JSLHR

Research Article

Age-Related Differences in the Perception of Emotion in Spoken Language: The Relative Roles of Prosody and Semantics

Boaz M. Ben-David,[a,b,c,d] Sarah Gal-Rosenblum,[a] Pascal H. H. M. van Lieshout,[b,c,d] and Vered Shakuf[a]

Purpose: We aim to identify the possible sources for age-related differences in the perception of emotion in speech, focusing on the distinct roles of semantics (words) and prosody (tone of speech) and their interaction.
Method: We implement the Test for Rating of Emotions in Speech (Ben-David, Multani, Shakuf, Rudzicz, & van Lieshout, 2016). Forty older and 40 younger adults were presented with spoken sentences made of different combinations of 5 emotional categories (anger, fear, happiness, sadness, and neutral) presented in the prosody and semantics. In separate tasks, listeners were asked to attend to the sentence as a whole, integrating both speech channels, or to focus on 1 channel only (prosody/semantics). Their task was to rate how much they agree the sentence is conveying a predefined emotion.
Results: (a) Identification of emotions: both age groups identified presented emotions. (b) Failure of selective attention: both age groups were unable to selectively attend to 1 channel when instructed, with slightly larger failures for older adults. (c) Integration of channels: younger adults showed a bias toward prosody, whereas older adults showed a slight bias toward semantics.
Conclusions: Three possible sources are suggested for age-related differences: (a) underestimation of the emotional content of speech, (b) slightly larger failures to selectively attend to 1 channel, and (c) different weights assigned to the 2 speech channels.

Participation in social situations is essential in healthy aging (Schneider, Pichora-Fuller, & Daneman, 2010). Communication difficulties often prompt withdrawal or avoidance of social situations, severely limiting the range of activities available for older adults (Heinrich et al., 2016; Smith & Kampfe, 1997). This can ultimately lead to a less satisfying lifestyle and depression (Polku et al., 2015), eventually affecting longevity (Holt-Lunstad, Smith, & Layton, 2010) and increasing the risk of dementia (Bennett et al., 2006). A key factor in effective communication and an emerging topic in aging research is deciphering emotions in speech (Paulmann, Pell, & Kotz, 2008). The evidence in the literature suggest that older adults have difficulties in identifying emotions in speech, which may, in turn, negatively affect social interactions (Carton, Kessler, & Pape, 1999; Mitchell & Kingston, 2011). However, evidence on the different processes underlying this age-related effect is inconsistent and lacking. The goal of the current study is to identify the possible sources for differences in the perception of emotion in speech between older and younger adults. Specifically, we test for age-related differences on (a) the identification of emotion in the prosodic (tone of speech) and semantic (words) auditory channels; (b) the extent of failures of selective attention to one auditory channel, while ignoring the other; and (c) the integration of the two auditory channels.

[a]Communication Aging and Neuropsychology Lab, Baruch Ivcher School of Psychology, Interdisciplinary Center, Herzliya, Israel
[b]Department of Speech-Language Pathology, University of Toronto, Ontario, Canada
[c]Rehabilitation Sciences Institute, University of Toronto, Ontario, Canada
[d]Toronto Rehabilitation Institute, University Health Networks, Ontario, Canada
Correspondence to Vered Shakuf: vshakuf@idc.ac.il

Editor-in-Chief: Frederick (Erick) Gallun
Editor: David Eddins
Received April 30, 2018
Revision received August 6, 2018
Accepted September 18, 2018
https://doi.org/10.1044/2018_JSLHR-H-ASCC7-18-0166
Publisher Note: This article is part of the Special Issue: Select Papers From the 7th Aging and Speech Communication Conference.

Disclosure: *The authors have declared that no competing interests existed at the time of publication.*

To illustrate the issue at hand, imagine a grandfather receiving a phone call from his young granddaughter saying, "I feel wonderful today," spoken with angry prosody. What will he think? Is she expressing happiness, anger, or a combination of the two? Phrasing the above question more generally, do younger and older adults differ in the way they perceive and identify verbal and prosodic emotional cues spoken by younger speakers? Do they assign the same weights to the two speech channels? This is not only a theoretically important question but also a practical one, given the important role of social interaction to well-being in aging (Hertzog, Kramer, Wilson, & Lindenberger, 2008).

Perception of Emotion in Speech: Semantics and Prosody

Perceiving spoken emotions in social interactions involves combining information from various sources. In the absence of visual cues (e.g., when talking over the phone) or when visual information is degraded due to aging (Ben-David & Schneider, 2010), the ability to derive emotional meaning in spoken language relies on how it is conveyed in two auditory speech channels: *semantics*, that is, the lexical content of the words, and *prosody*, (e.g., rhythm, stress, and intonation [suprasegmental features]). Perception of emotion in spoken language also requires an integration of the two speech channels, or an inhibition of one of them, as recently suggested by Ben-David, Multani, Shakuf, Rudzicz, and van Lieshout (2016; see also Paulmann et al., 2008).

To test the full complexity of the interplay between the two channels in the perception of emotion in speech, Ben-David et al. (2016) developed the Test for Rating of Emotions in Speech (T-RES), as presented in Figure 1. In the T-RES, participants are presented with spoken sentences in which the emotional semantic and prosodic content appear in different combinations from trial to trial. Listeners are asked to rate the extent to which each sentence conveys a predefined emotion. After testing 80 healthy young adults, Ben-David et al.'s (2016) findings spoke to three factors: (a) identification of emotions: When asked to identify the emotions expressed in the prosody and semantics separately, younger adults easily processed the emotional content; (b) selective attention: It is difficult for younger adults to perceive one channel without the influence of the other—this was noted by direct measures of selective attention, instructing listeners to ignore one channel (see also Ishii, Reyes, & Kitayama, 2003; Kitayama & Ishii, 2002; Mehrabian & Wiener, 1967); and (c) integration of channels: Listeners process the emotional content of spoken language as a whole, affected by the emotions conveyed by both prosodic and semantic channels. This was demonstrated by supremacy of congruency: A sentence that presents the same emotion in both channels was perceived by the younger adults as expressing a larger extent of that emotion than sentences that present mismatched emotions (see also Beaucousin et al., 2006; Mitchell, 2006; Nygaard & Queen, 2008; Paulmann, Jessen, & Kotz, 2009; Wurm, Vakoch, Strasser, Calin-Jageman, & Ross, 2001; for a discussion on the psychophysical properties of redundancy, see Ben-David & Algom, 2009). Most importantly, the two emotional channels do not play the same role in emotional processing. For the younger adults in the Ben-David et al. (2016) study, a prosodic dominance was documented: The prosody appeared to have a larger impact on emotional ratings than semantics (see also Jacob, Brück, Plewnia, & Wildgruber, 2014; Mehrabian & Wiener, 1967). These three factors have been examined in the literature as possible sources for age-related differences in the perception of emotion in speech.

Difficulties in the Perception of Emotion in Speech in Older Listeners

Age-Related Effects: Identification of Semantic and Prosodic Emotional Cues

The literature mainly focuses on the roles of each speech channel separately. Regarding semantics, an age-related decrease in the accuracy for categorizing emotions in textual stimuli was noted (Grunwald et al., 1999). However, with short stories, emotion identification was not affected by age (Phillips, MacLean, & Allen, 2002). The well-documented difficulties older adults experience in identifying spoken words, especially in noise (for a review, see Schneider et al., 2010), can also lead to difficulties in identifying the emotional content of the message.

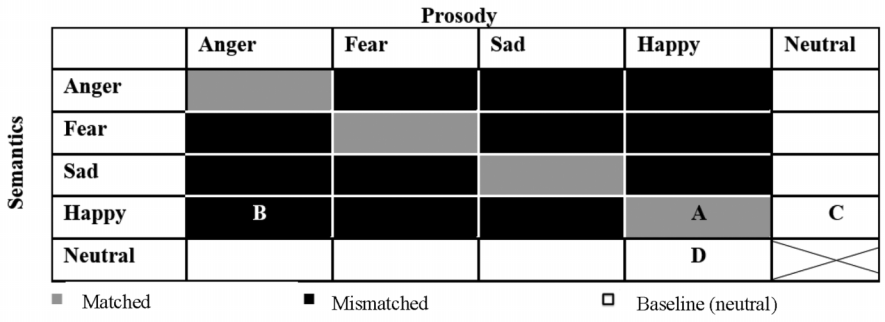
Regarding prosody, studies generally report an age-related decrease in accuracy (e.g., Brosgole & Weisman, 1995; Dupuis & Pichora-Fuller, 2015; Orbelo, Grim, Talbott, & Ross, 2005; Sen, Isaacowitz, & Schirmer, 2017), starting as early as the age of 40 years (Paulmann et al., 2008). These differences were found to persist even after controlling for hearing sensitivity (Mitchell, 2007; Schmidt, Janse, & Scharenborg, 2016), suggesting that they may be "primary in origin" (Mitchell, 2007, p. 1435). Indeed, age-related declines in the auditory system can impair older adults' ability to detect acoustic features that serve as prosodic cues (Dmitrieva & Gelman, 2012), such as duration (Gordon-Salant, Yeni-Komshian, Fitzgibbons, & Barrett, 2006) and neural temporal synchrony (Pichora-Fuller, Schneider, MacDonald, Pass, & Brown, 2007).

