# Clear speech adaptations in spontaneous speech produced by young and older adults [1]

*Valerie Hazan[a], Outi Tuomainen[a], Jeesun Kim[b], Christopher Davis[b], Benjamin Sheffield[c], Douglas Brungart[c]*

b) Department of Speech Hearing and Phonetic Sciences, UCL (University College London), 2 Wakefield Street, London WC$^1$N 1PF
c) The MARCS Institute, Western Sydney University, Locked Bag 1797, Penrith NSW 2751, Australia
d) Audiology and Speech-Pathology Center, Walter Reed National Military Medical Center, Bethesda, 4494 North Palmer Road, Bethesda, MD 20889, USA.

**Authors' manuscript of paper accepted for publication at the Journal of the Acoustical Society of America**

**Accepted for publication on: 17 August 2018**

## ABSTRACT

The study investigated the speech adaptations by older adults with and without age-related hearing loss made to communicate effectively in challenging communicative conditions. Acoustic analyses were carried out on spontaneous speech produced during a problem-solving task (diapix) carried out by talker pairs in different listening conditions. There were 83 talkers of Southern British English. 57 talkers were older adults (OA) aged 65-84: 30 with normal hearing (OANH) and 27 (OAHL) with presbycusis (mean PTA .250-4kHz: 27.7 dB HL). 26 talkers were younger adults (YA) aged 18-26 with normal hearing. Participants were recorded while completing the diapix task with a conversational partner (YA of the same sex) when (a) both talkers heard normally (NORM), (b) the partner had a simulated hearing loss (HLS) and (c) both talkers heard babble noise (BAB2). Irrespective of hearing status, there were age-related differences in some acoustic characteristics of YA and OA speech produced in NORM, most likely linked to physiological factors. In challenging conditions, while OANH talkers typically patterned with YA talkers, OAHL talkers made adaptations more consistent with an increase in vocal effort. Our study suggests that even mild presbycusis in healthy older adults can affect the speech adaptations made to maintain effective communication.

PACS number: 43.70 Mn

## I. INTRODUCTION

The effects of age and of hearing loss (presbycusis) on the ability to understand speech are well documented, but less attention has been given to whether these factors affects the production of speech when communication occurs in either good or challenging conditions. One key issue is whether older adults are as able as younger adults to make 'clear speech adaptations' in challenging conditions in order to overcome communication difficulties, and whether these are affected by the hearing status of the older adult. Also, are these adaptations as effective as those of younger adults? To examine these questions, it is essential to use experimental approaches that involve talker-listener interaction, as imagined interactions may not lead to the same behaviour as actual communication with a conversational partner (e.g., Scarborough and Zellou, 2013). In order to investigate the degree to which older adults are able to make such adaptations to ensure effective communication, we conducted acoustic-phonetic analyses of the speech produced by older adults and younger adult controls while carrying out a problem-solving task with another adult in easy and challenging communicative conditions.

Speech communication between two or more speakers is highly dynamic, with talkers adapting the degree of clarity of their speech according to the needs of their interlocutor. As argued in models of speech production that focus on the effects of speaker-listener interaction (e.g., Lindblom, 1990), when the information to be exchanged is highly predictable, little talker effort is required; a casual speaking style, characterised by a fast speaking rate and a high degree of hypo-articulation, may be acceptable. Depending on the context and intent of the message to be transmitted, successful communication can occur even in poor listening environments. However, there are many situations in which greater speech clarity is necessary. This is the case when the message to be transmitted has low predictability and when there is noise in the environment, when one or both speakers have a hearing loss, or when the two speakers interacting do not share a common native language. Young adults are adept at making acoustic and linguistic adaptations to their speech to facilitate communication in such situations. This has been shown extensively in studies where young adults are instructed to read sentences clearly (for a review, see Smiljanić and Bradlow, 2009). Some studies have involved

---

[1] Portions of this work were presented at Interspeech 2016, 8-12 September 2016, San Francisco, USA, at the 2nd Workshop on Psycholinguistic Approaches to Speech Recognition in Adverse Conditions (PASRAC2), October 2016, Nijmegen, the Netherlands, at the 5th Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan, 1-5 December 2016, Honolulu, Hawaii and at the Aging & Speech Communication Research Conference 2017, November, 2017, Tampa.

pairs of talkers interacting in different listening environments, and these adaptations have been shown to be attuned, to an extent, to the needs of the conversational partner, as the adaptations varied according to the type of challenging condition faced by the interlocutor (e.g., Cooke and Lu, 2010; Hazan and Baker, 2011). Adolescents too are able to produce clear speech adaptations when conversing in challenging conditions, although younger children may use less skilled strategies such as increasing their vocal effort (Hazan, Tuomainen and Pettinato, 2016).

For older adults, the evidence of whether they make clear speech adaptations is mixed. In an study using the interactive Map Task (Anderson et al., 1991) with dyads of young-young, older-older and young-older adults, Kemper et al. (1995) found that while younger adults adapted their speaking style when talking to older adults by varying characteristics such as fluency, grammatical complexity, semantic content, older adults showed little variation in their speech across conditions. However, in another study focusing on discourse strategies by older women communicating with young students and young adults with mild intellectual difficulties, there was evidence that older women were adjusting their discourse strategies to accommodate for their conversational partners (Gould and Shaleen, 1999).

In terms of studies focusing on the acoustic characteristics of speech, most evidence about clear speech adaptations comes from studies that involved participants reading sentences in an ‘instructed’ clear speaking style, i.e. when asked to speak as if to an imagined person with a hearing loss or who is not a native speaker. A recent lifespan study including young and older adults has directly investigated the acoustic characteristics of their casual and instructed clear speaking styles, among other conditions (Smiljanić and Gilbert, 2017a). When asked to read sentences clearly in a quiet environment, older adults reduced their speech rate and produced longer pauses than when speaking casually. However, they showed less of an increase in mean fundamental frequency and amount of energy in the mid frequency region of their speech than other talker groups. In a linked paper, the clear speech produced by all talker groups, including older adults, was more intelligible when presented in noise than their conversational speech (Smiljanić and Gilbert, 2017b). This suggests that there was a clear speech benefit (i.e. increased intelligibility relative to that for casual speech) for the speech of older talkers and that the changes that they made in their clear speaking style were effective.

