of stroke, Parkinson’s disease, or dementing illness. Participants were also free of medications that might compromise cognitive function.

The older adults in this study had good hearing for their ages (Morrell et al, 1996). Audiometric screening insured that all participants’ pure tone averages (PTAs; averaged across 500, 1000 and 2000 Hz), and speech reception thresholds (SRTs; lowest decibel level at which two-syllable words can be correctly identified 50% of the time), were less than 25 dB HL (Hearing Level) in the better ear. This level is considered to be clinically normal for speech (Hall & Mueller, 1997).

Stimulus Materials—The stimuli consisted of recorded spoken sentences, each of which contained two characters, with one character giving the listener information about the other character. The sentences were constructed such that which character is being described can be determined only by reference to the prosodic pattern with which a sentence was spoken. Each sentence was spoken with prosodic marking indicating either a canonical interpretation (e.g., “The doctor said the maid is thirsty” [the maid is thirsty]), or a noncanonical interpretation (e.g., “The doctor, said the maid, is thirsty” [the doctor is thirsty]). We refer to the first version as canonical because it is syntactically less computationally demanding than the noncanonical version that expresses the meaning using an embedded clause. It may also be the more expected form based on the frequency with which these simpler forms are encountered in everyday speech (MacDonald, Pearlmutter, & Seidenberg, 1994). At points of uncertainty as the sentence is unfolding in time there is thus a bias toward a canonical interpretation; a default interpretation unless prosody or a linguistic context makes it clear that this is not the intended meaning (Speer, Kjelgaard, & Dobroth, 1996).

Figure 1 shows waveforms and pitch contours for a standard utterance of a canonical (top panel) versus a noncanonical (bottom panel) version of this sample sentence. In the canonical version one sees a relatively long pause after the first clause (*the doctor said*) with a lengthened duration

of the clause-final word (*said*). In the noncanonical version one sees a reduced pause between the end of the initial noun phrase (*the doctor*) and the beginning of the embedded clause, *said the maid*, but an added short pause between the embedded clause and the final verb phrase (*is thirsty*). One also sees, in the top panel, the larger amplitude of the word *maid* when it is the person being described in the canonical form, versus its smaller amplitude when it is spoken as part of the embedded clause (as a quotative) indicating that it is the maid who is doing the describing. Below each panel is a pitch track of the fundamental frequency (Fo) where one also sees characteristic differences in pitch pattern across the canonical and noncanonical versions.

Each of the sentences but one contained seven words and took the form of, “The + *a character identified by occupation* + *said* + the + *a different character identified by occupation* + *is* + *descriptor*.” The italicized words varied between sentences; the nonitalicized words were the same across sentences. The sentences were constructed from 10 persons (*king, queen, nurse, knight, bride, juggler, teacher, dancer, doctor, and maid*) and 10 different descriptors (*hungry, worried, honest, sick, friendly, thirsty, nice, liar, kind, and tired*). Due to a preparation error one of the sentence sets contained as a descriptor a noun phrase (*a liar*) rather than an adjective, which with its article, resulted in an eight-word sentence. This collection of characters and descriptors was used to generate a set of 40 different sentences.

Each of the 40 sentences was recorded by a male speaker in either a prosodic pattern indicating the canonical or the noncanonical pattern. (The canonical and noncanonical versions of a sentence differed only in the prosodic pattern.) This procedure yielded a total of 80 sentences representing each of the 40 sentences spoken to indicate its canonical interpretation and its noncanonical interpretation. Waveform analyses of each of the stimulus sentences confirmed that the prosodic patterns in these utterances followed standard prosodic marking for these two interpretations (Shattuck-Hufnagel & Turk, 1996; Speer et al., 1996). The recordings of these original normal-prosody sentences were presented to a group of young adult listeners to insure that each rendering unambiguously signaled its appropriate interpretation. These recordings were then digitized onto computer sound files and then modified using filtering and speech-editing software so as to be heard in one of three altered prosodic patterns: reduction of *amplitude variation*, reduction of *pitch variation*, and reduction of *timing variation*.

Reduction of amplitude variation: Waveform analysis of the original spoken sentences showed that the character being described received word stress signaled by an amplitude level that was on average 30% larger than the mean amplitude of the other six words in the sentence. For this condition, SoundEdit speech editing software (Macromedia, San Francisco, CA) was used to reduce the amplitude of the character being described by 30%. This produced sentences in which differences in the pattern of amplitude variation among the words was minimized as an aid to parsing the sentence in order to identify the target character. Both pitch variation and timing features, however, remained unaltered.

Reduction in pitch variation: For this condition, Praat software (Boersma & Weenink, 2005) was used to reduce pitch variation. To perform this manipulation, Praat applies an algorithm that presents for editing a graphic display of mean pitch points at intervals of 0.01 s. These points were manually moved to the mean fundamental frequency of the speaker’s voice (130 Hz). This produced a global flattening of the fundamental frequency, such that the sentence would be heard without the normally accompanying pitch variation that might be used by a listener to identify which of the two characters in the sentence is the person being described. The procedure, however, preserved both the amplitude and timing variations of the original utterances.

Reduction in tuning variation: Analysis of the waveforms of the original normal-prosody sentences showed that the utterance duration of a character name was on average 30% longer when that character was the person being described (the target character) than when it was the nontarget character in the sentence. SoundEdit software was used to select the target character in the visual waveform of each sentence, and then to reduce the vowel duration of the word by 30%. In an analogous way, pauses that normally occurred immediately preceding the target character in the canonical sentences and immediately following the target character in the non-canonical sentences were also reduced by 30%. This procedure thus largely eliminated the distinction in word and pause duration that might ordinarily signal the target character. Both the amplitude and pitch variations of the original utterances were left intact.

Intelligibility check: We conducted a preliminary control experiment with a group of young adults to insure that our editing of the materials did not affect general speech intelligibility (e.g., Laures & Weismer, 1999). All of the sentences in all conditions could be reported correctly without error. None of the participants who served in this control experiment was tested in the main experiment.

Procedure—Testing was conducted individually in a sound-attenuated testing room, with the stimuli presented binaurally over earphones. Presentation levels were individually set for each participant in a pretest so as to be at a comfortable level that allowed the participant to hear the stimuli without effort. Once this level was selected, it was maintained for that participant throughout the experiment.