Age-Related Effects: Failures of Selective Attention

One cognitive decline that has been found to contribute to age differences in performance across a wide range of sensory modalities is the decrease in the efficiency of inhibitory processes (Hasher & Zacks, 1988; Healey, Campbell, & Hasher, 2008). A large body of the literature on aging (e.g., McDowd & Shaw, 2000; Troyer, Leach, & Strauss, 2006) suggests that the ability to inhibit the processing of irrelevant information decreases, often observed as a reduction in selective attention. In speech perception, inhibition deficits were documented in several studies (e.g., Jerger et al., 1993; Sommers & Danielson, 1999). With respect to the perception of emotion in speech, if inhibition is limited, older adults may be less able than

Figure 1. Test for Rating of Emotions in Speech.

General design of the stimuli



A. Matched	B. Mismatched	C. Baseline (semantics)	D. Baseline (prosody)
Semantically happy (e.g., "I got a raise in my salary") with matched happy prosody.	Semantically happy ("I got first place") with angry prosody.	Semantically happy ("I won the lottery") with neutral prosody.	Semantically neutral ("the pot has a blue lid") with happy prosody.

Rating tasks

General rating task	Prosodic rating task	Semantic rating task
Participants are asked to rate the overall emotion of the sentence as a whole.	Participants are asked to rate the sentence based only on the prosody ignoring the semantics.	Participants are asked to rate the sentence based only on semantics, ignoring the prosody.
Set 1 with 15 sentences	Set 2 with 15 sentences, presented in both tasks.	
This task measures the relative weighting of the semantic and prosodic channels.	These tasks (resembling Stroop-like tasks) gauge age-related differences in selective-attention. Baseline (neutral) sentences measure identification of emotions.	

Rating blocks

Each rating task is made up of 4 emotional rating blocks, comprising 12 experimental blocks. On each trial, listeners are asked "From 1 (strongly disagree) to 6 (strongly agree)…

Anger-rating	Sadness-rating	Fear-rating	Happiness-rating
"…to what extent do you agree that the speaker conveys **anger**?"	"…to what extent do you agree that the speaker conveys **sadness**?"	"…to what extent do you agree that the speaker conveys **fear**?"	"…to what extent do you agree that the speaker conveys **happiness**?"

younger adults to process the emotion conveyed in one channel (e.g., semantic) separately from the other (prosodic).

Age-Related Effects: Integration of Channels

Age-related differences in the perception of emotion in speech have also been attributed to a biased integration of the two speech channels, giving one speech channel a more dominant role than intended by the speaker (Ben-David et al., 2016). For example, Dupuis and Pichora-Fuller (2010) examined emotional judgments of spoken sentences, where the emotional prosody and semantics were either matched or mismatched. In mismatched trials, they found that younger adults' judgments were predominately prosody-based, whereas older adults were, in general, less consistent in their responses (see also Paulmann et al., 2008). Indirect support for age-related differences in the integration of emotional cues comes from a recent study testing the perception of emotion in facial expressions and body posture composites (Abo Foul, Eitan, & Aviezer, 2018). Younger adults were found to give more weight to the facial expressions, whereas older adults gave more weight to the body postures. The authors view these results as supporting a social expertise hypothesis (Hess, 2006), where lifelong social experience leads older adults to utilize different

strategies to improve their recognition of real-life social situations.

The Current Study

The current study aims to test the three factors that may lead to age-related differences in performance on spoken-emotions identification tasks. We compare performance between older and younger adults on the T-RES, using a Hebrew-adapted version (Shakuf, Gal-Rosenblum, & Ben-David, 2016). In the test, participants are asked to rate the extent to which an emotion is expressed by the prosody alone (prosodic rating task), the semantics alone (semantic rating task), or the sentence as a whole (general rating task).

1. Age-related effects: identification of semantics and prosodic emotional cues. Based on the literature presented above, we hypothesize that age-related differences in accuracy for identification of emotions in prosody and semantics are partially due to difficulties in inhibiting the information conveyed by the irrelevant channel. Thus, when no competing information is presented by the irrelevant channel, we can assume that age-related differences in identification will be minimized. One of the measures in the current study (prosodic and semantic rating of baseline sentences) is dedicated to test this. Listeners are asked to rate the emotions presented in one channel (e.g., prosody), whereas the other (semantics) conveys an emotionally neutral content. We hypothesize that age-related differences in this task would be significant yet minimal.
2. Age-related effects: failures of selective attention. The literature points to a decrease in the ability to inhibit irrelevant information in aging. Thus, we expect that, when competing information is conveyed by the irrelevant channel, we will find larger failures of selective attention in older adults. This is tested directly, when listeners are asked to rate the emotions presented in one channel (e.g., prosody), whereas the other (semantics) conveys a different emotion (prosodic and semantic rating tasks of mismatched sentences).
3. Age-related effects: integration of channels. First, the data in the literature suggest that listeners typically process the spoken message as a whole, affected by the information conveyed in both channels. This is tested directly by comparing ratings for sentences that present the same emotion in both channels with sentences that present mismatched emotions (supremacy of congruency in the general rating task). We expect to find that both older and younger adults will show an advantage for emotional-matched sentence (see Ben-David, Eidels, & Donkin, 2014). Second, we hypothesize that age-related differences in the perception of emotion in speech are not merely based on misidentification of either one of the channels but are engendered by assigning different weights to the two channels. This is tested directly in the prosodic dominance measure—an estimate of age-related differences in the relative weights listeners assign to prosodic and semantic channels when rating emotions in speech. We expect to find large age-related differences, where older adults are less biased toward the prosodic channel than younger adults are (Dupuis & Pichora-Fuller, 2010; Paulmann et al., 2008).

As a final step, we will discuss both theoretical and practical implications of our findings.

Method

Participants

Forty young adults, undergraduates from Interdisciplinary Center Herzliya (20 women and 20 men; M_{age} = 25 years, SD = 1.5 years), and 40 older adults from the Herzliya community (20 women and 20 men; M_{age} = 70 years, SD = 2.0 years) participated in this study. All participants were native Hebrew speakers as assessed by a self-report questionnaire. This was further confirmed by the vocabulary subset of the Hebrew version of Wechsler Adult Intelligence Scale–Fourth Edition (Wechsler, Coalson, & Raiford, 2008), with all participants scoring within the clinically normal range, with no significant differences between groups (averages of 61.1 and 62.5 for younger and older adults, respectively; $t < 1$). We only included participants with (a) good ocular health and no auditory or language problems, as assessed by a self-report questionnaire; (b) no indication of clinical depression as assessed by self-report scales (older: Geriatric Depression Scale [Zalsman, Aizenberg, Sigler, Nahshoni, & Weizman, 1998]; younger: Depression Anxiety and Stress Scale [Lovibond & Lovibond, 1995]); (c) pure-tone air-conduction thresholds within clinically normal limits for their age group, for the following frequencies: 250, 500, 1000, and 2000 Hz (younger: ≤ 15 dB; older: ≤ 35 dB; in the better ear); and (d) cognitive performance for older adults within clinically normal ranges, as assessed by the Montreal Cognitive Assessment (score > 26; Nasreddine et al., 2005).

The study received ethics approval from the Interdisciplinary Center Ethics Committee, and all participants signed an informed consent. Older adults were compensated by the equivalent of $10 per hour for their participation, and younger adults participated in the study for partial course credit.

Stimuli and Apparatus: T-RES

We used the Hebrew version of the original T-RES (Shakuf et al., 2016), with the following four emotional categories being tested: anger, sadness, fear, and happiness (Zupan, Neumann, Babbage, & Willer, 2009). These are generally deemed equally easily recognized and distinguished in both prosody and semantics and are widely used in the literature (Brodsky, Shnoor, & Be'er 2003; Laukka, 2003; but see Scherer, Banse, & Wallbott, 2001). A neutral category is also included as a baseline for performance. A full description of the tool and the procedures are presented

in Ben-David et al. (2016). For the Hebrew version, a set of 50 written semantic sentences in Hebrew (Shakuf & Ben-David, 2014), 10 in each of the five emotion categories, were validated as being distinctive in conveying their corresponding emotions. Sentences were equated on main linguistic characteristics (average frequency of usage in Hebrew of content words, average number of syllables per sentence) across the five affective categories (following the procedures detailed in Ben-David, van Lieshout, & Leszcz, 2011). They were recorded by a native Hebrew Israeli professional radio-drama actress, using the five different prosodies. A final set of 50 spoken sentences was selected as best exemplars of their respective prosodic categories by a group of trained raters (following the procedures discussed in Ben-David, Thayapararajah, & van Lieshout, 2013). Digital audio files were equated with respect to their root-mean-square amplitude and duration.