In some studies, the degree to which older talkers adapted their speech was solely determined by investigating the impact of their speech on listeners’ perception. In Smiljanić (2013), the degree to which older adults were able to make adaptations to their speech varied across the five individuals whose speech was included in the intelligibility tests. In Schum (1996), involving 10 older and 10 younger talkers with sentences in conversational and clear elicited styles, no statistically significant difference in clear speech benefit was shown across the two groups; however, an analysis of individual talkers also revealed that, in each group, some talkers showed a substantial clear speech benefit while little benefit was obtained for others. These studies suggest that clear speech strategies vary across older talkers, with some individuals being more effective at making adaptations than others. Although these studies provide valuable insights into older talkers’ ability to produce clear speech adaptations, none involved naturally-elicited speech adaptations within a communicative setting. Also, they do not enable us to examine the impact of hearing status within the older population as participant groups were not differentiated according to this status.

The acoustic characteristics of speech produced by older adults may differ from those of younger adults because of a number of factors (for detailed reviews, see Hooper and Cralidis, 2009; Gordon-Salant, 2014). First, there are physical changes to the vocal apparatus that occur with increasing age. There are documented increases in vocal tract length in older adults, resulting in increased vocal tract volume (Xue and Hao, 2003) while physiological changes to the larynx include a thinning of vocal folds and hardening of laryngeal cartilages (e.g. Sataloff et al., 1997; Baken, 2005). Aging has also been shown to have an effect on respiratory function for speech (e.g., Huber and Spruill, 2008) with some evidence that the effect may be greater in men than women. Goy et al. (2013) carried out voice analyses on a large and well-controlled speech corpus of young and older healthy adults. Although they concluded that age-related changes were smaller than previously reported in smaller-scales studies, they did replicate the well-documented reduction in fundamental frequency (F0) in older women although did not replicate the finding in many studies (e.g. Hollien and Ship, 1972; Torre and Barlow, 2009) of F0 increases in men; age-related effects on voice regularity and intensity varied with talker sex (Goy et al., 2013). Physical changes to the vocal apparatus also lead to changes in vocal tract resonances, resulting in changes in vowel formant frequencies (Linville and Rens, 2001; Xue and Hao, 2003), also seen in longitudinal analyses of formant frequencies of talkers recorded over several decades (Reubold, Harrington and Kleber, 2010).

Motor control also appears to be reduced in older speakers compared to young adults: older adults show greater within-speaker variation in articulatory movement and placement, at least for more complex speech items (Sadagopan and Smith, 2013). There are cognitive changes that may affect the willingness to make additional efforts to be understood and the empathy experienced towards a conversational partner (e.g. Zhang et al., 2013) although motivation has also been shown to affect the performance of older adults in laboratory settings (Spaniol, Voss, Bowen, and Grady, 2011). Finally, many older adults experience a degree of age-related hearing loss, or presbycusis, that affects hearing thresholds, frequency resolution and the ability to perceive speech effectively when masked by noise or other voices (for reviews, see Gordon-Salant, 2005, 2014). In sum, for older adults, reduced motor control (at least for precise articulations), greater listener effort and cognitive decline all make it likely that communication may, in many instances, require greater effort than for younger adults. What is less known is how these potential difficulties affect older adults’ ability and willingness to increase their speaking effort in order to produce clearer speech for the benefit of their interlocutor while conversing in challenging situations.

This study examines changes in acoustic-phonetic characteristics that have been shown to vary across casual and clear speaking styles. It focuses on three strategies for producing clear speech that have been documented across clear speech studies (for a review, see Smiljanić and Bradlow, 2009): reductions in articulation rate, increases in vocal effort, and hyper-articulation.

Reductions in articulation rate have consistently been shown to occur in many studies of clear speaking styles (e.g., Picheny et al., 1986; Uchanski et al., 1996; Hazan and Baker, 2011) although it has also been shown that a slower speech rate is not a necessary feature of clear speech (Krause and Braida, 2004). A first aim of this study is therefore to investigate changes in articulation rate across communicative conditions in older adults and young adult controls. Even when speaking without interference, differences in articulation rate across the lifespan have been documented in a number of studies (e.g., Duchin and Mysak, 1987; Jacewicz, Fox and Wei, 2010; Bilodeau-Mercure and Tremblay, 2016; Bóna, 2014; Smiljanić and Gilbert, 2017a). For example, a lifespan study showed that articulation rate measured from spontaneous speech monologues increased from childhood into adulthood and did not peak until adults were in their mid-40s (Jacewicz et al, 2010). Bóna (2014) examined articulation rate in older adults in a range of different speaking styles which had differing cognitive demands. Older adults had a slower articulation rate than young adults in all speaking styles, and greater frequency of pauses and slower speech rate in most speaking styles. Articulation rate did not differ across styles for older adults but did for younger adults.

Another strategy for speaking clearly when communicating in challenging environmental conditions such as when speaking in background noise, is to speak more loudly, which may involve increased vocal effort (Traunmueller and Eriksson, 2000; Garnier and Heinrich, 2014). In addition to raising intensity levels and changing the spectral balance of the long-term spectrum of speech, increased vocal effort also brings about increases in F0 and in the frequency of the first formant (F1) of vowels. In a study involving communication between two conversational partners in challenging conditions, children aged 9-12 years were seen to use this strategy: there were strong correlations, in their clear speech, between increases in F0 and in mean energy in the mid-frequency regions of their voice relative to their casual speech, while this correlation was absent in young adults (Hazan et al., 2016).

Finally, increasing the degree of articulation (hyper-articulation) is a strategy that has been documented in many studies of clear speech (e.g., Smiljanić and Bradlow, 2005; Ferguson and Kewley-Port, 2007; Hazan and Baker, 2011). Increases in hyper-articulation are most often demonstrated via vowel measures such as vowel space area or first and second formant ranges. Increases in the degree of articulation, as demonstrated by such measures, could partly result from the use of a lower articulation rate; indeed, such a strategy was less evident in clear speech produced at normal conversational speech rates (Krause and Braida, 2004).