To avoid fatigue, the experiment was conducted in two blocks of 40 sentences per block, with a short break between the two blocks. Across the two blocks half of the stimulus items were spoken with a prosodic pattern consistent with a canonical interpretation (the second character is the one being described) and half with a prosodic pattern indicating a noncanonical interpretation (the first character is the one being described). Within each canonicity condition (40 canonical, 40 noncanonical), participants heard 10 of the sentences with their original normal prosody and 10 each with amplitude, pitch, and timing variation reduced. Within each block, each cell of the design contained an equal number of items, with the combinations of characters and adjectives used in the stimulus items systematically varied. Sentences were presented to participants in a pseudorandom order mixed-list format in order to reduce any syntactic priming that might occur from a prior sentence interpretation (e.g., Bock & Griffin, 2000; Pickering & Branigan, 1998).

Participants were told that their task would be to say aloud, as quickly as possible, the name of the character being described in that sentence. Participants were told to give their answers as soon as they could do so with confidence, even if the sentence had not been fully heard. Responses were digitally recorded onto computer sound files for later measurement of response latencies to the correct responses. The relative value of a prosodic feature would be indicated by an increase in response latency to the correct response when variation in that particular feature was reduced. The main experiment was preceded by a brief practice session to familiarize the participants with the sounds of the prosodic patterns and the task instructions. None of these sentences was used in the main experiment.

Results

Response Accuracy—Table 1 shows the percentage of correctly parsed sentences for each of the prosody conditions for the young and older adults for the canonical and noncanonical sentences. It can be seen that both young and older adults were generally very accurate in identifying which of the two characters was being described. For the canonical sentences, both age groups were at or near ceiling for all conditions. For this reason, the canonical accuracy

data were not analyzed further. For the noncanonical sentences, response accuracy was also very high (90% or better in all conditions). Analysis of variance (ANOVA) performed on the noncanonical sentence data revealed a significant main effect of prosody condition, $F(3, 114) = 10.44$, $MSE = .005$, $p < .01$; however, neither the main effect of age nor the Age \times Prosody interaction approached significance, although the near ceiling level performance limits the interpretability of these null results.

For the noncanonical sentences, paired comparison testing against the original normal prosody sentences showed the young adults to exhibit a small but significant decline in accuracy when either pitch or timing variation was reduced ($p < .01$ in both cases). In the case of the older adults, none of the three prosody reduction conditions differed significantly from the original normal prosody condition. Because participants' accuracy levels were at or near ceiling for canonical sentences, and were also uniformly high for noncanonical sentences, we focus our interpretation on the more sensitive latency data.

Response Latency—Figure 2 shows the mean latencies to correct responses for the young and older adults for the canonical (left panel) and noncanonical (right panel) sentences when amplitude, pitch, and timing variation were each reduced. The mean latency levels for the original normal-prosody sentence utterances are indicated by the horizontal dotted lines. Because the condition in which timing variation was reduced resulted in sentences that were shorter in duration than the other three conditions, the latencies for all sentences were converted to proportions. This was accomplished by measuring response latencies from the start of the utterance to the onset of the response and then dividing by the total sentence duration. Proportional latencies shown as less than 1.0 in Figures 2 and 3 indicate that a correct response was produced before the sentence had been fully completed.

Canonical Sentences—For the canonical sentences (left panel of Figure 2), one sees a tendency for older adults to take somewhat longer to respond than the young adults. This tendency was reflected in an ANOVA conducted on the canonical sentence latency data that showed a marginal main effect of age, $F(1, 38) = 3.65$, $MSE = .078$, $p = .064$. The factor of primary interest in this experiment was a significant main effect of prosody condition, $F(3, 114) = 23.68$, $MSE = .005$, $p < .001$. Consistent with the similarity of the general pattern of effects for the young and older adults for the canonical sentences, the Age \times Prosody Condition interaction failed to reach significance, $F(3, 114) = 2.07$, *n.s.*

Despite the absence of an Age \times Prosody Condition interaction, we examined simple effects of prosody condition for the younger and older participants separately to assess in detail the similarity in patterns for the two age groups. Pairwise comparisons conducted on the young adults' data showed a significant increase in response latency relative to sentences heard with normal prosody when either pitch variation, $t(19) = 3.11$, $p < .01$, $r = .58$, or timing variation, $t(19) = 4.04$, $p < .001$, $r = .68$, was reduced. Of the two, reducing timing variation produced slightly, but significantly, longer latencies than reducing pitch variation, $t(19) = 3.97$, $p < .01$, $r = .67$. For the older adults, responses were also slowed significantly by either reduced pitch variation, $t(19) = 3.59$, $p < .01$, $r = .64$, or reduced timing variation, $t(19) = 3.01$, $p < .01$, $r = .57$. The increase in response latencies with reduced timing variation relative to reduced pitch variation was in the same direction for older as for young adults, but was not significant in the older adults' data, $t(19) = 1.01$, *n.s.*, $r = .23$. Another subtle difference between young and older group results was that reduced amplitude variation reliably speeded up the older adults' responses relative to the normal prosody condition, $t(19) = 2.30$, $p < .05$, $r = .47$, while the same effect was not significant in the young adults' data.

Noncanonical Sentences—The pattern of effects seen for the canonical sentences was amplified for the noncanonical sentences (Figure 2, right panel). An ANOVA conducted on

the latency data for the noncanonical sentences showed significant main effects of age, $F(1, 38) = 6.86$, $MSE = .190$, $p < .05$ and of prosody condition, $F(3, 114) = 49.72$, $MSE = .001$, $p < .001$. As with the canonical sentences, there was no significant Age \times Prosody Condition interaction, $F(3, 114) < 1$.

Simple effects in the form of pairwise comparisons were again conducted on the young and older adults' data separately. Results corroborated the similarity between the young and older group patterns and were largely similar to the patterns reported above for the canonical sentences. Relative to the normal-prosody condition, reduction in pitch variation slowed responses for the young adults, $t(19) = 2.98$, $p < .01$, $r = .56$, and for the older adults, $t(19) = 6.01$, $p < .001$, $r = .81$. Similarly, reduction in timing variation slowed responses for the young, $t(19) = 6.90$, $p < .001$, $r = .85$, and older adults, $t(19) = 6.53$, $p < .001$, $r = .83$. Reducing amplitude variation again failed to have a significant effect for the young adults and had only a marginal effect for the older adults, $t(19) = 1.86$, $p = .079$, $r = .39$.