The final experimental set was made of two subsets of 25 sentences, where each semantic category was represented once in each of the tested prosodies, generating a 5 (semantics) × 5 (prosody) matrix, as shown in the top panel of Figure 1. Closely following the original paradigm (Ben-David et al., 2016), spoken sentences that carry neutral information in both semantics and prosody were deemed as uninformative for this study and removed from the analysis, leaving a final set of 48 spoken sentences, 24 in Subset A and 24 in Subset B. Each subset consisted of emotional trials that present emotions on both channels (marked as black and gray filled cells in Figure 1) and baseline trials that present an emotion only on one channel and a neutral content on the other (marked as white filled cells in Figure 1).

Note that T-RES stimuli are recorded by a trained professional actress, rather than using natural emotional expressions. Despite the artificial nature of this condition, it provides better control over the variability in the acoustics of the recorded material and yields a more intense and prototypical expression of the specific emotion. This controlled condition adheres to a standard procedure reported in the literature (Mitchell, 2006) and may be especially relevant for testing seniors, given the high variability in their performance. Also, to minimize variability across stimuli and conditions (see Roche, Peters, & Dale, 2015), one speaker recorded all stimuli.

Procedure

Upon arrival, all participants received a short explanation regarding the experimental task and signed an informed consent form. Participants were tested individually, seated in a sound-attenuated booth. Instructions were presented on a 17-in. flat color monitor. Spoken sentences were presented via MAC-51 audiometer headphones. To partially mitigate the possible impact of hearing differences, stimuli were presented 40 dB above individual audiometric thresholds (pure-tone average) in quiet (see Schneider, Daneman, & Murphy, 2005).

An experimental session consisted of three rating tasks: semantic rating, prosodic rating, and general rating, separated by short breaks. In the semantic rating task, listeners were asked to rate exclusively the emotion conveyed by the semantic content ignoring the prosody of the spoken sentence (using sentences from Subset B). In the prosodic rating task, listeners were asked to attend exclusively to the prosody, ignoring the semantic content of the utterance (using sentences from Subset B). In the general rating task, listeners were asked to rate the emotion expressed by the speaker based on the spoken sentence as a whole (using sentences from Subset A). The three tasks are presented in Figure 1.

Each rating task was made of four emotion-rating blocks: anger-, fear-, sadness-, and happiness-rating blocks, comprising (3 rating tasks × 4 emotion-rating blocks =) 12 experimental blocks per session. In each block, participants were asked to listen to a spoken sentence and rate how much they agree that the speaker conveys a predefined emotion (based on the prosody, semantics, or sentence as a whole), using a 6-point Likert scale, ranging from *strongly disagree* (1) to *strongly agree* (6). For example, "How much do you agree that the speaker is_____?" (angry/fearful/sad/happy). The participants responded by pressing the respective number key (1–6) on a standard PC keyboard. Each of the 12 blocks consisted of 24 trials, making for 288 trials per session (for a total duration of less than 35 min).

An emotion-rating block began with the presentation of an instruction slide, followed by two practice trials. Practice trials were followed by a reminder of the instructions, and experimental trials were initiated by the participant. Each trial began with the presentation of the audio file, followed by the specific instructions presented on the monitor. No feedback was provided throughout the practice and experimental trials, as the T-RES gauges the listeners' subjective perception of emotion (i.e., There are no "right" or "wrong" answers).

To control for a possible effect of presentation order, three methods were utilized. First, each experimental session started with the general rating task to prevent biasing the listeners to pay attention to a specific channel. For a randomly selected half of the participants, this was followed by the semantic rating task and then the prosodic rating task. For the other half, this order was reversed. Second, in each task, the order of the four emotion-rating blocks was counterbalanced across participants using a Latin-square procedure (Grant, 1948) but maintained across all three rating tasks for an individual participant. The latter two counterbalancing procedures required the use of eight experimental groups (five participants per group, in each age group). For full details, see Table 2 in Ben-David et al. (2016). Finally, the order of the trials in each block was fully randomized.

Statistical Analyses

All of the following analyses used mixed-model repeated-measures analyses of variance (ANOVAs;

generalized linear model) with average ratings as the dependent variable, age group (2: younger vs. older adults) as a between-participant variable, and target emotion (4: anger, sadness, fear, or happiness) as a within-participant variable. Each test included one other within-participant variable, as specified in Appendix A.

1. Emotion-identification factor. To gauge age-related effects of identification of semantic and prosodic emotional cues, baseline sentences in the semantic and prosodic rating tasks were used. The ANOVA (Equation 1 in Appendix A) compared the difference between ratings of sentences that present the target emotion in the target channel with sentences that do not present the target emotion in either channel.
2. Selective attention factor. Age-related differences in failures of selective attention were tested using emotional sentences in the semantic and prosodic rating tasks. Specifically, the ANOVA (Equation 2 in Appendix A) compared the difference between ratings of sentences that present the target emotion only in the to-be-ignored channel with sentences that do not present the target emotion in either channel.
3. Integration of channels (trial-type factor). Age-related differences in integration of channels were tested comparing ratings of the following four types of trials in the general rating task (Equation 3 in Appendix A): (a) target-emotion–matched trials, where the target emotion is presented both in the prosody and in the semantics; (b) target-emotion–prosodic trials, where the target emotion is present only in the prosody; (c) target-emotion–semantics trials, where the target emotion is present only in the semantics; and (d) target-emotion–absent trials, where the target emotion is not present in either prosody or semantics. For example, when asked to rate happiness, that is, the target emotion is happiness: (a) target-emotion–matched trials can be the semantically happy sentence, "I got a raise in my salary," spoken with happy prosody; (b) target-emotion–prosodic trials may be the semantically sad sentence, "My best friend is leaving the country," spoken with happy prosody; (c) target-emotion–semantics trials can be the semantically happy sentence, "I got first place," spoken with angry prosody; and (d) target-emotion–absent trials may be the semantically angry sentence, "He cut in front of me," spoken with sad prosody.

Factors of order (4: order of emotional rating blocks, 2: whether the semantic rating task was introduced before or after the prosodic rating task) were included in all ANOVAs. As they did not yield any significant interaction with the tested effects, they will not be further discussed. The analyses of the effects of discrete emotions are presented in Appendix B, showing that the main trends were replicated in the separate analyses of each emotion. Level of significance across all analyses was set at .05. Partial eta squared (η_p^2) was used as the measure for effect size.

Results

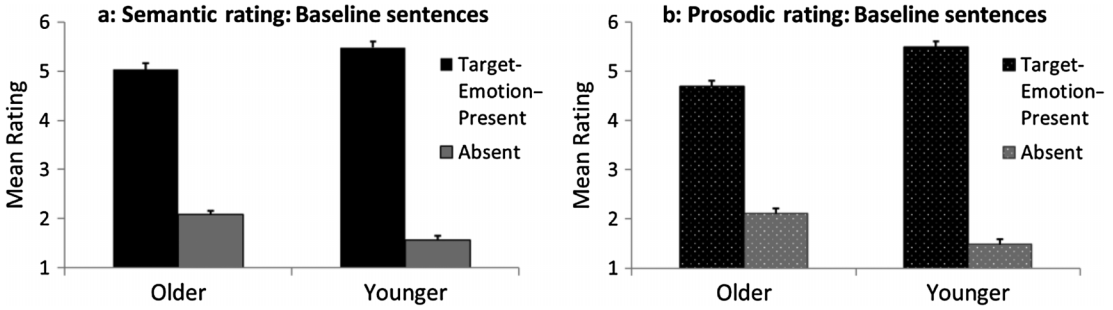
Identification of Emotions

In the prosodic and semantic rating tasks, we tested for possible age-related differences in the identification of emotions in the prosodic and semantic channels, respectively, using baseline sentences. The mixed-model repeated-measures ANOVA consisted of age group (2: older vs. younger) as a between-participant variable; target emotion (4: anger, fear, sad, and happy), target channel (2: semantics vs. prosody), and emotion identification (2: sentences that present the target emotion only in the target channel vs. sentences that do not) as within-participant variables; and order factors (2: task order × 4: block order) as covariates. For example, this analysis involves comparing prosodic-anger ratings of a semantically neutral sentence spoken with angry prosody (M = 5.7/6) with the average prosodic-anger ratings of all other semantically neutral sentences, spoken with sad, happy, and fearful prosody (M = 1.8/6).

Analysis, across emotions, shows a main effect for emotion identification (M = 3.4), $F(1, 74) = 1499.3$, $p < .001$, $\eta_p^2 = .95$, with no significant main effect for the target channel [comparing prosodic with semantic ratings, $F(1, 74) = 1.4$, $p = .24$]. Most importantly, emotion identification interacted significantly with age group, $F(1, 74) = 43.6$, $p < .001$, $\eta_p^2 = .37$, with the emotion-identification factor significantly larger for the younger than the older adults (3.95 vs. 2.8, respectively). A significant triple interaction of emotion identification, age group, and target channel was also noted, $F(1, 74) = 4.1$, $p = .047$, $\eta_p^2 = .05$, with a very small effect size. In follow-up analyses, for each group separately, emotion identification was found to be highly significant, with similarly large effect sizes, $\eta_p^2 > .90$ [younger adults: $F(1, 35) = 1175.5$, $p < .001$, $\eta_p^2 = .97$; older adults: $F(1, 35) = 427.8$, $p < .001$, $\eta_p^2 = .92$]. The interaction of emotion identification with target channel was only significant for the older adult group [younger: $F(1, 35) = 0.406$, $p = .5$, $\eta_p^2 = .01$; older: $F(1, 35) = 4.0$, $p = .05$, $\eta_p^2 = .10$], with a larger emotion-identification factor for semantic rating.