In our study, two challenging conditions, reflecting ecologically-valid communicative situations, were chosen to naturally elicit clear speaking styles. Both conditions involved exchanges of information between two participants aimed at successfully completing a task together. These exchanges would entail both participants in the interaction potentially showing some adaptation or accommodation to their communication partner, but our focus is on the talker who was told to lead the interaction and who did most of the talking (Talker A). The two challenging conditions varied in the degree to which the adaptations were made for the benefit of the interlocutor (interlocutor-oriented speaking style) or as a result of a direct impact of the challenging condition on the talker (self-oriented speaking style). In the first challenging condition, Talker A was told that their conversational partner had a simulated severe-to-profound loss (HLS). Talker A, whose communication channel was unaffected, had to speak clearly to ensure successful communication even if this was at the cost of increased effort for themselves, as predicated by Lindblom (1990), and so the clear speaking style was primarily interlocutor-oriented. The second condition, in which both conversational partners were exposed to babble noise (BAB2), reflects communication in noisy environments, known to be particularly challenging for older adults. Here, the adaptations were likely to be more self-oriented as the interlocutor (Talker B) was always a young adult and less likely to be affected by noise. In this condition, we would expect older adults, and especially adults with age-related loss, to be more affected by the interference, and, as a result, to potentially differ from younger adults in their adaptations. In short, our research objective was to establish whether young and older adults varied in speech adaptation strategies used in two ecologically-valid challenging communicative conditions, and whether this varied as a function of their hearing status. The condition in which conversational partners were communicating in good listening conditions is also informative in terms of showing whether conversational speech in easy communicative conditions varied as a function of age and hearing status.

## II. METHOD

### A. Participants

A total of 98 participants (as ‘Talker A’) were recruited via adverts circulated within UCL as well as to hiking groups and to the University of the Third Age in London and surrounding counties. Of these 98 participants, 7 were excluded because of their language background, age or hearing level, 3 withdrew from the study after the first screening session and 5 did not manage to complete all sessions by the end of the study. The final sample consisted of 83 Southern British English adult talkers. They were divided into two age groups: ‘older adults’ (OA) between 64-84 years of age (*N*=57; 30 F; Female *M*=71;4, Male *M*=74;1 expressed in years;months) and ‘younger adults’ (YA) between 19-26 years of age (*N*=26, 15 F; Female *M*=21;11, Male *M*=20;11). All participants reported no history of speech or language impairments. OA participants were further subdivided into two groups according to their hearing status. OANH participants (*N*=27; 14F) had normal hearing defined as a mean pure-tone hearing threshold better than 20 dB HL calculated over octave frequencies between 250-4000

Hz while OAHL participants (*N*=30; 16F) had a mild acquired hearing loss defined as a mean threshold of 20-45 dB HL calculated over the same frequency range with a symmetrical downward slope in the high-frequency range typical for an age-related hearing loss (presbycusis) profile (See Figure 1 for better ear mean thresholds per group). The OAHL group had a greater mean age than the OANH group (see Table 1): t(55)=2.18, p=.033, effect size, Hedges' g=0.58, medium effect size. YA participants all had normal thresholds as defined using the same procedure and criteria as the OANH participants.

A further 83 participants were recruited as 'secondary' participants ('Talker B'), who acted as conversational partners with Talker A participants but whose speech was not analysed. These talkers were always young adults (aged between 18-30 years; *M*=21;0) of the same sex as Talker A. They all passed a hearing screen at 25 dB HL or better at octave frequencies between 250-8000 Hz in both ears. The rationale for having young adults as Talker B for all participants was that this would enable us to have a more homogeneous conversational-partner group in terms of hearing and cognitive status as the conversational partner's communication ability also affects the interaction. The participant pairs did not know each other prior to testing.

Informed written consent was obtained. Ethical approval was obtained from the University College London (UCL) Research Ethics Committee.

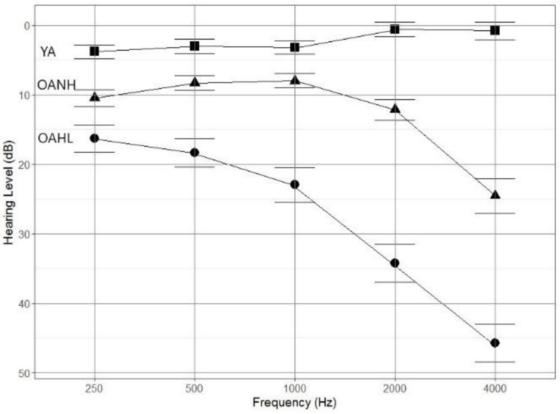

**FIGURE 1:** Mean pure tone thresholds in dB HL for frequencies between 0.250 and 4 kHz in the better ear according to participant group (YA – young adults, OANH – older adults with normal hearing, OAHL – older adults with hearing loss).

## B. Tasks

### *1. Experimental task: Diapix*

The experimental task, Diapix (van Engen et al., 2010) is a problem-solving 'spot the difference' picture task that has been used in a number of studies to elicit spontaneous speech within a communicative situation (e.g., Hazan and Baker, 2011, Hazan et al., 2016; McInerney and Walden, 2013; Sørensen et al., 2017). In addition to a reference condition where both talkers were interacting in good listening conditions (i.e. in quiet), this task was carried out in a number of different communicative conditions that were aimed at making it difficult for one or both talkers to understand each other in order to naturally elicit clear speech adaptations from Talker A. This study used a subset of these conditions: in addition to the NORM reference condition, the HLS condition where Talker B had a simulated hearing loss and a BAB2 condition where both talkers heard each other in a background of babble noise; both are described further below. For all these conditions, talkers could hear but not see each other (audio mode). Further recordings made at these recording sessions that are not reported here included a fourth condition, where only talker B heard babble noise (BAB1), and the same four conditions carried out in an audiovisual mode.

The diapixUK picture pairs developed by Baker and Hazan (2011) were used. Each participant was given a different version of the same picture-scene, and both were told that the pictures contained 12 differences that they had to find without having sight of their partner's picture. Talker A was told to lead the conversation and do most of the talking, whereas Talker B was mainly required to ask questions and make suggestions. They were instructed to start in the top left-hand corner of the picture and work clockwise; participants were given 10 minutes to find the differences after which the task was terminated. The task was terminated before this time limit if all differences had been found.