Overall, for the younger adults reducing timing variation was most disruptive, slowing responses to a significantly greater extent than a reduction in pitch variation, $t(19) = 3.97$, $p < .001$, $r = .67$, although both differed significantly from latencies in the reduced amplitude condition ($p < .01$ for pitch and $p < .001$ for timing). For older adults, reducing timing variation was also the most disruptive; the impact of reducing timing variation was greater than reducing pitch variation ($p < .01$), as well as greater than the effect of reducing amplitude ($p < .001$). The effect of reducing pitch variation was greater than that of reducing amplitude variation ($p < .05$). These results underscore a generally similar use of prosodic features by older and younger adults.

A separate analysis was conducted to examine a different question: whether older participants were less adaptable in terms of their response to the experimental task. This analysis was restricted to data from the normal prosody condition. An ANOVA including age (young, older) and sentence type (canonical, noncanonical) as factors showed a reliable Age Group \times Sentence Type interaction, $F(1, 38) = 5.62$, $MSE = .009$, $p < .05$, $r = .36$. The pattern of means seen in Figure 2 within this interaction showed that the young participants responded faster to the noncanonical sentences than the canonical sentences, whereas the older participants showed no difference.

This interaction is what one might expect if the young adults were better at developing a strategy of listening for the absence of a distinctive pause or pitch signal to mark the noncanonical interpretation. Because this signal should occur early in the utterance, participants could know the correct answer before the end of the stimulus sentence. One explanation is that the older adults as a group were less likely to develop and to capitalize on this strategic approach to the task and, as a result, produced latencies closer to 1.0; that is, responding at the end of the stimulus sentence—in both the canonical and noncanonical sentences. Another possibility is that the older adults were less confident than the young participants and, for that reason, did not capitalize on the available strategy.

Discussion

The data from Experiment 1 reveal two major features of young and older adults' use of prosody to aid in correct syntactic parsing of an otherwise ambiguous sentence. The first of these derives from the redundancy carried by the co-occurrence of the speech prosody components that ordinarily accompany spoken sentences (Shattuck-Hufnagel & Turk, 1996). The use of context-free ambiguous sentences such as the present stimuli provide a critical test of listeners' ability to take advantage of this redundancy in which sentence prosody contains more than one source of information about correct sentence parsing. As we indicated, a reduction in either amplitude, pitch, or timing information still allowed for high accuracy in parsing the sentences

for character identity for both young and older adults. This finding illustrates the richness of the prosodic information that may be available in an utterance. It also suggests that it is this characteristic “over-design” of natural speech that contributes to the robustness of speech comprehension in older adults that has been so frequently noted in the aging literature (see the review in Wingfield & Stine-Morrow, 2000).

A second major finding from Experiment 1 was a clear hierarchy in the relative value of the three prosodic features investigated. The most powerful feature was the lengthening of clause-final words that precede, and the brief pauses that follow, major clause boundaries (Shattuck-Hufnagel & Turk, 1996; Speer et al., 1996). The second most valuable prosodic marking was supplied by the pitch contour of the utterance, with amplitude variation being the least useful for correctly parsing this type of syntactically ambiguous sentence.

In summary, not only do older adults make good use of prosody for speech comprehension, confirming findings by Cohen and Faulkner (1986), Kjelgaard et al. (1999), and Wingfield et al. (1989, 1992), but we now see that older adults also give the same relative weighting to the specific components of prosody as do their younger counterparts.

Interestingly, members of both age groups were often able to give the correct response before the sentence had been fully completed (revealed as a proportional latency in Figure 2 of less than 1.0). This was somewhat more in evidence for the noncanonical sentences for which, after a well-marked first clause, one may already have enough information to make a correct parsing decision. Although canonical sentences are syntactically simpler and more common in everyday experience than their noncanonical forms, they do not have such a clear prosodic indicator coming so early in the utterance. This suggests that the longer latencies for canonical than noncanonical sentences may have reflected listeners’ delay in responding until the utterance of the second character had been encountered. It is important to note that this feature of the participants’ response latencies was particularly clear in the young adults’ latencies. That is, one difference that emerged between the young and older adults was that the young participants appeared more adept at developing and applying strategies to maximize performance in this novel experimental task.

Although these results were clear when a single feature was removed, we replicated the experiment with a new group of young and older adults, but this time contrasting original utterances spoken with appropriate prosody with the same sentences in which only a single prosodic feature was available intact for use. This second experiment provides a stronger test of the information value carried by each individual prosodic component. That is, the second study tests how well a listener can understand an utterance when he or she cannot take advantage of the redundancy ordinarily available from additional intact prosodic features.

EXPERIMENT 2

Methods

Participants—Participants were a different group of 20 older adults (8 men and 12 women), who ranged in age from 65 to 85 ($M=76.2$, $SD > = 4.7$), and 20 younger adults (7 men and 13 women), with ages ranging from 18 to 28 ($M = 21.7$, $SD = 2.6$). The groups were similar to those in Experiment 1 in terms of mean years of formal education (young adults = 15.7 years, $SD = 2.0$; older adults = 16.7 years, $SD = 2.7$, $t(38) = 1.25$, *n.s.*) and mean vocabulary scores (young adults = 54.4, $SD = 6.8$; older adults = 54.3, $SD = 7.5$, $t(38) = .045$, *n.s.*). All participants were again native speakers of American English, with audio-metric screening insuring that all participants had speech-range PTAs and SRTs of less than 25 dB HL in the better ear (Hall & Mueller, 1997).

Stimuli and Procedures—The stimuli for this experiment consisted of the same original normal-prosody sentences used for Experiment 1, except in this case computer editing was used to reduce variation in two of the prosodic features at a time, leaving only a single feature (amplitude, pitch, or timing variation) unaltered. Instructions were again to say aloud the name of the character being described as quickly as possible. As in Experiment 1, responses were digitally recorded onto computer sound files for latency measurements to the correct response. An intelligibility check was conducted for these sentences following the same procedures as described for Experiment 1 to insure that our prosody manipulations did not interfere with general intelligibility.

