Title
Perception of Linguistic and Affective Prosody in Younger and Older Adults

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 28(28)

ISSN
1069-7977

Authors
Taler, Vanessa
Baum, Shari
Saumier, Daniel

Publication Date
2006

Peer reviewed

Perception of Linguistic and Affective Prosody in Younger and Older Adults

Vanessa Taler (vtaler@alcor.concordia.ca)
Bloomfield Centre for Research in Aging, McGill University
3999 Cote St. Catherine Road, Montreal, Quebec, Canada H3T 1E2

Shari Baum (shari.baum@mcgill.ca)
School of Communication Sciences & Disorders, McGill University
1266 Pine Avenue West, Montreal, Que., Canada H3G 1A8

Daniel Saumier (daniel.saumier@mail.mcgill.ca)
Bloomfield Centre for Research in Aging, McGill University
3999 Cote St. Catherine Road, Montreal, Quebec, Canada H3T 1E2

Abstract

Previous research has suggested that older adults display deficits in affective-prosodic processing, while grammatical-prosodic processing remains intact. In the present study, groups of younger and older adults took part in a series of experiments assessing their comprehension of prosodic information at the affective, grammatical and perceptual levels. It was found that older and younger adults performed similarly on perceptual tasks. However, deficits were seen in older adults across a number of tasks: affective-prosodic processing, the use of temporal information to parse syntactic structure, and the use of lexical stress to distinguish adjective-noun pairs from compound nouns. These findings suggest a generalized prosodic deficit in older adults which cannot be ascribed to a primary auditory deficit.

Introduction

Prosody plays an important role in spoken language, signaling both emotional and grammatical content. It is often only prosodic information that allows a listener to distinguish between different sentence modalities, such as whether a speaker is asking a question or making a statement. Likewise, good comprehension of prosodic information is vital in determining a speaker's emotional state. As such, accurate comprehension of prosodic information is essential in psychosocial interactions and relationship well-being (Carton, Kessler & Pape, 1999). Given the crucial role of prosodic information in everyday communicative situations, it is of interest to investigate how processing of this information may be altered in healthy aging.

While language processing is typically found to be unaffected in healthy older adults, a number of studies have suggested that deficits are seen in processing affective (i.e., emotional) prosody in these individuals (Brosgole & Weisman, 1995; Cohen & Brosgole, 1988; Kiss & Ennis, 2001; Ross, Orbelo, Testa & Beatty, 2000; Orbelo, Grim, Talbot & Ross, 2005). Processing of grammatical prosody in older adults, on the other hand, has been the object of less study. However, research to date suggests that such processing is relatively spared in healthy older adults in terms of syntactic parsing (Kjelgaard, Titone & Wingfield, 1999; Wingfield, Lahar & Stine, 1989), stress perception, and other features (Cohen & Faulkner, 1986; Wingfield, Lindfield & Goodglass, 2000; Wingfield, Wayland & Stine, 1992).

Prosodic information may be conveyed by means of three acoustic parameters: fundamental frequency (F0), duration and amplitude (Lehiste, 1970). In terms of speech production, most current models treat prosody either as its own module, separate from the rest of the speech production system, or as a subcomponent of the phonological system (Garrett, 1980; Levelt, 1989). A separate prosodic tier specifying metrical structure is postulated in recent phonological models (Levelt 1989, Liberman and Prince 1977, Selkirk 1984). As such, deficits in prosody may be expected to dissociate from other linguistic deficits.

There are thus two ways in which our cognitive system may organize prosodic information. It may be the case that affective and grammatical prosody constitute separate cognitive modules, and as such may be differentially impaired. On the other hand, a distinction between affective- and grammatical-prosodic processing may not be reflected in our cognitive system; rather, the use of different prosodic cues (i.e., F0, duration and amplitude) could subsume modular processing.