Before starting the first diapix task, all participants practised for 5-10 minutes (until they found at least 6 differences) using a different set of pictures while seated in the same room. This gave them the occasion to meet each other face-to-face and to get used to each other's voices. In the case of older participants, this also made them aware that they were interacting with a younger adult rather than someone of the same age. Whether this affected the modifications that they made to their speech in relation to ones they would make with peers or with friends cannot be ascertained without further testing. During the experimental conditions, Talker A participants were given a short description of what their interlocutor was hearing (e.g. told that they had a simulated severe-to-profound hearing loss for the HLS condition) but they did not experience the condition directly so they had to adjust their speaking style based on any experience of communicating with people with hearing loss and on the direct (e.g. 'speak more slowly') and indirect (e.g. 'I didn't get that') feedback received from Talker B.

**TABLE I** Means and standard deviations (in parentheses) for results of cognitive and sensory tests for the three talker groups (YA, OANH, OAHL). These include the mini-mental state examination (MMSE) with a maximum score of 23, the forward and backward digit spans, word association score, pure tone audiogram (between 0.250 and 4 kHz in better ear) in dB HL and SNR threshold in dB for the WiNics words in noise perception test.

| | Age (years; months) | MMSE | F. Digit Span | B. Digit Span | Word Assoc. | PTA Ave (dB HL) | WiNics Thres. (dB) |
|---|---|---|---|---|---|---|---|
| YA (N=26) | 21;6 (2) | | 12 (3) | 8 (3) | 74 (14) | 2.3 (3.6) | -8.0 (1.4) |
| OANH (N=30) | 71;1 (4) | 22 (2) | 12 (2) | 8 (2) | 71 (10) | 12.8 (4.6) | -6.5 (1.2) |
| OAHL (N=27) | 74;4 (5) | 22 (2) | 12 (2) | 7 (2) | 62 (12) | 27.7 (6.7) | -4.8 (2.8) |

In the NORM condition, both participants communicated without any additional interference. In HLS, Talker B heard Talker A via a real-time hearing loss simulator modelling a profound sensorineural loss at levels 40-50-60-90-90 dB HL at frequencies 250-500-1000-4000-8000 Hz; (HELPS; Zurek and Desloge, 2007). In the BAB2 condition, both speakers communicated in the same 8-talker babble noise, also heard through their headphones. This noise was scaled to 75 dB SPL to match the microphone target level using an Automated Gain Control (AGC) aimed at achieving a set SNR of 0 dB for all participants. The babble noise (from Cooke and Lu, 2010) had been created using recordings from 4 male and 4 female talkers speaking aloud while solving Sudoku puzzles with all pauses removed; voices were normalised to the same RMS level before adding. In the HLS and BAB2 conditions, the use of AGC for Talker A's voice meant that talking louder would not lead to a change in signal-to-noise ratio for Talker B although it could lead to a change in the spectral balance of Talker A's voice. Although this may not totally reflect what would occur in natural communication, participants in Talker A role were not aware of the use of AGC and so still used a strategy of talking more loudly. The use of AGC ensured that some ongoing communication difficulty remained without having to use excessive noise levels; it is also typically used in hearing aids.

### 2. *Cognitive and sensory function screening tasks*

In order to investigate the relationship between sensory and cognitive function and speech adaptations, a test battery of measures was carried out on the primary participants (Talker A).

*a.* *Sensory function.* To obtain hearing thresholds, pure tone audiograms were measured in both ears at octave frequencies between 250-8000 Hz (BSA, 2011). For some older adults, thresholds could not be obtained at 8000 Hz so mean thresholds were calculated over the 250-4000 Hz frequency range. Psychophysical tests of gap and frequency modulation (as described in Schoof and Rosen, 2014) were carried out but these are not reported due to the high number of participants for whom reliable thresholds could not be obtained. Thresholds for word intelligibility in background noise were obtained using the WiNics task (as described in Hazan et al., 2009) which was modelled on the coordinate response measure (Moore, 1981). Participants heard the following carrier phrase: "Show the dog where the [colour] [number] is," using 6 colours (black, red, white, blue, green and pink) and 8 digits (1,2,3,4,5,6,8,9). An adaptive procedure was used to vary the signal-to-noise ratio (SNR) and to track correct trials. The SNR corresponding to a target intelligibility level (79.4%) was calculated from the mean of the reversals excluding the first two.

b. *Cognitive function.* All older adults passed the shorter version of the MMSE dementia screening (>18 out of maximum 20). To obtain measures of short-term memory, a forward digit span (DSF) task was used. For working memory, we used the backward digit span (DSB) test that measures information storage and rehearsal. In the DSF/DSB tasks, the participant repeated auditorily-presented number sequences in the same or reverse order and were scored as correct or incorrect for each sequence (maximum scores, DSF=16, DSB=14).The efficiency of word search and retrieval from the stored lexicon was measured using the Word Association (WA) task in which participants had to say as many words as possible from a category in 60 seconds; the final score was the total number of items across the three categories. All tasks were from Semel, Wiig, & Secord (2006).

## C. Procedure

Each 'Talker A' participant took part in three 2-hour sessions. At the initial individual screening session, the sensory and cognitive tasks were run. They were then invited back within an interval of between 2 days and 2 months (median interval: 10 days) to complete two communicative sessions with a conversational partner (Talker B). The communicative tasks were carried out in adjacent sound-treated rooms, with a window connecting the two rooms so that the task could be carried out either with the two participants seeing each other (audiovisual mode) or not (audio mode). The participants wore Eagle G157b lapel microphones and Vic Firth SIH-1 headphones. The speech of each participant was recorded at a sampling rate of 44,100 Hz (16 bit) using an EMU 0404 USB audio interface and Adobe Audition and Rode NT1-A condenser microphones.

Recordings were controlled by the experimenter via a Dell laptop computer external to the two rooms using Matlab (2016) scripts. Two-channel recordings were made with the speech of each talker on a separate channel to facilitate the transcription and acoustic analysis stages. For each task, the pairs always started with a NORM condition and the adverse conditions were randomised within groups.