Results

Response Accuracy—Table 2 shows the mean percentage of correct character identifications for the canonical and noncanonical sentences when they were heard in the original normal prosody or when only a single prosodic feature (amplitude variation, pitch variation, or timing variation) was present intact. As can be seen, both the young and older adults were very accurate in correctly parsing the sentences in all four prosody conditions. For the canonical sentences, both age groups were again at or near ceiling for all conditions. For this reason the canonical accuracy data were not analyzed further.

For the noncanonical sentences, response accuracy was also high. An ANOVA performed on the noncanonical sentence data revealed a significant main effect of prosody condition, $F(3, 114) = 16.19$, $MSE = .01$, $p < .001$. This was true for both age groups, with no main effect of age ($F < 1$) and no Age \times Prosody interaction ($F < 1$). For these noncanonical sentences, paired comparison testing against the original normal prosody sentences showed the young adults to exhibit a small but significant decline in accuracy when either only amplitude variation ($p < .05$) or timing variation ($p < .05$) were available intact. For the older adults, accuracy was slightly but significantly reduced relative to the original normal prosody condition when the sentences were heard with only amplitude variation ($p < .01$), timing variation ($p < .05$) or pitch variation ($p < .05$) presented intact. Because participants' accuracy levels were again at or near ceiling for canonical sentences, and were also very high for noncanonical sentences in some prosody conditions, we again focus our interpretation on the more sensitive latency data.

Response Latencies—Figure 3 shows the mean latencies to correct responses for the young and older adults for the canonical (left panel) and noncanonical (right panel) sentences when only amplitude, pitch, or timing variation were preserved intact. The mean latency levels for the original normal-prosody sentences are indicated by the horizontal dotted lines. Latencies are again shown converted to proportions.

Canonical Sentences—For the canonical sentences (left panel, Figure 3), one sees a tendency for older adults to take somewhat longer to respond than the young adults. This was reflected in a marginal main effect of age, $F(1, 38) = 3.28$, $MSE = .085$, $p = .078$. There was an effect of prosody condition on response latencies, $F(3, 114) = 3.47$, $MSE = .004$, $p < .05$, although this effect was primarily in longer response latencies for the older adults when only amplitude variation was available. This was evidenced by a significant Age \times Prosody Condition interaction, $F(3, 114) = 4.66$, $MSE = .004$, $p < .01$.

These trends were supported by simple effects pairwise comparisons, which showed no significant differences among the four prosody conditions for the young participants. That is, we see that the young participants were able quickly to parse the canonical sentences based on minimal prosodic information. In the case of the older adults, the amplitude variation condition resulted in longer latencies to correct responses than in the normal-prosody condition, $t(19) =$

2.45, $p < .05$, $r = .49$. The pitch and timing conditions did not differ significantly from each other or from the normal-prosody condition.

Noncanonical Sentences—The differentially lesser value of amplitude variation relative to the other prosodic features that was seen for the older adults responding to the canonical sentences was observed for both age groups for the noncanonical sentences. When amplitude variation was the only prosodic feature available intact, it was the least valuable as a guide for syntactic parsing. The effect of age was not significant, $F(1, 38) = 1.75$, $n.s.$, but there was a significant main effect of prosody condition, $F(3, 114) = 33.24$, $MSE = .004$, $p < .001$. A nonsignificant Age \times Prosody Condition interaction, $F(3, 114) = 1.61$, $n.s.$, was consistent with the similarity in the pattern of prosody effects for the young and older participants.

Simple effects pairwise comparisons showed that correct responses when only amplitude variation remained intact were significantly slowed relative to responses in the normal-prosody condition for both the younger, $t(19) = 7.56$, $p < .001$, $r = .87$, and older, $t(19) = 5.10$, $p < .001$, $r = .76$, participants. For the young adults availability of just intact pitch variation, $t(19) = 2.16$, $p < .05$, $r = .44$, or just intact timing variation, $t(19) = 4.04$, $p < .01$, $r = .68$, also resulted in significantly slowed responses relative to responses in the normal-prosody condition, and there was a small but significant difference between the pitch and timing variation conditions, $t(19) = 2.14$, $p < .05$, $r = .44$, both of which were faster than the amplitude variation condition ($p < .001$ and $p < .01$, respectively). The same effects were not reliable for the older adults: pitch and timing variation were equally valuable as aids to correct syntactic parsing, with neither pitch nor timing variation differing significantly in their latencies from the normal-prosody condition nor from each other.

As in Experiment 1, the response latencies for the normal prosody condition were examined separately in a 2 (age group: young, older) \times 2 (sentence type: canonical, noncanonical) ANOVA. There was a marginal interaction between age group and sentence type, $F(1, 38) = 3.34$, $MSE = .006$, $p = .076$, $r = .28$, reflecting a slightly more pronounced tendency for the young participants to be faster to respond to the noncanonical sentences than to the canonical sentences, relative to the older adults. This finding replicates the normal-prosody pattern as seen in Experiment 1.

Discussion

Assessing listeners' weighting of the three major prosodic features for syntactic parsing can be conducted in two ways. In Experiment 1, we examined response latencies to correct character identification when one prosodic feature was reduced while the other two features remained intact. Experiment 2 sought converging evidence by examining response latencies when only a single prosodic feature was available intact for potential use. The generally high accuracy in both experiments demonstrated that any of these prosodic features could be used to indicate the correct sentence interpretation. In both experiments, however, the relative value of a feature would be evidenced when its removal increased response latencies relative to a control condition where all three normally co-occurring prosodic features were available for potential use.

Focusing on the noncanonical sentences, which would be the most sensitive to prosody effects, one sees that the results of Experiment 2 reinforced the findings of Experiment 1 with regard to the subsidiary role played by amplitude variation for correct parsing. In this case, we note that when only amplitude variation was available intact, the speed of correct character identification was slower than that for the normal-prosody condition and slower than for either of the other two prosodic feature conditions.