The present study addresses these issues by examining comprehension of prosodic information in older adults across a variety of domains. We used a battery of tasks designed to tap prosody processing at the perceptual, affective and grammatical levels. Comprehension of affective prosody was assessed in a task examining detection of emotional valence at the sentence level, both in the presence and in the absence of semantic information. Given the multiple roles of prosody in signaling grammatical information, we investigated use of grammatical-prosodic information at both the syntactic and lexical levels. First, we examined older adults' capacity to utilize prosody to determine sentence modality (interrogative, declarative or imperative). As in the affective-prosodic task, stimuli included sentences containing semantic information and sentences that did not contain such information. Second, we looked at older adults' use of prosodic information to assign syntactic structure in otherwise ambiguous sentences. Third, we investigated their use of lexical stress in word recognition. At the perceptual level, we examined older and younger adults' use of pitch and keyword and pause duration in two categorization tasks.

It was hypothesized that older adults would exhibit deficits in comprehension of affective prosody, but that their performance in the tasks assessing prosody processing at the perceptual and grammatical levels should parallel that of younger adults. Such a result would support the claim that grammatical and affective prosody are represented separately in our cognitive system.

Methods and Participants

Participants

Ten older adults (average age = 79.1 years ± 6.8; average education: 13.1 years ± 3.3) and 8 younger adults (average age = 24.6 years ± 4.4; average education: 15.3 years ± 1.3) took part in the study. All subjects were native speakers of English with no history of neurological or psychiatric illness. Older adults were recruited from the Memory Clinic of the Jewish General Hospital of Montreal, a tertiary referral centre, and underwent a complete neuropsychological battery to exclude dementia. Undergraduate students in psychology and linguistics were recruited to serve as a younger control group. In order to confirm that participants' hearing thresholds were adequate to perceive the stimuli, they were each required to repeat a series of five sentences played through computer speakers. All participants successfully completed this hearing screen.

Methods

The study comprised seven subtests, and took approximately two hours to complete. Testing was completed in a single session, unless the participant requested that testing be split into two sessions. In each subtest, stimuli were recorded by a native speaker of English, and were played to participants on a computer with an external speaker. Stimuli were repeated as many times as necessary upon participant request.

Perceptual categorization of phrases varying along the statement-question continuum. In order to assess the establishment of category boundaries between prosodic categories, a natural declarative statement ("He wants to leave now.") was acoustically manipulated by linearly increasing the fundamental frequency contour of the sentence's final word to a level that is equivalent to the F0 level of the final word of its question counterpart ("He wants to leave now?") This resulted in a series of eight sentences between the two endpoints in which the F0 of the final word differed by equal 11% steps. An identical procedure was applied to transform the natural question counterpart ("He wants to leave now?") into a statement. Participants heard each of the resultant 20 sentences twice and were asked to categorize each as either a statement or a question.

Sentence prosody: linguistic prosody processing. This task aimed to assess participants' ability to detect grammatical modality using prosodic cues, both in the presence and in the absence of semantic information. In this task, participants were required to identify intonation meaning from a set of sentences varying in linguistic modality. Sentences were either interrogative (e.g., "Has your daughter begun school yet?"), declarative (e.g., "The doctor examined the patient"), or imperative (e.g., "Run to the store for some bread and milk.").

The first set of sentences was in English (n=8 in each modality for a total of 24 sentences); semantic information was thus available in this condition. Each sentence was then low-pass filtered at 500Hz to remove all intelligible linguistic information, while conserving intonational variations across the utterances; thus, semantic information was not available in this set of stimuli. Finally, a set of sentences was recorded using nonsense words, where prosodic information indicated an interrogative, declarative or imperative intonation (n=8 of each). Again, these sentences conveyed no semantic information, meaning that judgments had to be made on the basis of prosodic information alone.