### D. Data processing

For all recordings, each channel was automatically transcribed using a cloud-based speech recognition system (https://www.speechmatics.com/). An in-house Matlab script was used to create a Praat textgrid from the JSON-formatted transcription produced by the system. These automated transcriptions and word-level alignments were then manually checked and corrected for errors.

The following acoustic characteristics were analysed using the audio signal recorded for Talker A in each condition. The analysis methods were identical to those used with a diapix corpus for child speech; further details about the analysis procedure are available in Hazan et al. (2016).

#### 1. Articulation rate and pause frequency

Articulation rate was calculated as the number of syllables produced by Talker A divided by the total duration (in seconds) of the speech regions. Syllable counts were calculated from the orthographic transcriptions of the spontaneous speech using the qdap package in R (Rinker, 2013), after exclusion of segments labelled as unfinished words, hesitations, fillers and agreements (e.g. ‘yeah’, ‘yup’, ‘err’, ‘hmm’). A normalised measure of pause rate was calculated as the number of within-talker pauses of over 300 ms in duration divided by the number of words (excluding fillers and interrupted words).

#### 2. Mean energy 1-3 kHz (ME1-3kHz)

Long-term average spectrum (LTAS) analyses were carried out using a Praat script. First, for each file, the intensity of all labelled speech segments was calculated and those above a set level (88 dB) were excluded for the LTAS calculations, as likely to be instances of shouting. The remaining speech segments were concatenated and the intensity of the resulting waveform scaled to a set level (75 dB). The signal was then band-pass filtered between 1 and 3 kHz and the mean intensity of the resulting waveform calculated to give a measure of the amount of energy present in the 1-3 kHz frequency range relative to the total energy in the spectrum. An increase in the relative energy in this mid-frequency band also reflects a reduction of spectral tilt, documented in speech produced with vocal effort (e.g., Glave and Rietveld, 1975; Sluijter and Van Heuven, 1996).

#### 3. Fundamental frequency (F0)

For each file, a Praat script was used to concatenate intervals which were not marked as silences, laughter, noise or breath intake and F0 calculations were done using the ‘pitch’ function in Praat, with a time step of 100 pitch values per second. A formula (De Looze and Hirst, 2008) was used to calculate ceiling and floor limits specific to each talker, as more successful at excluding rogue values than the use of default values for male and female talkers. For each talker, median F0 values were calculated per condition; values in Hertz were converted to semitones relative to 1 Hz. This conversion was to facilitate comparisons of median frequency changes in the challenging conditions across male and female talkers.[1]

#### 4. Vowel formant ranges

Formant ranges were obtained from the formant values measured for three corner vowels in content words: [iː], [æ] and [ɔː]. On average, 39 [iː], 31 [æ] and 22 [ɔː] vowel tokens were included in the calculations of vowel measures per talker for NORM, 37 [iː], 30 [æ] and 20 [ɔː] tokens for HLS and 34 [iː], 27 [æ] and 17 [ɔː] tokens for BAB2. Formant tracking algorithms by Burg in Praat were used to obtain the first two formant values (F1, F2) in content words with [i;], [æ], [ɔː] vowels over 50 ms in duration. Formant frequencies were measured from the midpoint of the vowel; the reference frequencies of the formant tracking algorithm was based on sex (M: 500 and 1485 Hz; F: 550 and 1650 Hz). First and second formant frequency estimates for [ɔː] were all checked manually as they were the most prone to tracking errors. For all three vowels, outliers were determined by using a ±2 *SD* cut-off criterion within each individual and were removed. Formant estimates were converted to Equivalent Rectangular Bandwidth (ERB) values to reduce the effect of anatomical differences due to sex and age. Although other vowel normalisation methods may more successfully reduce differences relating to anatomical differences, they require a full set of vowels per speaker or accurate measures for F3 (Flynn and Foulkes, 2011) and so could not be used with this corpus. Median F1/F2 ERB values were calculated per vowel per talker using the normalizeVowels function in phonR package (McCloy, 2015) that employs the Glasberg and Moore (1990) formula. For each talker, a measure of F1 range (in ERB) was derived by subtracting F1 [iː] from F1 [æ], giving an indication of the degree of differentiation in vowel height. F2 range (in ERB), reflecting front/back distinction, was obtained by subtracting F2 [ɔː] from F2 [iː].

#### 5. Task transaction measures

These measures were calculated to reflect the effectiveness of the interaction as frequent requests for repetitions or misunderstandings would lead to longer time needed to complete the task. The measure used was the time it took (in minutes) to find the first eight differences in the pictures (Time8).

## III RESULTS

### A. Profile of three listener groups

The three participant groups were compared in terms of their sensory and cognitive abilities (see Table 1) by looking at the effect of group on each of the measures described above in a series of univariate ANOVAs. A multivariate ANOVA was not used due to some missing data points for some tests.

#### 1. Sensory tests

As hearing thresholds were a selection criterion, the effect of group was significant [$F(2,79)=162.28$; $p<.001$] for pure tone average (PTA) in the better ear (0.25-4kHz); Least Significant Difference (LSD) post-hoc tests showed that all three groups differed from the others (OAHL>OANH>YA). For the WiNics test, a measure of speech perception in noise, the effect of group was significant [$F(2,79)=18.07$; $p<.001$] with the three groups differing significantly in their thresholds (YA<OANH<OAHL). In order to test that the difference in mean age between the two older adult groups was not causing this effect, data were recalculated after removing data for participants aged over 80 years in the OAHL group, which removed the significant age difference between OANH and OAHL groups (OANH=71 vs OAHL=72 years, $p=.424$). Group effects for the WiNics test remained significant at $p<.001$.

### 2. *Cognitive tests*

There was no significant group effect ($p>.604$) for forward digit span, backward digit span (run for all groups) and the mini-mental test (run with OA groups only). In the word association test, post-hoc tests showed that the OAHL group ($M=62$) scored more poorly than the OANH ($M=71$) and YA ($M=74$) groups, which did not differ significantly [$F(2, 79)=7.35$; $p=.001$]. These effects remained when calculations were done with age-matched OA groups, although with a revised effect for the word association test of $p=0.021$.