A graphic description of these comparisons between target-emotion–present and target-emotion–absent baseline trials is presented in Figure 2, separately for semantic rating (Panel A) and prosodic-rating (Panel B) tasks. In summary, analysis indicates that both age groups clearly identified the presented emotions in both prosody and semantics, as indicated by the extremely high effect sizes. However, younger adults gave sentences that present the target emotion in the target channel higher ratings than older adults did and gave sentences that do not present the target emotion in either channel lower ratings than older adults did. Moreover, older adults' ratings appear to be slightly higher when emotions appeared in the semantic than in the prosodic channel. The general trends were confirmed for each separate emotion, as detailed in Appendix B.

Figure 2. Graphic description of ratings for older and younger adults (left and right pair of bars, respectively) in the semantic rating (a) and prosodic rating (b; dotted bars) tasks for target-emotion–present (dark bars) and target-emotion–absent (gray bars) baseline trials. The large difference between the latter two demonstrates identification of emotions. Error bars represent standard errors.



Failures of Selective Attention

An analysis of prosodic and semantic ratings indicated the expected age-related difference in failure to inhibit the irrelevant channel (Hasher & Zacks, 1988). The mixed-model repeated-measures ANOVA consisted of age group (2: older vs. younger) as a between-participant variable; target emotion (4: anger, fear, sad, and happy), target channel (2: semantics vs. prosody), and selective attention (2: sentences that present the target emotion only in the to-be-ignored channel vs. sentences that do not present the target emotion in either channel) as within-participant variables; and order factors (2: task order × 4: block order) as covariates.

For example, this analysis involves comparing average prosodic-anger ratings of semantically angry sentences spoken with sad/fear/happy/neutral prosody (M = 2.6/6) with the average prosodic-anger ratings of sad/fear/happy/neutral semantics spoken with sad/fear/happy/neutral prosody (M = 2.0/6). In both cases, anger prosody is not present; therefore, prosodic-anger ratings should be similar and very low. If listeners can selectively attend to one channel, the difference between the two types of sentences (selective-attention factor) will be zero. Yet, if listeners cannot ignore the semantics-anger content, significant differences ensue. The extent of this difference gauges the extent of failure to inhibit the irrelevant channel.

Results across both groups and emotions show a significant main effect for failures of selective attention, $F(1, 74) = 87.8$, $p < .001$, $\eta_p^2 = .54$. An interaction of selective-attention factor with age group, $F(1, 74) = 5.0$, $p = .029$, $\eta_p^2 = .06$, suggests larger failures of selective attention in the older group (note the low effect size). The lack of significant interaction of selective attention and task type, $F(1, 74) = 0.16$, $p = .69$, suggests that failures of selective attention did not significantly differ between the two tasks, across age groups. No significant triple interaction with age group was found either, $F(1, 74) = 2.2$, $p = .14$. Follow-up analyses confirm that failures of selective attention were significant in both age groups [younger adults: $F(1, 35) = 44.9$, $p < .001$, $\eta_p^2 = .57$; older adults: $F(1, 35) = 54.3$, $p < .001$, $\eta_p^2 = .61$].

A graphic description of failures of selective attention in semantic and prosodic rating tasks for older and younger adults is presented in Figure 3. In summary, as can be seen in Figure 3, failures of selective attention are evident in both age groups, for both tasks, but to a slightly greater extent in older adults (means of 0.43 vs. 0.255). Significant failures of selective attention in both age groups were also found in dedicated analyses for each emotion, as presented in Appendix B.

Integration of Channels

Figure 4 presents a graphic depiction of ratings in the general rating task, averaged across the four emotion rating blocks, separately for older and younger adults for target-emotion–matched trials (the target emotion appears in both channels), target-emotion–prosody trials (the target emotion appears only in the prosody), target-emotion–semantics trials (the target emotion appears only in the

Figure 3. Graphic description of the extent of failures of selective attention for older and younger adults (left and right pair of bars, respectively) in the semantic rating (empty bars) and prosodic rating (dotted bars) tasks. Each bar represents the difference between emotional trials that present the target emotion in the to-be-ignored channel only and emotional trials that do not present it in either channel. Error bars represent standard errors.

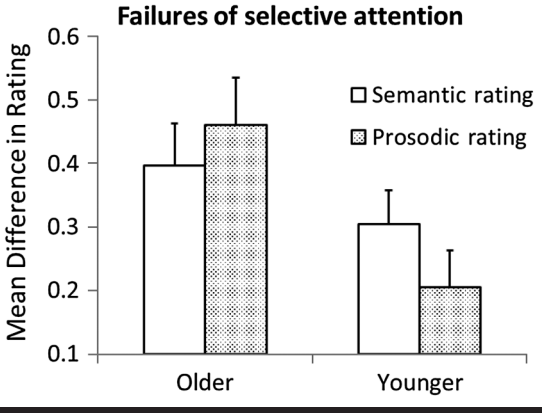
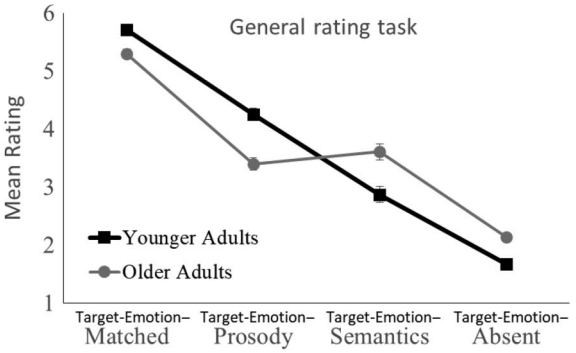


Figure 4. Graphic description of ratings in the general rating task for older (gray line) and younger (black line) adults. Ratings are averaged for target-emotion–matched trials (the target emotion appears in both channels), target-emotion–prosody trials (the target emotion appears only in the prosody), target-emotion–semantics trials (the target emotion appears only in the semantics), and target-emotion–absent trials (the target emotion does not appear in either the semantics or the prosody). Error bars represent standard errors.



semantics), and target-emotion–absent trials (the target emotion does not appear in either the semantics or the prosody). Figure 4 suggests that average performance across groups mimics the linear trend observed for younger adults in the original study (Ben-David et al., 2016). Target-emotion–matched trials received the highest emotional ratings, followed by target-emotion–prosody trials, target-emotion–semantics trials, and, finally, target-emotion–absent trials. The most notable feature of the graph is the apparent age-related interaction: Younger adults rated target-emotion–prosody sentences higher than target-emotion–semantics sentences (means of 4.25/6 vs. 2.87/6, respectively; replicating Ben-David et al.'s [2016] results), whereas this trend appears to be reversed to some degree for older adults (means of 3.39/6 vs. 3.61/6, respectively).

These observations were supported by the mixed-model repeated-measures ANOVA that consisted of age group (2: older vs. younger) as a between-participant variable; target emotion (4: anger, fear, sad, and happy) and the linear trend of trial type (1: target-emotion–matched > target-emotion–prosody > target-emotion–semantics > target-emotion–absent trials) as within-participant variables; and order factors (2: task order × 4: block order) as covariates. The ANOVA indicated a significant linear trend across age groups of target-emotion–matched > target-emotion–prosody > target-emotion–semantics > target-emotion–absent trials, $F(1, 74) = 1923.3$, $p < .001$, $\eta_p^2 = .96$, that interacted significantly with age group, $F(3, 74) = 61.7$, $p < .001$, $\eta_p^2 = .41$, with no main effect for age group, $F(1, 74) = 0.18$, $p = .7$.

To clarify this interaction, a follow-up analysis tested the possible effects of age group on the three effects identified in the original T-RES study with younger adults (Ben-David et al., 2016): (a) In supremacy of congruency, the extensive advantage for target-emotion–matched trials over target-emotion–prosody trials was significant across age groups, $F(1, 74) = 407.5$, $p < .001$, $\eta_p^2 = .85$, and did not significantly interact with age group, $F(1, 74) = 5.2$, $p = .06$, $\eta_p^2 = .05$. (b) Semantics appears to play a similar role for both groups. Target-emotion–semantics trials were rated significantly higher than target-emotion–absent trials, $F(1, 74) = 209.7$, $p < .001$, $\eta_p^2 = .74$, to a similar extent in both groups [no significant main effect for age group, $F(1, 74) = 0.43$, $p = .5$]. (c) However, findings indicate that the effect for prosodic dominance (the advantage of target-emotion–prosody over target-emotion–semantics trials) was not significant when tested across groups, $F(1, 74) = 3.0$, $p = .088$ $\eta_p^2 = .04$, but rather interacted significantly with age group, $F(1, 74) = 19.9$, $p < .001$, $\eta_p^2 = .21$, with no main effect for age group, $F(1, 74) = 0.78$, $p = .38$. Separate tests show that prosodic dominance was significant for younger adults (4.0 vs. 3.1), $F(1, 35) = 17.8$, $p < .001$, $\eta_p^2 = .34$. For older adults, the reverse effects were indicated, with slightly higher ratings for target-emotion–semantics over target-emotion–prosody trials (3.2 vs. 3.7), $F(1, 35) = 4.4$, $p = .04$, $\eta_p^2 = .11$. These general trends were replicated in dedicated analyses for each emotional rating block, as presented in Appendix B.