Nine sentence categories, each comprising 8 stimuli, were thus included in the experiment: statement- English (S-E), statement-nonsense words (S-NE), statement-low-pass filtered (S-F), command-English (C-E), command-nonsense words (C-NE), command-low-pass filtered (C-F), question-English (Q-E), question-nonsense words (Q-NE), and question-low-pass filtered (Q-F). The resultant 72 sentences were recorded by a female native speaker of English and presented in a pseudo-random order. Participants were asked to decide for each sentence whether they had heard a question, a statement or a command.

Sentence prosody: affective prosody processing. This task was identical to that described above, except that, rather than grammatical function, participants were required to identify affective tone. In this task, affective tone was either happy (H; e.g., "It was so nice to see you again"), sad (S; e.g., "We were not chosen for the team"), or angry (A; e.g., "She will never clean up her mess!"). Again, 72 stimuli in total were included in the task, 24 in English, the same 24 stimuli low-pass filtered at 500Hz, and 24 stimuli with happy, sad or angry intonation, but using nonsense words.

Lexical-stress perception. Lexical-stress perception was assessed using a task in which participants were required to differentiate between 12 compound words (e.g., 'greenhouse) and their matched noun phrases (e.g., green 'house) in a word-picture matching task. For each stimulus, participants were shown an array of three pictures representing the referent of the stimulus, its matched adjective-noun phrase or compound, and an unrelated foil. For example, participants heard the compound noun 'greenhouse, and were asked to select from among a picture of a greenhouse, a picture of a house painted green, and a picture of a chair. Both compound and adjective-noun stimuli were presented, for a total of 24 items.

Temporal cues marking phrasal boundaries I. This test aimed to assess participants' use of temporal information, specifically keyword and pause duration, in assigning phrasal boundaries. In this test, the phrase "pink and black and green" was manipulated by inserting a temporal boundary after "pink" or "black", and varying the pre-boundary word durations or pause durations at the boundary in a step-wise manner (Aasland & Baum, 2003; Baum, Pell, Leonard & Gordon, 1997). The duration of the keyword "pink" ranged from 286ms to 446ms in five 40ms steps; likewise, the duration of the keyword "black" ranged from 284ms to 448ms. That is, for the pink series, pause length was set to 286ms (step 1), 326ms (step 2), 366ms (step 3), 406ms (step 4), or 446ms (step 5); for the black series, pause length was set to 284ms (step 1), 324ms (step 2), 364ms (step 3), 404ms (step 4), or 444ms (step 5). For both series, boundary pauses ranged from 0ms to 160ms in five 40ms steps (i.e., 0ms, 40ms, 80ms, 120ms, 160ms). Subjects listened to each stimulus and reported whether they had heard [[pink and black] and green] or [pink and [black and green]], either verbally or by spatially manipulating three coloured paper squares to place the pink and black or the black and green squares together.

Temporal cues marking phrasal boundaries II. Stimuli in this task comprised 14 pairs of sentences, each of which differed only in the presence of a pause changing the phrase structure of the sentence (e.g., "Madam, Flower is the name of my cat" vs. "Madam Flower is the name of my cat"). Participants heard each sentence and were then asked comprehension questions in order to verify how the sentence was parsed. Two types of ambiguous sentences were presented. A set of seven sentences were designed in which either the first two or final two words were names. These could either constitute a compound name signaling the agent or theme of the sentence (e.g., "I am going to see Billy Jean") or two separate names signaling the agent/theme of the sentence and the addressee (e.g., "I am going to see Billy, Jean"). A second set of seven sentences was constructed which could be interpreted as containing either a parenthetical (e.g., "Tom, my boyfriend, is out of town) or an addressee (e.g., "Tom, my boyfriend is out of town"; n=7). Sentences were presented in pseudo-random order.