## B. Effects of talker group and sex on spontaneous speech in good communicative conditions.

First, the recordings obtained for Talker A for the NORM condition were analysed to investigate the effects of talker group (YA, OANH, OAHL) and talker sex (F, M) on the acoustic characteristics of spontaneous speech produced in good listening conditions (NORM, see Table 2 for descriptive data).

Analyses were based on linear mixed-effects modelling using the lme function in the nlme package for R (R Development Core Team, 2013). The best-fitting model for each individual analysis was chosen with hierarchical approaches, that is, adding one predictor at a time to a baseline model that includes no predictors other than the intercept. Talker Group (3: YA, OANH, OAHL) and Sex (2: F, M) were entered one by one as fixed effects and Participant was a random effect. Likelihood ratio tests were used to test the goodness-of-fit between the models. For the Talker Group comparisons, treatment coding was used, with the YA group chosen as the reference level (see Table A in supplemental materials).

### *1. Articulation rate*

There were significant main effects of talker group ($\chi^2(2)=17.20$, $p<.001$) and sex ($\chi^2(1)=9.46$, $p=.002$). Speech produced by both OA groups (OANH $M=3.57$; OAHL $M=3.62$) was articulated more slowly than speech produced by the YA group ($M=4.03$; $p<.001$) but there was no effect of hearing status within the OA group ($p=.711$). Men were faster talkers ($M=3.87$) than women ($M=3.60$).

### 2. *Normalised pause frequency*

There was a significant main effect of talker group ($\chi^2(2)=7.03$, $p=.030$) and a significant talker group by sex interaction ($\chi^2(1)=6.54$, $p=.038$). For male talkers, the three talker groups (YA, OANH, OAHL) did not differ ($p=.145$) but for female talkers there was a significant main effect of group ($p=.028$): female OANH talkers paused significantly less than female YA talkers but no other post-hoc paired comparisons were significant.

### 3. *Mean Energy (1-3 kHz)*

There were significant main effects of talker group ($\chi^2(2)=24.58$, $p=<.001$) and sex ($\chi^2(1)=10.03$, $p=.002$). There was more relative energy in the 1-3 kHz frequency range ($p<.001$) for the speech of YA ($M=65.52$) than OANH ($M=62.98$) and OAHL talkers ($M=61.59$) and there was a trend for OANH speech to have higher energy than OAHL speech ($p=.059$). There was also, as expected, more mid-frequency energy in the speech of female ($M=64.21$) than male ($M=62.26$) talkers.

### 4. *Median F0*

There were significant main effects of talker sex ($\chi^2(2)=108.89$, $p<.001$) and group by sex interaction ($\chi^2(1)=10.18$, $p=.006$): female talkers had a higher median F0 than males ($M=90.51$ vs 83.20 for male talkers); The interaction was due to OANH women having lower median F0 than both YA ($p=.003$) and OAHL female talkers ($p=.010$). No significant group differences were found for male talkers ($p=.312$).

### 5. *Vowel formant ranges*

For F1 range, there was a significant main effect of talker group ($\chi^2(2)=17.74$, $p=<.001$): OAHL talkers ($M=4.34$) had significantly smaller F1 range than YA talkers ($M=5.11$, $t=-4.26$, $p<.001$) and OANH talkers ($M=4.86$, $t=-2.97$, $p=.004$). YA and OANH talkers did not differ ($t=-1.45$, $p=.150$). For F2 range, there were significant main effects of talker group ($\chi^2(2)=21.25$, $p=<.001$) and sex ($\chi^2(1)=21.02$, $p=<.001$). YA talkers ($M=7.87$) had significantly smaller F2 range than OANH talkers ($M=9.02$, $t=4.53$, $p<.001$) and OAHL talkers ($M=8.87$, $t=3.84$, $p<.001$). OANH and OAHL did not differ ($t=-0.60$, $p=.551$).

In summary, some age-related differences were found in the spontaneous speech produced when conversing with an interlocutor in good communicative conditions. Overall, the speech of OA talkers was slower and had less relative energy in the 1-3 kHz range than the speech of YA talkers. The speech characteristics of OA talkers only differed significantly as a function of their hearing status for vowel F1 range; there was also a trend for OANH talkers to have more mid-frequency energy, and for OANH women to have lower median F0 than their OAHL peers. For female talkers, there were some differences between groups for certain measures: for example, young women paused more than

**TABLE II.** For the NORM condition, means and standard deviation measures of articulation rate (syllables/sec), normalised pause frequency (pauses per minute), mean relative energy between 1 and 3 kHz (dB), median F0 (semitones relative to 1 Hz) and vowel F1/F2 ranges (ERB).

| | | Artic. rate (Syll/sec) | | Pause Freq. (pause/min) | | ME1-3 kHz (dB) | | Median F0 (semitones) | | F1 range (ERB) | | F2 range (ERB) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | M | S.D. | M | S.D. | M | S.D | M | S.D | M | S.D | M | S.D |
| YA | F (N=15) | 3.9 | *0.2* | 0.1 | *0.04* | 65.9 | *2.8* | 91.4 | *1.7* | 5.0 | *0.3* | 8.1 | *0.9* |
| | M (N=11) | 4.2 | *0.8* | 0.08 | *0.03* | 65.0 | *2.1* | 82.2 | *2.3* | 5.2 | *0.7* | 7.6 | *0.9* |
| | All (N=26) | 4.0 | *0.6* | 0.09 | *0.04* | 65.5 | *2.5* | 87.4 | *5.0* | 5.1 | *0.5* | 7.9 | *0.9* |
| OANH | F (N=17) | 3.4 | *0.3* | 0.07 | *0.02* | 63.9 | *2.8* | 89.3 | *2.1* | 5.0 | *0.5* | 9.5 | *0.6* |
| | M (N=13) | 3.8 | *0.3* | 0.08 | *0.03* | 61.7 | *2.7* | 83.8 | *2.6* | 4.6 | *0.6* | 8.4 | *1.0* |
| | All (N=30) | 3.6 | *0.4* | 0.07 | *0.03* | 63.0 | *3.0* | 87.0 | *3.6* | 4.9 | *0.6* | 9.0 | *1.0* |
| OAHL | F (N=13) | 3.5 | *0.4* | 0.09 | *0.03* | 62.7 | *2.3* | 91.1 | *1.6* | 4.5 | *1.0* | 9.4 | *0.9* |
| | M (N=14) | 3.7 | *0.3* | 0.1 | *0.03* | 60.4 | *2.5* | 83.4 | *2.5* | 4.2 | *0.5* | 8.4 | *0.7* |
| | All (N=27) | 3.6 | *0.4* | 0.1 | *0.03* | 61.6 | *2.6* | 87.1 | *4.4* | 4.3 | *0.8* | 8.9 | *0.9* |