To further validate this interaction, a separate analysis was conducted for baseline sentences, comparing semantically neutral sentences that carry the target emotion in the prosody with prosodically neutral sentences that carry the target emotion in the semantics. Here, even when emotion integration (or, possibly, inhibition) was minimized, prosodic dominance significantly interacted with age group, $F(1, 74) = 42.2$, $p < .001$, $\eta_p^2 = .36$, with stronger dominance found for younger adults.

Discussion

The current study explored age-related differences in the perception of emotion in spoken language. Our findings, comparing 40 younger and 40 older native Hebrew speakers, highlight the following main trends: (a) Identification of emotions, both older and younger adults were able to identify the emotions presented in one channel, when no competing information was presented in the other. However, small but significant age-related differences were found, suggesting that younger adults rated target-emotion–present sentences higher than older adults. (b) Failure of selective attention, ratings of both age groups indicated failures of selective attention (in both channels), with slightly larger failures found for older than for younger adults. (c) Integration of channels, specifically, channel dominance, when rating emotions in speech, younger adults showed a bias toward prosody, whereas older adults weighted prosody and semantics more equally, with a slight bias for semantics. As discussed in the next passages, our study findings suggest three possible sources for age-related differences in the perception of emotional speech: (a) underestimation of the emotional content of speech, in both prosodic and semantic channels; (b) slightly larger failures to selectively attend to one channel and ignore the other; and (c) different weights assigned to the two speech channels.

Identification of Emotions

Several studies in the literature attribute age-related differences in the perception of emotional speech mainly to misidentification of emotional cues conveyed by one of the two speech channels (prosody: Schmidt et al., 2016; semantics: Grunwald et al., 1999; but see Phillips et al., 2002). Our data show that older listeners were able to identify emotions when asked to focus on one channel (prosodic and semantic rating tasks), when no competing emotional information was presented in the other, to-be-ignored, channel. The same trends were observed in both prosody and semantics and across emotional categories. In other words, when no inhibition was called for, older adults were successful at identifying emotions in speech. However, the extent of ratings differed between age groups. Younger adults gave higher ratings to target-emotion–present sentences than older adults did (means of 5.5/6 vs. 4.9/6, respectively), $F(1, 74) = 21.0$, $p < .001$, $\eta_p^2 = .22$, and lower ratings to target-emotion–absent sentences than older adults did (means of 1.5/6 vs. 2.1/6, respectively), $F(1, 74) = 25.1$, $p < .001$, $\eta_p^2 = .25$. In other words, older adults may underestimate the emotion presented in speech and overestimate the existence of the emotion when it is not present. These findings highlight the strength of using a rating method over a forced-choice response. With a rating method, we obtained a more detailed view of the full subjective perception of emotional prosody and semantics, rather than merely a classification of emotions. Indeed, findings point to an age-related difference in the extent of perception of spoken emotions.

Failures of Selective Attention

Both groups' performance indicated failures of selective attention. Failures were evident both when listeners were instructed to focus on the prosody and when instructed to focus on the semantics. This replicates findings from the original T-RES study (Ben-David et al., 2016), suggesting that the two speech channels are perceived more holistically and are difficult to pull apart, even when the task calls for it. Following Garner's conceptualization of the attributes of multidimensional stimuli (Garner & Felfoldy, 1970), the perception of emotional prosody and semantics can be considered as integral dimensions.

As expected, larger failures were documented for older than for younger adults. A decrease in the efficiency of inhibitory processes is well documented in the literature on aging, across a wide range of domains (Ben-David et al., 2014; Hasher & Zacks, 1988; Healey et al., 2008; but see Ben-David & Schneider, 2009, 2010). The current finding is consistent with a recent study where age-related differences in speech perception were uncovered only when inhibition demands were increased (Ben-David, Tse, & Schneider, 2012).

With respect to the perception of emotion in speech, when inhibition is limited, older adults have more difficulties than younger adults to process the emotion conveyed in one channel (e.g., semantic) separately from the other (prosodic). As a result, in real-life situations, older adults may inappropriately attempt to integrate the emotional information provided in both channels, and this can lead to misperceiving emotional content. Consider again our example, a grandfather receiving a phone call from his young granddaughter saying, "I feel wonderful today," spoken with angry prosody. Social convention may lead one to focus on the prosodic channel. However, inhibitory deficits may reduce the grandfather's ability to suppress the inconsistent (semantic) information, rendering the utterance happier than intended.

Integration of Channels: Channel Dominance

The most notable finding of the current study is the significant age-related difference in the integration of the two speech channels. For younger adults, emotional ratings appear to be impacted mainly by the (emotional content of the) prosodic channel, with a smaller contribution of the semantics. This effect extends our previous findings on English, an Indo-European language (Ben-David et al., 2016), to Hebrew, a Semitic language. Older adults are less biased in their weighting of channels, with a slight advantage to the semantic channel. Consider once more our example ("I feel wonderful today" spoken with angry prosody). Whereas the grandchild may perceive it as conveying mostly anger with only a modicum of happiness, the grandparent will perceive it as expressing both anger and happiness or even conveying slightly more happiness. This may lead the older listener to misinterpret the true intention of the younger speaker. Our findings augment previous evidence in the literature that shows that older adults are less focused on the prosodic emotional content than younger adults are (Dupuis & Pichora-Fuller, 2010; Paulmann et al., 2008).

Finally, a clear advantage in ratings for matched semantic–prosody combinations was indicated for both older and younger adults. This finding adds to the existing literature that demonstrates more efficient processing of matched combinations as well as higher valence ratings (Beaucousin et al., 2006; Ben-David et al., 2014; Dupuis & Pichora-Fuller, 2015; Mitchell, 2006; Nygaard & Queen, 2008).

Sensory and Cognitive Aging

Two key factors have been discussed as possible determinants of communication difficulties in older age (Carton et al., 1999): sensory and cognitive. Both factors can be used to partially explain the results of the current study.

Sensory Aging

Age-related changes in the auditory system (peripheral and central) degrade the quality of information transmitted to the central nervous system, thereby impeding speech understanding (e.g., Humes, 1996, 2007; Humes, Lee, &

Coughlin, 2006). A growing body of research directly associates sensory decline with age-related decline in cognitive and linguistic performance (Lindenberger & Baltes, 1994; Lindenberger & Ghisletta, 2009). This information degradation theory (Schneider & Pichora-Fuller, 2000) may, in part, explain age-related differences in emotional speech perception. Specifically, age-related sensory (auditory) degradation can impair the ability to identify emotion in both speech channels: (a) semantics, that is, age-related auditory dysfunction can reduce the ability to understand an emotion-related word (e.g., misperceiving "I feel mad" for "I feel sad"; for similar effects of offset overlap confusion, see Hadar, Skrzypek, Wingfield, & Ben-David, 2016), and (b) prosody, that is, failing to perceive (or misperceiving) emotional prosodic cues, due to reduced sensory processing (e.g., Dmitrieva & Gelman, 2012), especially under degraded listening conditions (Schneider & Pichora-Fuller, 2000; Schimdt et al., 2016).

Notably, in the current study, listening conditions were ideal, as speech was presented with no masking background, in a sound-attenuated booth, and the amplitude was adjusted to the participants' pure-tone audiometric thresholds. This somewhat minimizes the potential effects of age-related sensory degradation (see Ben-David et al., 2012).

Cognitive Aging

Several studies suggest that differences in inhibition and speech processing can be minimized when sensory differences are controlled for (see reviews in Monge & Madden, 2016, and Schneider et al., 2010). For example, Ben-David et al. (2012) gauged spoken word identification as the word unfolds in time, based on eye movements. In ideal listening conditions, no age-related differences in the moment-to-moment processing of spoken words were noted, nor differences in the processing of to-be-inhibited alternatives. However, the current data demonstrate that age-related differences in identification of spoken emotions (semantics and prosody) and in inhibition persisted, even after speech levels were tailored for ideal listening conditions. This may suggest that sensory sources are not fully mitigated by merely adjusting sound levels. Alternatively, age-related differences found in the current study may not be fully attributed to sensory changes. Future studies should test adjusted signal-to-noise ratios with different types of noise (e.g., Ben-David et al., 2012) to further examine the contribution of sensory factors to the perception of emotion in speech. We wish to also consider that age-related differences in the processing of emotion in speech may reflect a difference in strategy instead of a deficit. For example, sensory degradation may reduce the availability of acoustic prosodic cues, although seniors' lexicons easily exceed those of their younger counterparts (for a recent meta-analysis, see Ben-David et al., 2016). As a consequence, older listeners may adapt and rely more heavily on the semantic content of speech. This may relate to Hess's (2006) model on social cognitive perception. Namely, as older adults are motivated to conserve resources, they are more sensitive than younger adults in their allocation of resources to enhance performance (see Hess, 2014). Thus, they may opt to rely on intact abilities, to compensate for difficulties in other abilities, using their extensive experience.