Results

Perceptual categorization of phrases varying along the statement-question continuum. Results are shown in Figure 1; percentage of statement responses is depicted on the y-axis, and step is depicted on the x-axis, with "1" being the most statement-like exemplar and "10" the most question-like. A repeated-measures ANOVA with Origin (statement vs. question) and Step (1-10) as within-subject variables and Group (older vs. younger) as a between-subject variable revealed significant main effects of Origin ($F(1,1)=4.48$, $p<0.05$) and Step ($F(1,10)=151.796$, $p<0.001$). No main effect of Group was found, suggesting that younger and older adults perform this task in a similar fashion. However, an Origin X Step X Group interaction was seen ($F(1,9) = 3.227$, $p<0.01$), reflecting the fact that older adults exhibit more variability when responding to statements which have been altered to sound like questions.

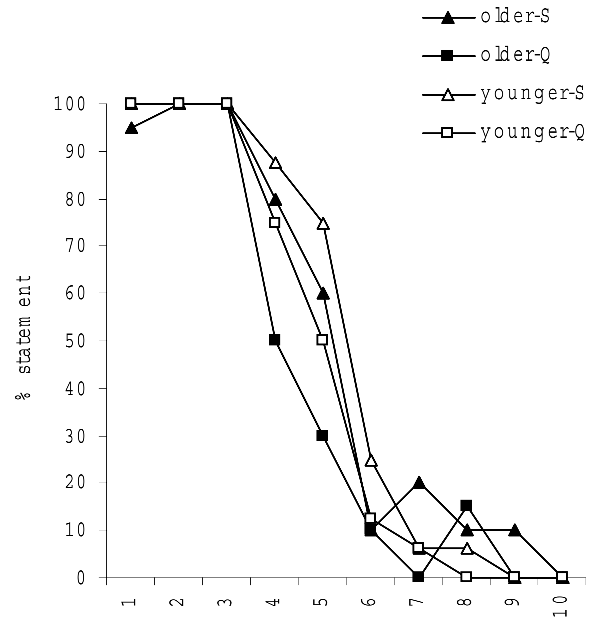

Figure 1: Results, perceptual categorization

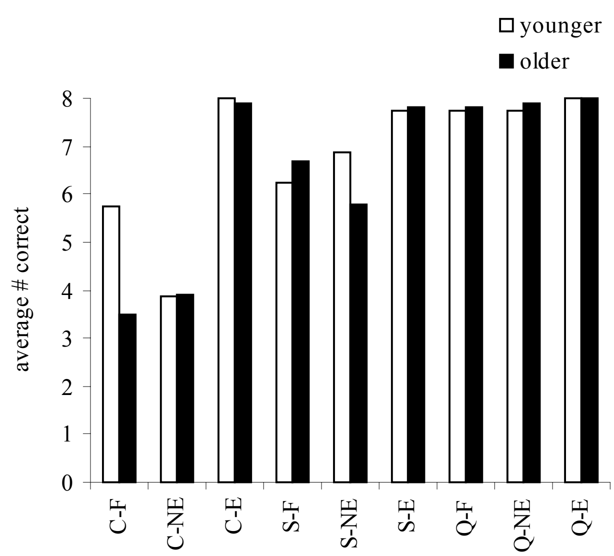

Figure 2: Results, linguistic prosody processing

Sentence prosody: linguistic prosody processing. Results of this task are shown in Figure 2 below. Again, older and younger adults performed similarly on almost all stimulus categories (chi-square, $p>0.05$ for C-NE, C-E, S-F, S-NE, S-E, Q-F, Q-NE, Q-E). Older adults did, however, manifest impaired performance in the C-F condition. That is, in the low-pass filtered condition, they were significantly more likely to misidentify a command as a statement (chi square = 11.43, $p<0.001$).