**TABLE III:** Articulation rate and pause frequency measures expressed as mean percentage change (and standard deviation) in the HLS and BAB2 conditions relative to the NORM condition, calculated per individual talker.

| | % change | Articulation rate | | | | Pause frequency | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | HLS | | BAB2 | | HLS | | BAB2 | |
| | | Mean | S.D. | Mean | S.D. | Mean | S.D. | Mean | S.D. |
| YA | F (N=15) | -13.2 | *5.0* | -3.9 | *7.1* | 7.6 | *22.6* | -19.1 | *16.7* |
| | M (N=11) | -9.9 | *13.6* | -3.1 | *13.7* | 17.2 | *27.4* | -13.9 | *32.3* |
| | All (N=26) | -11.7 | *9.6* | -3.6 | *10.1* | 11.8 | *24.8* | -16.9 | *23.9* |
| OANH | F (N=17) | -8.9 | *7.5* | -2.2 | *8.0* | 43.4 | *47.5* | 12.4 | *39.7* |
| | M (N=14) | -8.6 | *13.0* | -5.8 | *7.5* | 39.7 | *50.4* | 14.1 | *29.5* |
| | All (N=30) | -8.8 | *10.1* | -3.8 | *7.9* | 41.8 | *48.0* | 13.1 | *35.1* |
| OAHL | F (N=13) | -12.5 | *6.3* | -2.1 | *8.4* | 38.2 | *52.8* | -10.1 | *23.0* |
| | M (N=14) | -3.5 | *7.7* | -5.8 | *5.4* | 11.0 | *17.2* | -4.9 | *22.0* |
| | All (N=27) | -8.0 | *8.3* | -4.0 | *7.1* | 24.6 | *40.9* | -7.4 | *22.2* |

OANH talkers. However, no significant between-group differences were observed for male talkers.

### C. Effects of talker age and sex and task type on speaker adaptations made in challenging communicative conditions.

The next analysis focused on adaptations made by talkers when carrying out the diapix task in two conditions that reflect typical communication experiences of older adults: HLS, reflecting conversation with an interlocutor with hearing loss, and BAB2 reflecting communication in noisy environments. As the degree of adaptation is dependent on the degree of communication difficulty, which may vary across conditions, the focus is not on a comparison across the HLS and BAB2 conditions but rather on the effects of the age and hearing status of Talker A in each of these two communicative conditions relative to the normative condition (NORM).

We are focusing on the use of three potential strategies: slowing down one’s speech, increasing vocal effort and hyperarticulating. A reduction in rate can be marked in two complementary ways: by a reduction in articulation rate and an increase in pausing frequency. A strategy of increasing vocal effort would be marked by concomitant increases in mid-frequency energy, F0 and vowel F1. Hyperarticulation would be marked by a change in vowel formant ranges.

Statistical analyses of the data were based on linear mixed-effects modelling using the lmer function from the lme4 (Bates, Maechler, Bolker, & Walker, 2014) package for R (R Development Core Team, 2013). For each dependent variable (Time8, articulation rate, normalised pausing, ME 1–3 kHz, median F0, vowel formant ranges), we began with a saturated model that included interaction terms for all as fixed effects with random intercepts and slopes (Barr, Levy, Scheepers, & Tily, 2013). Due to nonconvergence, we simplified the models hierarchically from most complex to least complex, followed by forward entry of random slopes for the fixed effects that were retained in the initial backward elimination. The resulting converged models for all four variables included the following fixed effects: Condition (3: NORM, HLS, BAB2), Sex (2: female, male), Talker Group (3: YA, OANH, OAHL), and Participant as random effect but no random slopes. For talker group comparisons, we used treatment coding, with the YA group as the reference level, and for condition comparisons, with the NORM condition as reference level. For full statistical results and model summaries, see Tables 1-6 in supplemental materials.

We compared model residuals via chi-square tests ($\alpha = .05$) from the most complex models (containing the largest interaction term) to the least complex models (containing only single terms). If an interaction term was significant, we included all lower level effects involved in the interaction in the final model.

### 1. *Task difficulty*

First, we investigated the degree of communicative difficulty in the three conditions, as reflected in the Time8 transaction measure described above. The final model included fixed effects of talker group and condition and a random effect of participant (see Table 1 in Supplementary Materials): the OAHL group took longer to find the first 8 differences (*M*=6.14) than the OANH (*M*=5.29; *t*=2.87, *p*=.004) and YA (*M* =4.92; *p*<.001) groups who did not differ (*p*=.341). This effect remained when the two OA groups were matched for mean age, so is not due to the OAHL group being older on average. Also, Time8 was longer for the HLS (*M* =6.43) than the NORM (*M* =5.03; p<.001) or BAB2 (*M* =4.91; *t*=9.44, *p*<.001) conditions which did not differ (p=.334).

We examined whether, within the OA groups, Time8 was correlated with any of the sensory and cognitive measures. For example, talkers with elevated hearing thresholds could have greater difficulty interacting in the BAB2 condition or adults with poorer verbal fluency could use less effective repair strategies when put under time pressure in a challenging environment. Correlations were examined separately for the OANH and OAHL groups. No significant correlations were obtained with any of the sensory or cognitive measures.

Overall, the HLS condition introduced a level of difficulty for Talker B that was not totally resolved by talkers adapting their speech as they took longer to do the task than in other conditions. However, this was the case for both the YA and OA groups. For the BAB2 condition, transaction time did not differ from the NORM condition; this most likely reflects the fact that even though the babble noise might have affected some participants in Talker A role, at least one of the conversational partners, Talker B, who was always a young adult, may have been relatively unaffected by the interference from the babble noise. An alternative explanation is that the fact that Talker A was directly affected by the interference led them to make greater adaptations.