The relation between the two most valuable features of pitch and timing variation was somewhat different from that observed in Experiment 1. In Experiment 2, pitch and timing variation did not differ significantly from each other in their effect on response speed except for the young participants responding to noncanonical sentences. One possible explanation for this discrepancy in pattern across Experiments 1 and 2 may rest on listeners' modes of responding to each in the two experiments. Although in natural speech, pitch and timing features typically co-occur (Shattuck-Hufhagel & Turk, 1996; Speer et al., 1996), their interdependence may be asymmetric. For example, if timing is the most salient cue, when it co-occurs with pitch or amplitude variation as in Experiment 1, its removal might be expected to cause most disruption in a mixed list format that encourages participants to listen for all cues. In contrast, in Experiment 2, where only a single cue was available intact, participants may have developed a different perceptual set and had been prepared to listen to a single source of information on each trial. In this case, pitch variation was as reliable, or slightly more reliable, as a cue as was timing variation, although the difference between the utility of pitch and timing was slight.

One way in which the older adults appeared to differ qualitatively from the young adults in Experiment 2 was when only amplitude variation was available intact in the canonical sentences. Here the older adults took longer to respond in the amplitude-only condition relative to the normal prosody condition, whereas the young participants did not. It may be that this rapid responding by the young adults, as well as the young adults' differentially faster responses for noncanonical sentences relative to the older adults, was a reflection of the young adults' development of a situation-specific strategy for responding. This was seen in the young adults' somewhat greater tendency across all four prosody conditions to respond before the end of the stimulus sentence, as reflected in the proportional latency measures of less than 1.0. That is, the young and older adults in both experiments responded to speech prosody and each of their individual components in highly similar ways, unless there was an experiment-bound opportunity to develop a situation-specific strategy for optimal performance.

GENERAL DISCUSSION

As we earlier noted, older adults, at least in the absence of significant neuropathology or sensory decline, make good use of naturally occurring prosody to aid rapid speech processing at the sentence level (Cohen & Faulkner, 1986; Kjølgaard et al., 1999; Wingfield et al., 1989, 1992). The present study adds to this finding by showing that older adults not only make good use of speech prosody to aid comprehension at the sentence level, but that older adults maintain the same general pattern of relative weighting given to the specific features of the speech prosody as do young adults.

The stimulus sentences were intentionally chosen to be semantically reversible, such that either target could be the one described, distinguishable only by the prosodic marking. These can be contrasted with nonreversible sentences, such as "The nurse said the book is readable" in which only one interpretation is possible. As such, the sentence types used in the current study allowed a critical test of listeners' ability to use prosody and each of its usually co-occurring components. This finding adds to our understanding of the support available to older adults for language comprehension that helps maintain this faculty at a level that often exceeds what might otherwise be predicted from declining efficiency in sensory processing (Schneider & Pichora-Fuller, 2000), and supporting cognitive function such as working memory (Salthouse, 1994).

Older adults appeared slightly less adept than young adults at developing task-specific strategies that allowed faster responding to noncanonical sentences. In addition, older adults exhibited slightly less flexibility than young adults in adjusting to the intact amplitude

variation-alone condition in Experiment 2. These effects are consistent with reports that older adults sometimes show reduced flexibility in employing a task-specific strategy relative to young adults (e.g., Chasseigne, Mullet, & Stewart, 1997). Of course, this age effect may also reflect decreased confidence in the older adults, leading to slower response latencies. However, regardless of the source of the effect, this difference related to age appears in the context of a main finding of similarity in the use of prosody for sentence parsing by young and older adults, both in general, and in regard to the relative value of each of the specific prosodic features tested.

One of these similarities was the effectiveness of timing variation as an indicator of the clausal structure of a sentence and hence its correct interpretation. It should be noted in this regard that our experimental materials did not allow assessment of the individual effects of variation in the duration of key words versus the duration of the pauses that also mark clausal boundaries, as the two were linked when effects of timing variation were examined. We can only assert that together they play a powerful role for both age groups as an indicator of syntactic parsing for the types of sentences employed in this study.

The notion that right hemisphere functions show differential decline in adult aging, an idea that has a long but controversial history (cf., Cherry & Hellige, 1999; Nebes, 1990; Orbelo et al., 2005; see the review in Kempler, 2005) receives little encouragement from these data. Lesion and functional imaging studies have suggested that the right hemisphere is especially effective in processing fundamental frequency (F₀), which would be necessary for utilization of pitch contour, whereas temporal cues dependent on timing distinctions are more effectively processed in the left hemisphere (e.g., Van Lancker & Sidtis, 1992; Zatorre et al., 1992). Although there were some differences in relative weightings of prosodic feature use when listeners had two versus one prosodic component available intact, the similarity in pattern of effects across the two age groups was striking. To the extent that these presumed hemispheric processing asymmetries are correct, the older adults' effective utilization of pitch variation does not suggest a differential right hemisphere weakness, at least in this regard.

Our present results illustrate the dissociability of components of prosody in natural speech, components that ordinarily co-occur to add useful redundancy to the speech signal. This overall redundancy contributes to the robustness of syntactic parsing in adult aging. This resilience has been shown in the present data to include a close parallel in the relative usefulness for syntactic parsing of the three ordinarily co-occurring prosodic features tested, and adds to our understanding of how spared function in normal aging compensates for other loss. The usefulness of these features in cases of significant sensory or cognitive impairment remains to be determined.

Acknowledgments

The authors acknowledge support from NIH grant AG04517 from the National Institute on Aging (A.W.), and an American Psychological Association Diversity Program in Neuroscience Pre-Doctoral Fellowship (K.J.H.). The authors also gratefully acknowledge support from the W.M. Keck Foundation.

References

- Bock K, Griffin ZM. The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General* 2000;129:177–192. [PubMed: 10868333]
- Boersma, P.; Weenink, D. Praat: Doing phonetics by computer (Versopm 4.3.30) [computer program]. 2005. Retrieved from <http://www.praat.org>
- Chasseigne G, Mullet E, Stewart TR. Aging and multiple cue probability learning: The case of inverse relationships. *Acta Psychologica* 1997;97:235–152.