Sentence prosody: affective prosody processing. In this task, the younger and older adults' performance did not differ in the E (English) condition, where semantic information was

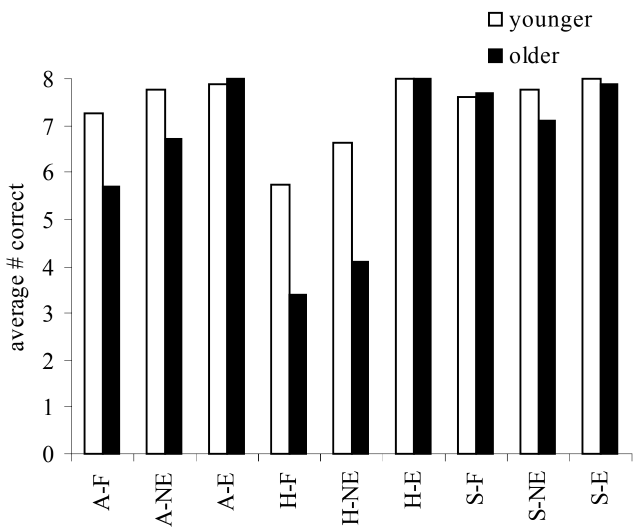

Figure 3: Results, sentence prosody: affective prosody processing

available. However, significantly lower performance was seen in the nonsense (NE) condition, both for happy (H; chi-square=15.6, $p<0.001$) and angry (A; chi-square = 6.56, $p<0.05$) stimuli. Likewise, older adults were significantly less likely to identify happy and angry sentences in the low-pass filtered (F) condition (H: chi-square = 12.43, $p<0.001$; A: chi-square = 8.30, $p<0.01$).

Lexical-stress perception. Older adults' performance on this task differed significantly from that of younger adults overall (chi-square = 8.98, $p<0.01$). Separate analyses of performance on adjective-noun stimuli (e.g., green 'house) and compound nouns ('greenhouse) revealed that, when hearing a compound noun, older and younger adults were equally likely to select the picture depicting the correct referent (chi-square = 0.66, $p>0.05$). However, older adults were significantly more likely to select the picture corresponding to the compound reading when hearing an adjective-noun combination (chi-square = 11.58, $p<0.001$). Results are presented in Figure 4.

Temporal cues marking phrasal boundaries I. A repeated-measures ANOVA was conducted on the data, with series (black vs. pink), pause duration (0ms, 40ms, 80ms, 120ms, 160ms) and keyword duration (five durations between 286ms and 446ms for the keyword "pink"; five durations between 284ms and 444ms for the keyword "black") as within-subject variables and group as a between-subject variable. This analysis revealed a main effect of series ($F(1,1) = 1124.672$, $p<0.001$) as well as interactions between series and keyword duration ($F(1,4) = 3.581$, $p<0.01$) and between series and pause duration ($F(1,4) = 2.687$, $p<0.04$). A three-way interaction between series, keyword and pause was also observed ($F(1,16) = 2.231$, $p<0.005$). The main effects and interactions indicate that stimulus has an effect. No interactions with group and any other variable were seen, indicating that older and younger adults interpret durational cues in a similar fashion. Results are presented by keyword duration in Figure 5a, and by pause duration in Figure 5b below. The percentage of responses in which participants selected the reading [pink and [black and green]] is depicted on the y-axis; duration (from shortest to longest) is depicted on the x-axis. In Figure 5a, "B-series" refers to the duration of the keyword "black", and "P-series" refers to the duration of the keyword "pink". In Figure 5b, "B-series" refers to the duration of the pause following the keyword "black", and "P-series" refers to the duration of the pause following the keyword "pink".