Replication of the T-RES

The results demonstrate the replicability of the main trends of the original English version of the T-RES as tested with Canadian younger adults, with the Hebrew version tested with Israeli younger adults. (a) In terms of identification of emotions, in both versions, younger adults easily processed the emotional content of the prosody and semantics separately; (b) in terms of selective attention, both the English and Hebrew versions yielded significant failures of selective attention; and (c) in terms of integration of channels, in the general rating task, the same linear trend was observed in both versions for younger adults: target-emotion–matched > target-emotion–prosody > target-emotion–semantics > target-emotion–absent trials. In other words, ratings demonstrated a prosodic dominance and a supremacy of congruency for younger adults in both versions.

Caveat and Future Studies

In our study, to somewhat mitigate age-related and individual differences in auditory sensitivity, speech stimuli were presented at a set amplitude above individual pure-tone audiometric thresholds. Clearly, this manipulation does not fully compensate for possible auditory differences. Moreover, our tests for selective attention do not reveal the full scope of age-related and individual differences in cognitive abilities (for a discussion, see Wingfield, 2016). Future studies may wish to focus on these issues, directly weighing the separate role of cognitive and sensory factors in generating the age-related differences in perception of emotion in speech, as this was beyond the scope of the current study.

We also note that we did not ask participants to overtly report on the strategy they used in rating the emotions presented by the spoken sentences. We hope that future research can delve into this and examine if listeners' self-reports match the statistical analysis of their rating performance. Furthermore, as this is the first application of the T-RES with an older population, we did not fully address the possible discreet-emotion effects. Appendix B shows that the main trends were mostly replicated in each emotion, but not to the same extent. This sets the foundation for future studies that may wish to examine whether older and younger adults perform differently in the presence of specific emotions. Finally, as one speaker recorded all stimuli in the T-RES, it may potentially decrease the generalizability of our data. On the other hand, we maintain that this choice provides a more accurate gauge of age-related differences in performance. Future research may further assess performance with multiple speakers as well.

Summary

This study presents a novel outlook on age-related differences in processing emotion in spoken language. We focused on the different weights older and younger adults assign to the different channels and not only on age-related differences in identification of channel-specific cues. Age-related differences were obtained even in ideal listening conditions. Results may point to a cognitive source, a sensory source, or a difference in strategy between age groups.

Apart from the theoretical relevance of the current study, there are also clear practical implications. Based on UN projections, by the year 2025, about a quarter of Europeans (107 million) will be over the age of 65 years (Muenz, 2007). This signifies a complete reversal in the makeup of the European population from 50 years ago. Correct identification of emotional cues in speech is crucial for efficient social interactions (e.g., Kotz & Paulmann, 2007; Mayo, Florentine, & Buus, 1997; Pell, Paulmann, Dara, Alasseri, & Kotz, 2009; for a review, see Juslin & Laukka, 2003) and is associated with one's ability to maintain healthy and successful relationships (Carton et al., 1999). Specifically, due to mobility challenges, many seniors rely more on communication over the phone (where no visual cues are available) or text messages (where no prosodic cues are available). A better understanding of the sources of difficulties in processing emotional speech in aging can promote healthy aging and independence.

The current study further demonstrated that the T-RES paradigm can be used with different populations and age groups, across cultures and languages (on adapting speech tasks for aging individuals, see Ben-David & Icht, 2016; across languages, Icht & Ben-David, 2015). Future studies can use this paradigm to test processes underlying the perception of emotion in speech in other populations for whom emotion processing deficits have been identified such as individuals with autism spectrum disorder, attention-deficit/hyperactivity disorder, traumatic brain injury, cochlear implants, and other auditory pathologies.

Acknowledgments

The research was partially supported by a Marie Curie Career Integration Grant (FP7-PEOPLE-2012-CIG) from the European Commission awarded to the first author.

References

Abo Foul, Y., Eitan, R., & Aviezer, H. (2018). Perceiving emotionally incongruent cues from faces and bodies: Older adults get the whole picture. *Psychology and Aging, 33*(4), 660–666.

Beaucousin, V., Lacheret, A., Turbelin, M.-R., Morel, M., Mazoyer, B., & Tzourio-Mazoyer, N. (2006). FMRI study of emotional speech comprehension. *Cerebral Cortex, 17*(2), 339–352.

Ben-David, B. M., & Algom, D. (2009). Species of redundancy in visual target detection. *Journal of Experimental Psychology: Human Perception and Performance, 35*(4), 958–976.

Ben-David, B. M., Eidels, A., & Donkin, C. (2014). Effects of aging and distractors on detection of redundant visual targets and capacity: Do older adults integrate visual targets differently than younger adults? *PLoS One, 9*(12), e113551.

Ben-David, B. M., & Icht, M. (2016). Voice changes in real speaking situations during a day, with and without vocal loading: Assessing call center operators. *Journal of Voice, 30*(2), 247.e1–247.e11.

Ben-David, B. M., Multani, N., Shakuf, V., Rudzicz, F., & van Lieshout, P. H. (2016). Prosody and semantics are separate but not separable channels in the perception of emotional speech: Test for Rating of Emotions in Speech. *Journal of Speech, Language, and Hearing Research, 59*(1), 72–89.

Ben-David, B. M., & Schneider, B. A. (2009). A sensory origin for color-word Stroop effects in aging: A meta-analysis. *Aging, Neuropsychology, and Cognition, 16*(5), 505–534.

Ben-David, B. M., & Schneider, B. A. (2010). A sensory origin for color-word Stroop effects in aging: Simulating age-related changes in color-vision mimics age-related changes in Stroop. *Aging, Neuropsychology, and Cognition, 17*(6), 730–746.

Ben-David, B. M., Thayapararajah, A., & van Lieshout, P. H. (2013). A resource of validated digital audio recordings to assess identification of emotion in spoken language after a brain injury. *Brain Injury, 27*(2), 248–250.

Ben-David, B. M., Tse, V., & Schneider, B. A. (2012). Does it take older adults longer than younger adults to perceptually segregate a speech target from a background masker? *Hearing Research, 290*(1–2), 55–63.

Ben-David, B. M., van Lieshout, P. H., & Leszcz, T. (2011). A resource of validated affective and neutral sentences to assess identification of emotion in spoken language after a brain injury. *Brain Injury, 25*(2), 206–220.

Bennett, D., Schneider, J., Arvanitakis, Z., Kelly, J., Aggarwal, N., Shah, R., & Wilson, R. (2006). Neuropathology of older persons without cognitive impairment from two community-based studies. *Neurology, 66*(12), 1837–1844.

Brodsky, J., Shnoor, Y., & Be'er, S. (2003). *The elderly in Israel: The 2003 statistical abstract* [Report]. Jerusalem, Israel: Mashav (Hebrew).

Brosgole, L., & Weisman, J. (1995). Mood recognition across the ages. *International Journal of Neuroscience, 82*(3–4), 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*(1), 91–100.

Dmitrieva, E., & Gelman, V. Y. (2012). The relationship between the perception of emotional intonation of speech in conditions of interference and the acoustic parameters of speech signals in adults of different gender and age. *Neuroscience and Behavioral Physiology, 42*(8), 920–928.

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

Dupuis, K., & Pichora-Fuller, M. K. (2015). Aging affects identification of vocal emotions in semantically neutral sentences. *Journal of Speech, Language, and Hearing Research, 58*(3), 1061–1076.

Garner, W. R., & Felfoldy, G. L. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology, 1*(3), 225–241.

Gordon-Salant, S., Yeni-Komshian, G. H., Fitzgibbons, P. J., & Barrett, J. (2006). Age-related differences in identification and discrimination of temporal cues in speech segments. *The Journal of the Acoustical Society of America, 119*(4), 2455–2466.

Grant, D. A. (1948). The Latin square principle in the design and analysis of psychological experiments. *Psychological Bulletin, 45*(5), 427–442.

Grunwald, I. S., Borod, J. C., Obler, L. K., Erhan, H. M., Pick, L. H., Welkowitz, J., … Whalen, J. (1999). The effects of age and gender on the perception of lexical emotion. *Applied Neuropsychology, 6*(4), 226–238.

Hadar, B., Skrzypek, J. E., Wingfield, A., & Ben-David, B. M. (2016). Working memory load affects processing time in spoken word recognition: Evidence from eye-movements. *Frontiers in Neuroscience, 10,* 221.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. *Psychology of Learning and Motivation, 22,* 193–225.