- Cherry BJ, Hellige JB. Hemispheric asymmetries in vigilance and cerebral arousal mechanisms in younger and older adults. *Neuropsychology* 1999;13:111–120.
- Cohen G, Faulkner D. Does “elderspeak” work? The effect of intonation and stress on comprehension and recall of spoken discourse in old age. *Language and Communication* 1986;6:91–98. 151.
- Culler A, Darwin CJ. Phoneme monitoring reaction-time and preceding prosody: Effects of stop closure duration and of fundamental frequency. *Perception and Psychophysics* 1981;29:217–224. [PubMed: 7267273]
- Hall, J.; Mueller, G. Audiologist desk reference. San Diego, CA: Singular Publishing; 1997.
- Kausler, DM. Learning and memory in normal aging. San Diego, CA: Academic Press; 1994.
- Kemper, S. Language and aging. In: Craik, FIM.; Salthouse, TA., editors. *Handbook of aging and cognition*. Hillsdale, NJ: Erlbaum; 1992. p. 213-270.
- Kemper S. Elderspeak: Speech accommodations to older adults. *Aging and Cognition*, I 1994:17–28.
- Kemper S, Harden T. Experimentally disentangling what’s beneficial about elderspeak from what’s not. *Psychology and Aging* 1999;14:656–670. [PubMed: 10632152]
- Kempler, D. *Neurocognitive disorders in aging*. Thousand Oaks, CA: Sage; 2005.
- Kjelgaard MK, Titone D, Wingfield A. The influence of prosodic structure on the interpretation of temporary syntactic ambiguity by young and elderly listeners. *Experimental Aging Research* 1999;25:187–207. [PubMed: 10467511]
- Kraljic T, Brennan SE. Prosodic disambiguation of syntactic structure: For the speaker or for the addressee? *Cognitive Psychology* 2005;50:194–231. [PubMed: 15680144]
- Laures JS, Weismer G. The effects of flattened fundamental frequency on intelligibility at the sentence level. *Journal of Speech Language and Hearing Research* 1999;42:1148–1156.
- Nebes, RD. Hemispheric specialization in the aged brain. In: Trevarthen, C., editor. *Brain circuits and function of the mind: Essays in honor of R. W. Sperry*. New York: Cambridge University Press; 1990.
- Orbelo DM, Grim MA, Talbott RE, Ross ED. 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* 2005;18:25–32. [PubMed: 15681625]
- Pickering MJ, Branigan HP. The representation of verbs: Evidence from syntactic priming. *Journal of Memory and Language* 1998;39:633–651.
- Salthouse TA. The aging of working memory. *Neuropsychology* 1994;8:535–543.
- Salthouse TA. The processing-speed theory of adult age differences in cognition. *Psychological Review* 1996;103:403–428. [PubMed: 8759042]
- Schneider, BA.; Pichora-Fuller, MK. Implications of perceptual deterioration for cognitive aging research. In: Craik, FIM.; Salthouse, TA., editors. *Handbook of aging and cognition*. 2. Mahwah, NJ: Erlbaum; 2000. p. 155-220.
- Shattuck-Hufnagel S, Turk AE. A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research* 1996;25:193–247. [PubMed: 8667297]
- Snedeker J, Trueswell J. Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language* 2003;48:103–130.
- Speer SR, Kjelgaard MM, Dobroth KM. The influence of prosodic structure on the resolution of temporary syntactic closure ambiguities. *Journal of Psycholinguistic Research* 1996;25:247–268.
- Streeter LA. Acoustic determinants of phrase boundary perception. *Journal of the Acoustical Society of America* 1978;64:1582–1592. [PubMed: 739094]
- Van Lancker D, Sidtis JJ. The identification of affective-prosodic stimuli by left- and right-hemisphere-damaged subjects: All errors are not created equal. *Journal of Speech and Hearing Research* 1992;35:963–970. [PubMed: 1447930]
- Verhaeghen P. Aging and vocabulary score: A meta-analysis. *Psychology and Aging* 2003;18:332–339. [PubMed: 12825780]
- Wingfield A, Lahar CJ, Stine EAL. Age and decision strategies in running memory for speech: Effects of prosody and linguistic structure. *Journal of Gerontology: Psychological Sciences* 1989;44:106–113.

- Wingfield A, Lindfield KC, Goodglass H. Effects of age and hearing sensitivity on the use of prosodic information in spoken word recognition. *Journal of Speech, Language, and Hearing Research* 2000;43:915–925.
- Wingfield, A.; Stine-Morrow, EAL. Language and speech. In: Craik, FIM.; Salthouse, TA., editors. *Handbook of Aging and Cognition*. 2. Mahwah, NJ: Erlbaum; 2000. p. 359-416.
- Wingfield A, Wayland SC, Stine EAL. Adult age differences in the use of prosody for syntactic parsing and recall of spoken sentences. *Journal of Gerontology: Psychological Sciences* 1992;47:350–356.
- Zatorre RJ, Evans AC, Meyer E, Gjedde A. Lateralization of phonetic and pitch discrimination in speech processing. *Science* 1992;1:37–46.