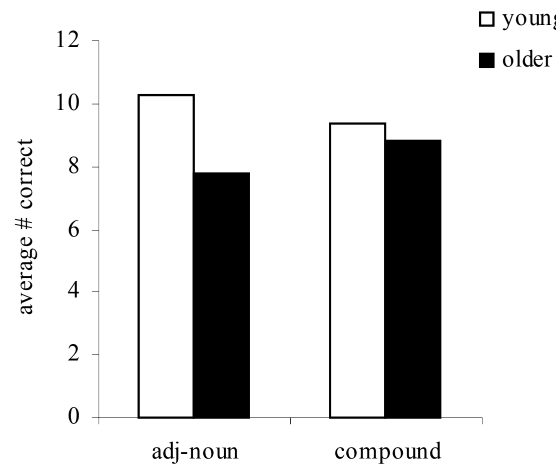

Figure 4: Results, lexical stress perception

Temporal cues marking phrasal boundaries II. This task assessed participants' use of temporal cues to determine the syntactic structure of a sentence. Results are shown in Figure 6. Older adults performed significantly worse on this task than did younger adults, interpreting an average of 22.5 of 28 sentences correctly, versus younger adults' 25.9 (chi-square, $p<0.001$). This difference was significant both for the appositive condition ($p<0.01$) and for the compound condition ($p<0.025$).

Discussion

Overall, differences were seen in older and younger adults' processing of prosodic information across a number of tasks. As predicted, older adults' performance on the task assessing affective-prosodic processing revealed a deficit in the capacity of these individuals to determine sentence modality on the basis of prosodic information. While they exhibited no difficulty in determining sentence modality when semantic information was available (i.e., when the sentences were in English), their performance declined sharply when semantic information was unavailable, either due to its removal via low-pass filtering of the original English sentence, or when the stimulus was recorded using nonsense stimuli. This result is consistent with the finding reported in the literature that older adults exhibit deficits in affective-prosodic processing (Brosgole & Weisman, 1995; Cohen & Brosgole, 1988; Kiss & Ennis, 2001; Ross et al., 2000).

With respect to the tasks assessing grammatical-prosodic processing, variability was observed across tasks. In the first task, in which participants used prosodic information to determine sentence modality, older adults manifested a deficit in only one condition: low-pass filtered commands, which they were significantly more likely to misclassify as statements. However, no deficit was seen in distinguishing statements from questions; furthermore, younger and older adults exhibited equal difficulty in identifying commands in the NE (nonsense stimuli) condition. Thus, we postulate that younger adults' superior performance in the low-pass filtered condition may in fact be due not to a difference in prosody processing itself, but rather to their ability to extract some linguistic (i.e., semantic) information from the low-pass filtered sentences.