Healey, M. K., Campbell, K. L., & Hasher, L. (2008). Cognitive aging and increased distractibility: Costs and potential benefits. *Progress in Brain Research, 169,* 353–363.

Heinrich, A., Gagne, J.-P., Viljanen, A., Levy, D. A., Ben-David, B., & Schneider, B. A. (2016). Effective communication as a fundamental aspect of active aging and well-being: Paying attention to the challenges older adults face in noisy environments. *Social Inquiry into Well-Being, 2*(1), 51–69.

Hertzog, C., Kramer, A. F., Wilson, R. S., & Lindenberger, U. (2008). Enrichment effects on adult cognitive development: Can the functional capacity of older adults be preserved and enhanced? *Psychological Science in the Public Interest, 9*(1), 1–65.

Hess, T. M. (2006). Adaptive aspects of social cognitive functioning in adulthood: Age-related goal and knowledge influences. *Social Cognition, 24*(3), 279–309.

Hess, T. M. (2014). Selective engagement of cognitive resources: Motivational influences on older adults' cognitive functioning. *Perspectives on Psychological Science, 9*(4), 388–407.

Holt-Lunstad, J., Smith, T. B., & Layton, J. B. (2010). Social relationships and mortality risk: A meta-analytic review. *PLoS Medicine, 7*(7), e1000316.

Humes, L. E. (1996). Speech understanding in the elderly. *Journal of the American Academy of Audiology, 7,* 161–167.

Humes, L. E. (2007). The contributions of audibility and cognitive factors to the benefit provided by amplified speech to older adults. *Journal of the American Academy of Audiology, 18*(7), 590–603.

Humes, L. E., Lee, J. H., & Coughlin, M. P. (2006). Auditory measures of selective and divided attention in young and older adults using single-talker competition. *The Journal of the Acoustical Society of America, 120*(5), 2926–2937.

Icht, M., & Ben-David, B. M. (2014). Oral-diadochokinesis rates across languages: English and Hebrew norms. *Journal of Communication Disorders, 48,* 27–37.

Icht, M., & Ben-David, B. M. (2015). Oral-diadochokinetic rates for Hebrew-speaking school-age children: Real words vs. nonwords repetition. *Clinical Linguistics & Phonetics, 29*(2), 102–114.

Ishii, K., Reyes, J. A., & Kitayama, S. (2003). Spontaneous attention to word content versus emotional tone: Differences among three cultures. *Psychological Science, 14*(1), 39–46.

Jacob, H., Brück, C., Plewnia, C., & Wildgruber, D. (2014). Cerebral processing of prosodic emotional signals: Evaluation of a network model using rTMS. *PLoS One, 9*(8), e105509.

Jerger, S., Pirozzolo, F., Jerger, J., Elizondo, R., Desai, S., Wright, E., & Reynosa, R. (1993). Developmental trends in the interaction between auditory and linguistic processing. *Attention, Perception, & Psychophysics, 54*(3), 310–320.

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin, 129*(5), 770–814.

Kitayama, S., & Ishii, K. (2002). Word and voice: Spontaneous attention to emotional utterances in two languages. *Cognition & Emotion, 16*(1), 29–59.

Kotz, S. A., & Paulmann, S. (2007). When emotional prosody and semantics dance cheek to cheek: ERP evidence. *Brain Research, 1151,* 107–118.

Laukka, P. (2003). Categorical perception of emotion in vocal expression. *Annals of the New York Academy of Sciences, 1000*(1), 283–287.

Lindenberger, U., & Baltes, P. B. (1994). Sensory functioning and intelligence in old age: A strong connection. *Psychology and Aging, 9*(3), 339–355.

Lindenberger, U., & Ghisletta, P. (2009). Cognitive and sensory declines in old age: Gauging the evidence for a common cause. *Psychology and Aging, 24*(1), 1–16.

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the *Depression Anxiety Stress Scales* (DASS) with the *Beck Depression and Anxiety Inventories*. *Behaviour Research and Therapy, 33*(3), 335–343.

Mayo, L. H., Florentine, M., & Buus, S. (1997). Age of second-language acquisition and perception of speech in noise. *Journal of Speech, Language, and Hearing Research, 40*(3), 686–693.

McDowd, J. M., & Shaw, R. J. (2000). Attention and aging: A functional perspective. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (2nd ed., pp. 221–292). Mahwah, NJ: Erlbaum.

Mehrabian, A., & Wiener, M. (1967). Decoding of inconsistent communications. *Journal of Personality and Social Psychology, 6*(1), 109–114.

Mitchell, R. L. (2006). How does the brain mediate interpretation of incongruent auditory emotions? The neural response to prosody in the presence of conflicting lexico-semantic cues. *European Journal of Neuroscience, 24*(12), 3611–3618.

Mitchell, R. L. (2007). Age-related decline in the ability to decode emotional prosody: Primary or secondary phenomenon? *Cognition and Emotion, 21*(7), 1435–1454.

Mitchell, R. L., & Kingston, R. A. (2011). Is age-related decline in vocal emotion identification an artefact of labelling cognitions? *International Journal of Psychological Studies, 3*(2), 156–163.

Monge, Z. A., & Madden, D. J. (2016). Linking cognitive and visual perceptual decline in healthy aging: The information degradation hypothesis. *Neuroscience & Biobehavioral Reviews, 69,* 166–173.

Muenz, R. (2007). *Aging and demographic change in European societies: Main trends and alternative policy options* (World Bank SP Discussion Paper No. 703). Washington, DC: World Bank.

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., … Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society, 53*(4), 695–699.

Nygaard, L. C., & Queen, J. S. (2008). Communicating emotion: Linking affective prosody and word meaning. *Journal of Experimental Psychology: Human Perception and Performance, 34*(4), 1017–1030.

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.

Paulmann, S., Jessen, S., & Kotz, S. A. (2009). Investigating the multimodal nature of human communication: Insights from ERPs. *Journal of Psychophysiology, 23*(2), 63–76.

Paulmann, S., Pell, M. D., & Kotz, S. A. (2008). How aging affects the recognition of emotional speech. *Brain and Language, 104*(3), 262–269.

Pell, M. D., Paulmann, S., Dara, C., Alasseri, A., & Kotz, S. A. (2009). Factors in the recognition of vocally expressed emotions: A comparison of four languages. *Journal of Phonetics, 37*(4), 417–435.

Phillips, L. H., MacLean, R. D., & Allen, R. (2002). Age and the understanding of emotions: Neuropsychological and sociocognitive perspectives. *Journals of Gerontology: Series B: Psychological Sciences and Social Sciences, 57*(6), P526–P530.

Pichora-Fuller, M. K., Schneider, B. A., MacDonald, E., Pass, H. E., & Brown, S. (2007). Temporal jitter disrupts speech intelligibility: A simulation of auditory aging. *Hearing Research, 223*(1), 114–121.

Polku, H., Mikkola, T. M., Portegijs, E., Rantakokko, M., Kokko, K., Kauppinen, M., … Viljanen, A. (2015). Life-space mobility and dimensions of depressive symptoms among community-dwelling older adults. *Aging & Mental Health, 19*(9), 781–789.

Roche, J. M., Peters, B., & Dale, R. (2015). "Your tone says it all": The processing and interpretation of affective language. *Speech Communication, 66,* 47–64.

Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology, 32*(1), 76–92.

Schmidt, J., Janse, E., & Scharenborg, O. (2016). Perception of emotion in conversational speech by younger and older listeners. *Frontiers in Psychology, 7,* 781.

Schneider, B. A., Daneman, M., & Murphy, D. R. (2005). Speech comprehension difficulties in older adults: Cognitive slowing or age-related changes in hearing? *Psychology and Aging, 20*(2), 261–271.

Schneider, B. A., Pichora-Fuller, K., & Daneman, M. (2010). Effects of senescent changes in audition and cognition on spoken language comprehension. In S. Gordon-Salant, R. Frisina, A. Popper, & R. Fay (Eds.), *The aging auditory system* (pp. 167–210). New York, NY: Springer. https://doi.org/10.1007/978-1-4419-0993-0_7

Schneider, B. A., & Pichora-Fuller, M. K. (2000). Implications of perceptual deterioration for cognitive aging research. In F. I. M. Craik & T. A. Salthouse (Eds.), *Handbook of cognitive aging II* (pp. 155–219). Mahwah, NJ: Erlbaum.

Sen, A., Isaacowitz, D., & Schirmer, A. (2017). Age differences in vocal emotion perception: On the role of speaker age and listener sex. *Cognition and Emotion, 32*(6), 1189–1204.

Shakuf, V., & Ben-David, B. M. (2014). *Who is afraid of the bear? Hebrew test for recognition of emotions in spoken language for the diagnosis and evaluation of abilities and disabilities.* Paper presented at the Israeli Speech Hearing Language Association, Tel Aviv, Israel.