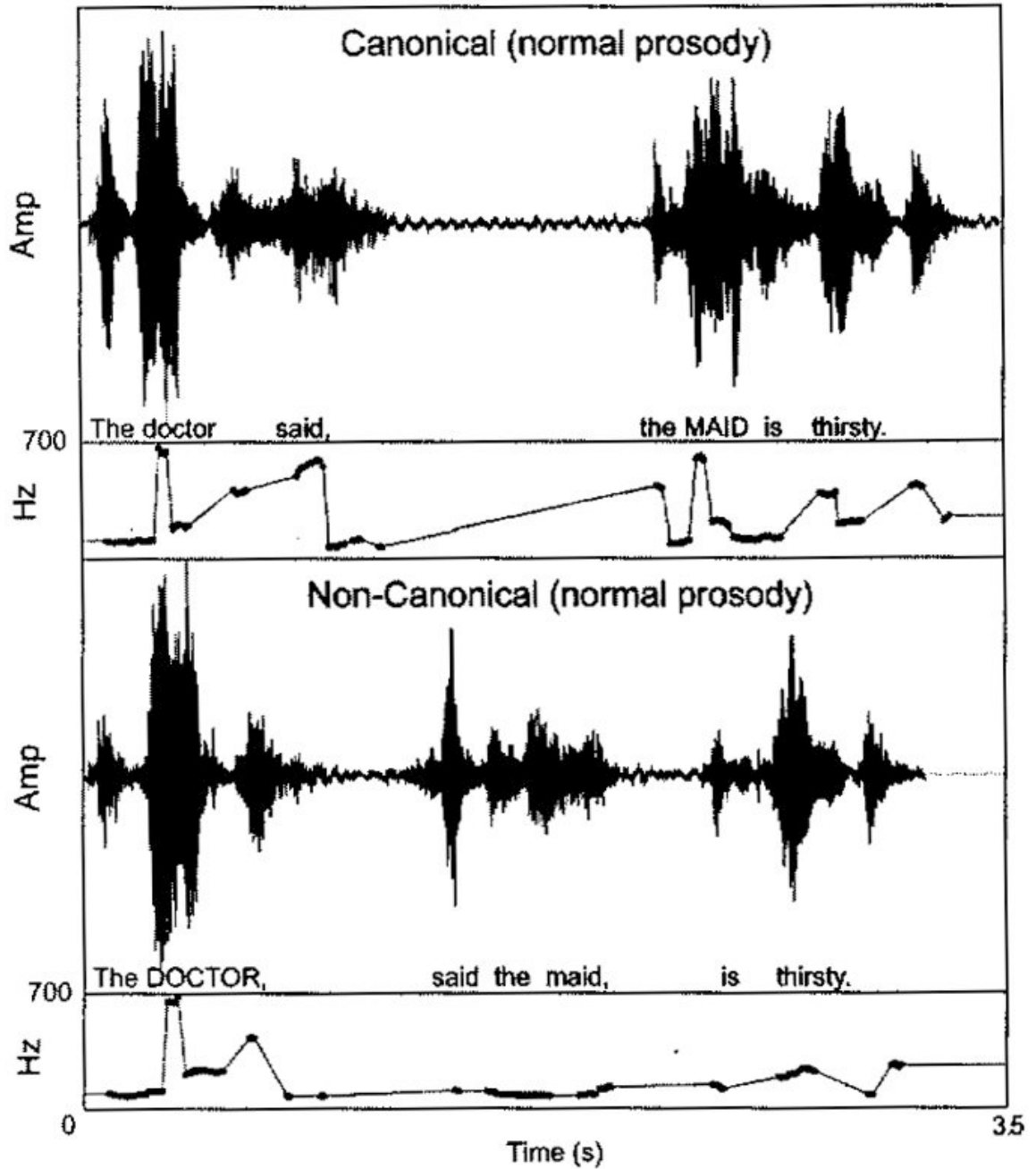


Figure 1. Waveforms and pitch contours for a sentence spoken with prosody to indicate a canonical (*top panel*) or noncanonical (*bottom panel*) interpretation.

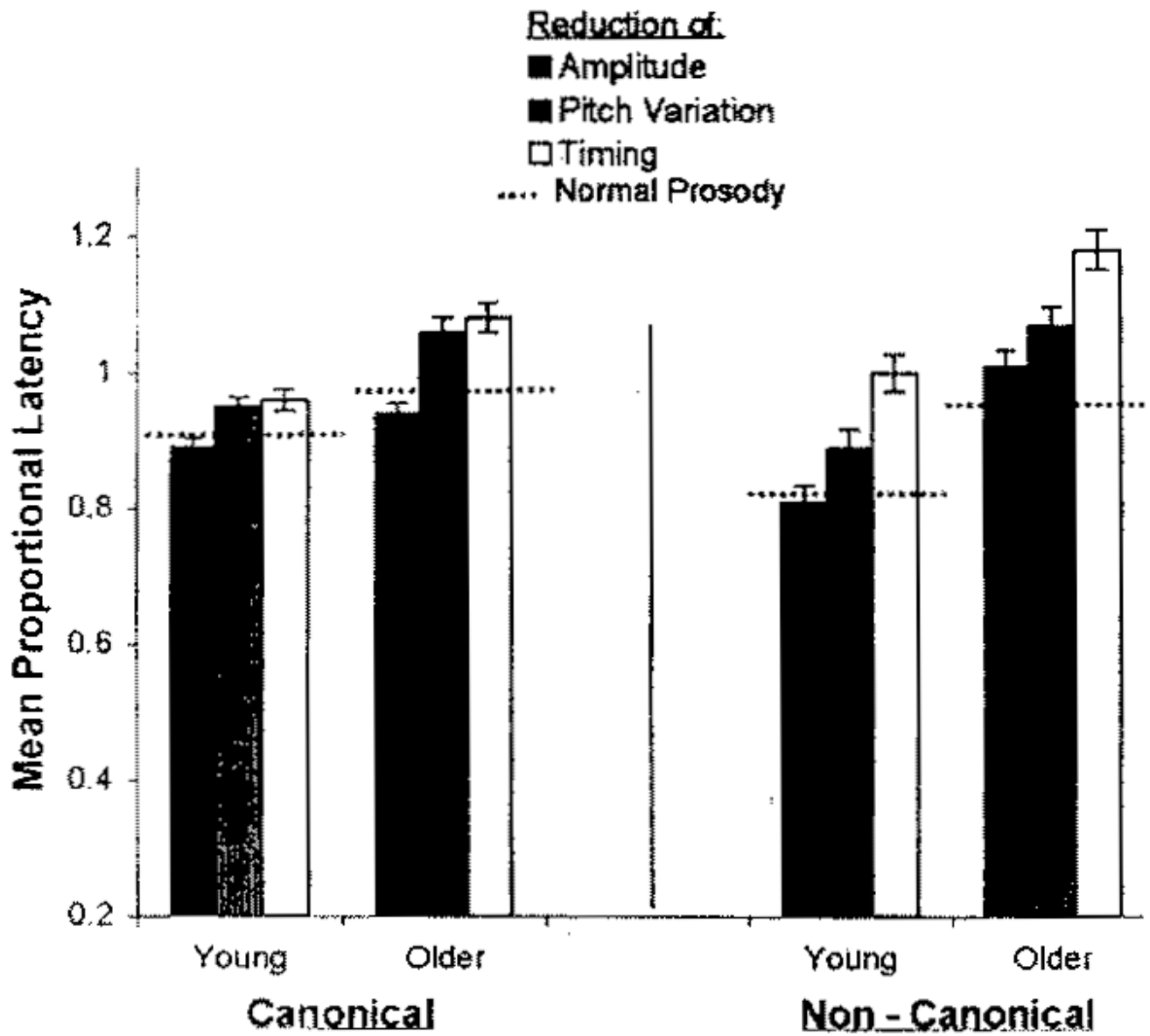


Figure 2. Mean proportional latency to identifying the described character in canonical (*left panel*) and noncanonical (*right panel*) sentences spoken with normal prosody (*horizontal dotted lines*) and when one prosodic feature was reduced, with the remaining features presented intact. Error bars represent one standard error.