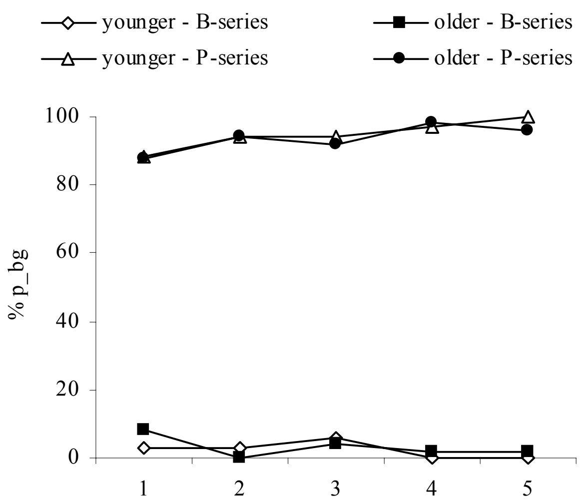

Figure 5a: Results, phrasal boundaries: effect of keyword duration

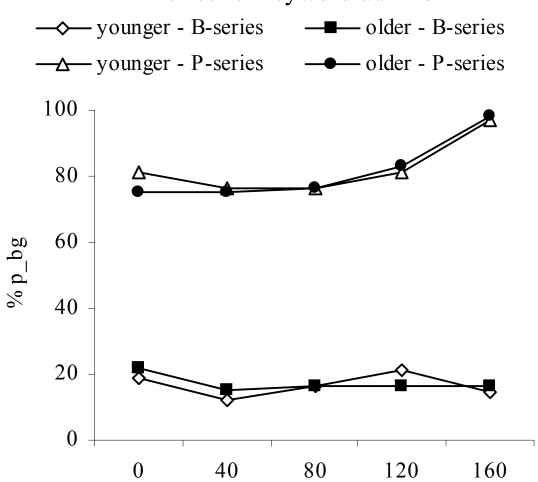

Figure 5b: Results, phrasal boundaries: effect of pause duration

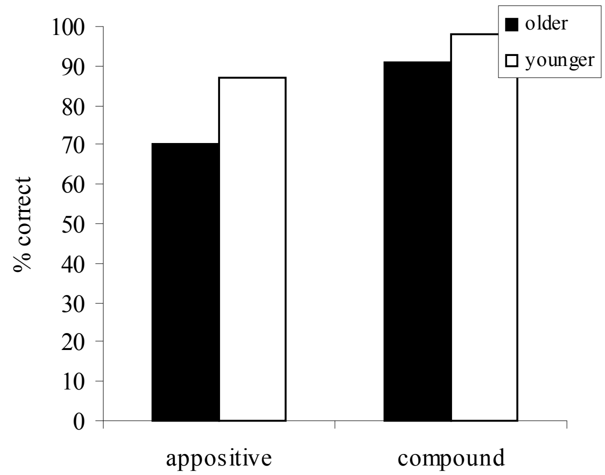

Figure 6: Results, phrasal structure: syntactic parsing

In the second task, in which participants were required to use lexical stress to distinguish adjective-noun pairs from matched compound nouns in a word-picture matching task, older adults exhibited an impairment relative to the younger adults. Specifically, no difference was seen across the two groups in their classification of compound nouns, but older adults were more likely to incorrectly select the picture depicting the referent of the compound noun when hearing an adjective-noun combination (35% error rate in the older adult group, versus a 14.6% error rate in the younger adults. For example, when hearing "light 'house", they were more likely than younger adults to select a picture of a lighthouse rather than a picture of a house floating in the air. One possible account for this finding is that older adults' responses were driven by referent plausibility, rather than purely by phonological information. For example, when older adults hear the stimulus "light 'house", they are more likely to select a picture with greater real-world plausibility (a lighthouse) rather than the implausible entity (a floating house), even if the prosodic information indicates that this is the correct referent.

Finally, in the third grammatical task, in which participants heard sentences whose syntactic structure was disambiguated by prosodic information, older adults exhibited significantly lower performance than younger adults. This was seen both in the condition in which participants were required to use prosodic information to distinguish appositive constructions from addressee + SVO structures, and in the condition in which prosodic information signaled a compound name or agent/theme + addressee.

The possibility that this result may be due to differences between the two groups in perceptual prosodic processing is belied by the finding that similar performances were seen on the two tasks assessing processing at the perceptual level. Minimal differences were seen in the two participant groups' use of sentence-final F0 to categorize sentences as statements vs. questions; only in statements altered to sound like questions did older adults display slightly more variability than younger adults. In terms of the use of keyword and pause duration to determine stimulus grouping, no difference was seen between the two groups. This suggests that any differences seen in the remaining tasks are not the result of a primary auditory-perceptual deficit.

In contrast to previous findings, the present results point to a impairment in processing of grammatical-prosodic information in healthy older adults. This deficit does not appear to affect the ability to detect sentence modality, which remains relatively spared. However, surprisingly, the use of prosody to parse syntactic structure appears affected, as does the use of lexical stress to distinguish between minimal pairs of adjective-noun combinations and noun-noun compounds. It should be noted, however, that in performing the latter task older adults may have relied on semantic or pragmatic information, such as judgments of real-world plausibility, more heavily than did younger adults. This reliance on semantic/pragmatic information may possibly reflect a compensatory strategy on the part of older adults to compensate for prosodic processing difficulties. Further research is clearly required to address this possibility; moreover, more robust perceptual testing may be required to fully discount the possible role of a perceptual impairment in the processing of grammatical-prosodic information.