Shakuf, V., Gal-Rosenblum, S., & Ben-David, B. M. (2016). The psychophysics of aging. In emotional speech, older adults attend to semantic, while younger adults to the prosody. In I. Skotnikova, O. Korolkova, I. Blinnikova, V. Doubrovski, V. Shendyapin, & N. Volkova (Eds.), *Fechner Day 2016: Proceedings of the 32nd Annual Meeting of the International Society for Psychophysics* (p. 89). Moscow, Russia: Institute of Psychology of Russian Academy of Sciences, Publishing House.

Smith, S. M., & Kampfe, C. M. (1997). Interpersonal relationship implications of hearing loss in persons who are older. *Journal of Rehabilitation, 63*(2), 15–21.

Sommers, M. S., & Danielson, S. M. (1999). Inhibitory processes and spoken word recognition in young and older adults: The interaction of lexical competition and semantic context. *Psychology and Aging, 14,* 458–472.

Troyer, A. K., Leach, L., & Strauss, E. (2006). Aging and response inhibition: Normative data for the Victoria Stroop Test. *Aging, Neuropsychology, and Cognition, 13*(1), 20–35.

Wechsler, D., Coalson, D., & Raiford, S. (2008). *WAIS-IV. Wechsler Adult Intelligence Scale–Fourth Edition: Technical and interpretative manual.* San Antonio, TX: NCS Pearson.

Wingfield, A. (2016). Evolution of models of working memory and cognitive resources. *Ear and Hearing, 37,* 35S–43S.

Wurm, L. H., Vakoch, D. A., Strasser, M. R., Calin-Jageman, R., & Ross, S. E. (2001). Speech perception and vocal expression of emotion. *Cognition & Emotion, 15*(6), 831–852.

Zalsman, G., Aizenberg, D., Sigler, M., Nahshoni, E., & Weizman, A. (1998). Geriatric Depression Scale–Short Form—Validity and reliability of the Hebrew version. *Clinical Gerontologist, 18*(3), 3–9.

Zupan, B., Neumann, D., Babbage, D. R., & Willer, B. (2009). The importance of vocal affect to bimodal processing of emotion: Implications for individuals with traumatic brain injury. *Journal of Communication Disorders, 42*(1), 1–17.

Appendix A

Formulae for Statistical Analyses

1. Identification of emotions. In the semantic and prosodic rating tasks, two measures were used. To gauge identification of emotion, the difference between ratings of sentences that present the target emotion in the target channel were compared with sentences that do not present the target emotion in either channel. This was conducted with baseline sentences. Formally:

$$\begin{aligned}\text{Identification}_{\text{semantics}_x} &= \\ \text{semantics}_x\text{rating}|\text{semantics}_x - \frac{1}{3}\sum\nolimits_y \text{semantics}_x\text{ratings}|\text{semantics}_y) &\\ \text{Identification}_{\text{prosody}_x} &= \\ \text{prosody}_x\text{rating}|\text{prosody}_x - \frac{1}{3}\sum\nolimits_y \text{prosody}_x\text{ratings}|\text{prosody}_y)&\end{aligned} \tag{1}$$

where *x* represents one of the four rated emotions and *y* represents the other three emotions that are not *x*.

2. Selective attention. To test selective attention, the difference between ratings of sentences that present the target emotion only in the to-be-ignored channel were compared with sentences that do not present the target emotion in either channel. Formally:

$$\begin{aligned}&\text{SelectiveAttention}_{\text{semantics}_x} \\ &\quad = \frac{1}{4}\sum\nolimits_y \left[(\text{prosody}_x\text{rating}|\text{prosody}_y \ \&\ \text{semantics}_x)\right. \\ &\quad \left. - \frac{1}{4}\sum\nolimits_z (\text{prosody}_x\text{rating}|\text{prosody}_y \ \&\ \text{semantics}_z)\right] \\ &\text{SelectiveAttention}_{\text{prosody}_x} \\ &\quad = \frac{1}{4}\sum\nolimits_y \left[(\text{semantics}_x\text{rating}|\text{semantics}_y \ \&\ \text{prosody}_x)\right. \\ &\quad \left. - \frac{1}{4}\sum\nolimits_z (\text{semantics}_x\text{rating}|\text{semantics}_y \ \&\ \text{prosody}_z)\right]\end{aligned} \tag{2}$$

where *x* represents one of the four rated emotions and *y* and *z* represent the other three emotions and the neutral emotion.

3. Integration of channels. In the general rating tasks, we measured average rating for target-emotion–matched trials, target-emotion–prosodic trials, target-emotion–semantics trials, and target-emotion–absent trials, as formally presented here:

$$\begin{aligned}\text{Target}-\text{Emotion}-\text{Matched}_x &= \text{rating}_x|\text{semantics}_x \ \&\ \text{prosody}_z \\ \text{Target}-\text{Emotion}-\text{Prosody}_x &= \frac{1}{3}\sum\nolimits_y (\text{rating}_x|\text{semantics}_y \ \&\ \text{prosody}_x) \\ \text{Target}-\text{Emotion}-\text{Semantics}_x &= \frac{1}{3}\sum\nolimits_y (\text{rating}_x|\text{semantics}_x \ \&\ \text{prosody}_y) \\ \text{Target}-\text{Emotion}-\text{Absent}_x &= \frac{1}{9}\sum\nolimits_{yz} (\text{rating}_x|\text{semantics}_y \ \&\ \text{prosody}_z)\end{aligned} \tag{3}$$

where *x* represents one of the four rated emotions and *y* and *z* represent the other three emotions.

Appendix B

Analysis per Emotion

1. Identification of emotions. Analysis shows that identification interacted significantly with the type of target emotion, $F(1, 74) = 1499.3$, $p < .001$, $\eta_p^2 = .95$, but most importantly, this did not interact with age group. Separate analyses for each target emotion indicate the same major trend—a significant effect for emotion identification [anger: $F(1, 74) = 902.9$, $p < .001$, $\eta_p^2 = .91$; sad: $F(1, 74) = 1045.1$, $p < .001$, $\eta_p^2 = .93$; happy: $F(1, 74) = 774.5$, $p < .001$, $\eta_p^2 = .91$; fear: $F(1, 74) = 425.2$, $p < .001$, $\eta_p^2 = .85$], that interacted significantly with age group [anger: $F(1, 74) = 23.5$, $p < .001$, $\eta_p^2 = .24$; sad: $F(1, 74) = 28.4$, $p < .001$, $\eta_p^2 = .28$; happy: $F(1, 74) = 11.5$, $p = .001$, $\eta_p^2 = .13$; fear: $F(1, 74) = 23.8$, $p < .001$, $\eta_p^2 = .24$].

2. Selective attention failures. Examining selective attention, we see a very similar trend: The selective attention measure significantly interacted with the target emotion, $F(3, 72) = 25.3$, $p \leq .001$, $\eta_p^2 = .51$. However, the triple interaction of Selective Attention × Age Group × Target Emotion was not significant, $F(3, 72) = 1.5$, $p = .24$. Separate analyses for each target emotion indicate the same major trend—a significant effect for selective attention [anger: $F(1, 74) = 121.2$, $p < .001$, $\eta_p^2 = .62$; sad: $F(1, 74) = 7.9$, $p = .002$, $\eta_p^2 = .12$; happy: $F(1, 74) = 16.7$, $p < .001$, $\eta_p^2 = .18$; fear: $F(1, 74) = 30.0$, $p < .001$, $\eta_p^2 = .29$]. An age-related increase in failures of selective attention was observed in all emotions but reached significance only in happiness ratings [anger: $F(1, 74) = 1.1$, $p = .29$; sad: $F(1, 74) = 0.29$, $p = .59$; happy: $F(1, 74) = 13.7$, $p < .001$, $\eta_p^2 = .16$; fear: $F(1, 74) = 1.5$, $p = .22$].

3. Integration of channels. First, note that the linear trend observed in the cross-emotional analysis reappeared in the separate emotion-specific analyses: matched > prosody > semantics > absent trials [anger: $F(1, 74) = 901.5$, $p < .001$, $\eta_p^2 = .92$; sad: $F(1, 74) = 852.6$, $p < .001$, $\eta_p^2 = .92$; happy: $F(1, 74) = 985.3$, $p < .001$, $\eta_p^2 = .93$; fear: $F(1, 74) = 590.6$, $p < .001$, $\eta_p^2 = .89$]. This linear trend interacted significantly with age group in all emotions [anger: $F(1, 74) = 44.0$, $p < .001$, $\eta_p^2 = .37$; sad: $F(1, 74) = 28.3$, $p < .001$, $\eta_p^2 = .28$; happy: $F(1, 74) = 23.1$, $p < .001$, $\eta_p^2 = .24$; fear: $F(1, 74) = 14.5$, $p < .001$, $\eta_p^2 = .16$]. Next, separate emotion-specific analyses of prosodic dominance show that the interaction of prosodic dominance and age group was replicated as well [anger: $F(1, 74) = 24.9$, $p < .001$, $\eta_p^2 = .25$; sad: $F(1, 74) = 19.0$, $p < .001$, $\eta_p^2 = .20$; happy: $F(1, 74) = 11.1$, $p = .001$, $\eta_p^2 = .13$; fear: $F(1, 74) = 3.5$, $p = .06$, $\eta_p^2 = .05$].