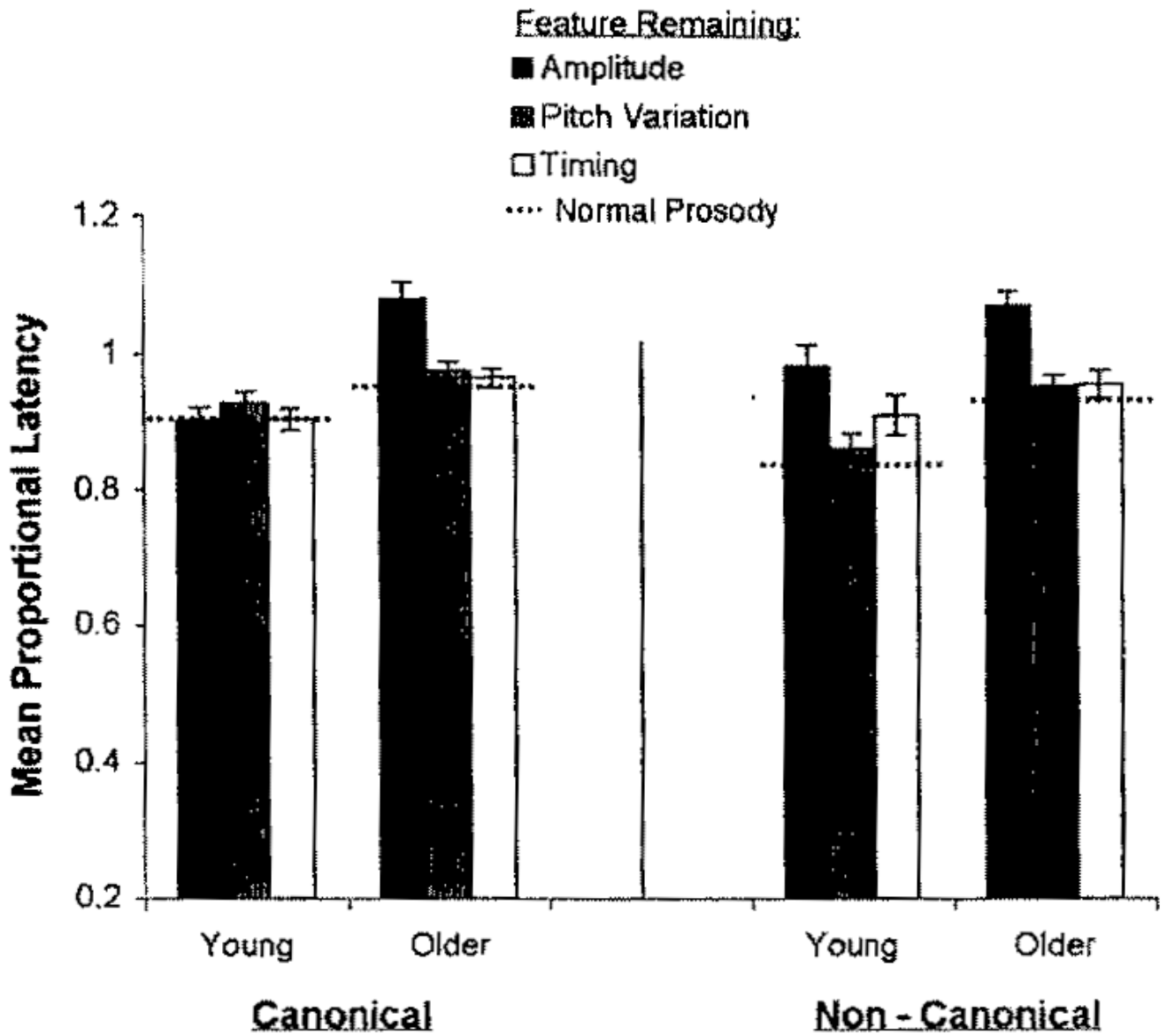


Figure 3. Mean proportional latency to identifying the described character in canonical (*left panel*) and noncanonical (*right panel*) sentences spoken with normal prosody (*horizontal dotted lines*) and when only one of three prosodic features was presented intact. Error bars represent one standard error.

Table 1

Percentage of correctly parsed sentences

Prosody condition	Canonical sentences		Noncanonical sentences	
	Young adults	Older adults	Young adults	Older adults
Original normal-prosody	98.5 (4.9)	97.5 (6.4)	99.0 (3.1)	94.5 (16.1)
Amplitude reduced	99.5 (2.2)	100.0 (–)	99.5 (2.2)	97.5 (5.5)
Pitch reduced	98.5 (3.7)	99.0 (4.5)	90.0 (1.3)	91.5 (14.6)
Timing reduced	99.5 (2.2)	99.0(3.1)	94.5 (7.6)	92.0 (12.0)

Note. Numbers in parentheses are standard deviations.

Table 2

Percentage of correctly parsed sentences

Prosody condition	Canonical sentences		Noncanonical sentences	
	Young adults	Older adults	Young adults	Older adults
Original normal-prosody	100.0 (–)	98.0 (4.1)	95.0 (8.9)	98.0 (7.0)
Amplitude remaining	100.0(–)	97.5 (5.5)	81.0(15.5)	85.0 (16.4)
Pitch remaining	100.0 (–)	99.0 (3.8)	96.0 (6.0)	95.0 (9.5)
Timing remaining	99.5 (2.2)	98.5 (3.7)	88.5(11.8)	87.0(15.6)

Note. Numbers in parentheses are standard deviations.