We thus argue, on the basis of a series of well-controlled tasks assessing various aspects of prosodic processing, that older and younger adults differ in their processing of prosodic information at both the affective and the grammatical levels. This is in contrast to previous studies which have suggested that the prosodic impairment seen in older adults is specific to affective processing. Rather, our results suggest that older adults may exhibit a generalized deficit in comprehension of prosodic information. The present findings do not support the hypothesis that affective and grammatical prosody constitute separate cognitive modules. We suggest that prosody may best be viewed as a centralized module, possibly subdivided by acoustic cue rather than function. This is consistent with recent neuroimaging research (Gandour, Tong, Wong, Talavage et al., 2004) suggesting that processing of prosodic information is not localizable to a single neural region, but rather solicits a large network of cortical regions distributed across the two cerebral hemispheres.

Acknowledgments

The current research was supported by a grant from the Centre for Research on Language, Mind and Brain (CRLMB) to Shari Baum, Howard Chertkow and Daniel Saumier. We wish to thank the medical staff and support staff of the JGH/McGill Memory clinic for their assistance, particularly Dr. Howard Chertkow, Shelley Solomon, Renée Kaminski, and Chris Hosein.

References

Aasland, W.A., & Baum, S.R. (2003) Temporal parameters as cues to phrasal boundaries: a comparison of processing by left- and right-hemisphere brain-damaged individuals. *Brain and Language, 87,* 385-99.

Baum, S.R., & Pell, M.D. (1999). The neural bases of prosody: Insights from lesion studies and neuroimaging. *Aphasiology, 13,* 581-608.

Baum, S.R., Pell, M.D., Leonard, C.L., & Gordon, J.K. (1997). The ability of right- and left-hemisphere-damaged individuals to produce and interpret prosodic cues marking phrasal boundaries. *Language and Speech, 40,* 313-30.

Brosgole, L., & Weisman, J. (1995). Mood recognition across the ages*. International Journal of Neuroscience, 82,* 169–189.

Carton, J.S., Kessler, E.A., & Pape, C.L. (1999). Nonverbal decoding skills and relationship well-being in adults. *Journal of Nonverbal Behavior, 23,* 91-100.

Cohen, E., & Brosgole, L. (1988) Visual and auditory affect recognition in senile and normal elderly persons. *International Journal of Neuroscience, 43,* 89–101.

Cohen, E., & Faulkner, D. (1986). Does “elderspeak” work? The effect of intonation and stress on comprehension and recall of spoken discourse in old age. *Language and Communication, 6,* 91-98.

Gandour, J., Tong, Y., Wong, D., Talavage, T., Dzemidzic, M., Xu, Y., Li, X., & Lowe, M. (2004). Hemispheric roles in the perception of speech prosody. *Neuroimage, 23,* 344-357.

Garrett, M. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.) *Language Production.* (London: Academic Press).

Kiss, I. & Ennis, T. (2001). Age-related decline in perception of prosodic affect. *Applied Neuropsychology, 8,* 251-254.

Kjelgaard, M.M., Titone, D.A., & Wingfield, A. (1999). The influence of prosodic structure on the interpretation of temporary syntactic ambiguity by young and elderly listeners. *Experimental Aging Research, 25,* 187-207.

Lehiste, I. (1970). *Suprasegmentals.* (Cambridge, MA: MIT Press).

Levelt, W. (1989). *Speaking: From Intention to Articulation* (Cambridge, MA: MIT Press).

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,* 25–32.

Ross, E.D., Orbelo, D.M., Testa, J., & Beatty, W. (2000) Age-related changes in processing affective prosody. *Neurology, 54(suppl 3),* A418–A419.

Wingfield, A., Lahar, C.J., & Stine, E.A.L. (1989). Age and decision strategies in running memory for speech. *Journal of Gerontology: Psychological Sciences, 44,* 106-113.

Wingfield, A., Wayland, S.C., & Stine, E.A.L. (1992). Adult age differences in the use of prosody for syntactic parsing and recall of spoken sentences. *Journal of Gerontology: Psychological Sciences, 47,* 350-356.