### 2. *Acoustic-phonetic adaptations*$^2$

In analysing the different acoustic measures, as mentioned above, the focus is on the effects of age and hearing status and their interactions on the use of each strategy. First, the statistical analyses are presented for each of the three strategies identified. Next, these analyses are summarised for each of the HLS and BAB2 conditions to highlight the effects of age and hearing status on speech adaptations in each of these communicative conditions. The descriptive data shown in Tables 3 and 4 are expressed as percentage change in the values of individual acoustic measures for the HLS and BAB2 condition relative to the NORM condition calculated per individual talker. This is to highlight the degree of adaptation across conditions rather than absolute values, and reflects the fact that the NORM condition is used as intercept in the statistical evaluations.

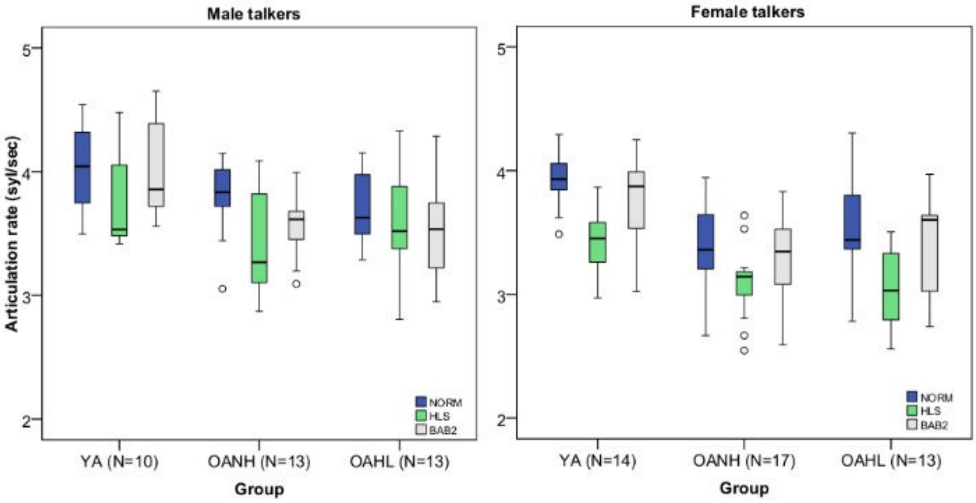

**FIGURE 2.** Mean articulation rates in the NORM, HLS and BAB2 conditions for the three talker groups. Data are split for male (left panel) and female (right panel) talkers.

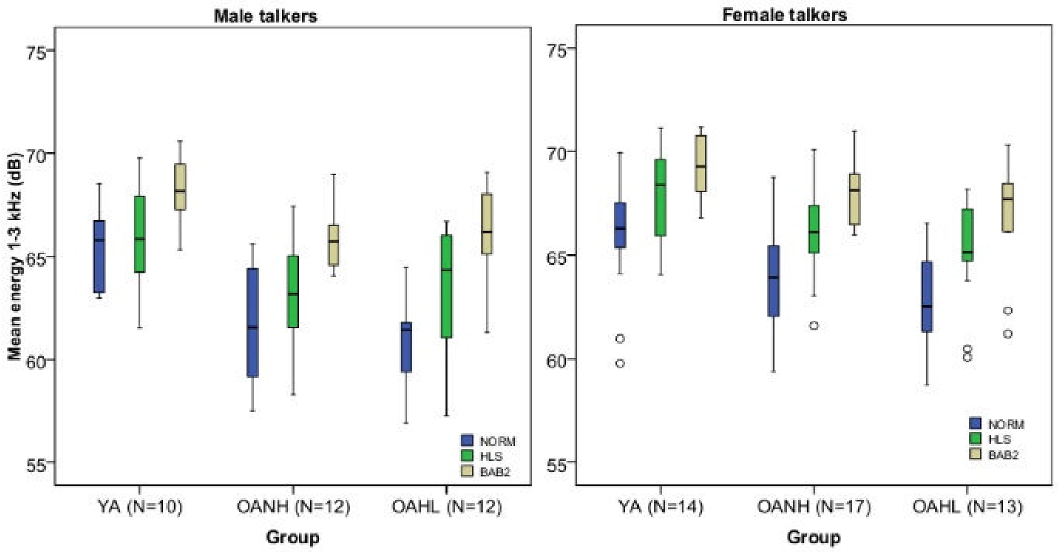

**FIGURE 3.** Measures of mean energy in the 1-3 kHz range in the NORM, HLS and BAB2 conditions for the three talker groups. Data are split for male (left panel) and female (right panel) talkers.

**TABLE IV***:* Data for ME1-3 kHz and median F0 measures expressed as mean percentage change (and standard deviation) in the HLS and BAB2 conditions relative to the measures in the NORM condition, calculated per individual speaker.

| | | ME1-3 kHz | | | | Median F0 | | | |
|---|---|---|---|---|---|---|---|---|---|
| | % change | HLS | | BAB2 | | HLS | | BAB2 | |
| | | Mean | S.D. | Mean | S.D. | Mean | S.D. | Mean | S.D. |
| YA | F (N=15) | 3.3 | *2.7* | 5.3 | *3.6* | 1.4 | *1.1* | 2.6 | *1.3* |
| | M (N=11) | 1.4 | *2.5* | 4.2 | *2.1* | 2.9 | *2.4* | 6.5 | *2* |
| | All (N=26) | 2.4 | *2.7* | 4.8 | *3.1* | 2.1 | *1.9* | 4.3 | *2.5* |
| OANH | F (N=17) | 3.7 | *2.8* | 6.7 | *2.8* | 1.6 | *1.9* | 5.1 | *1.7* |
| | M (N=13) | 2.3 | *2.2* | 7 | *4* | 2.7 | *2.4* | 6.4 | *1.9* |
| | All (N=30) | 3.1 | *2.6* | 6.8 | *3.3* | 2.2 | *2.2* | 5.7 | *1.9* |
| OAHL | F (N=13) | 4 | *2.2* | 6.9 | *2.5* | 2.3 | *1.8* | 5.1 | *2.1* |
| | M (N=14) | 4.4 | *4.6* | 9.1 | *3.6* | 3.5 | *3.3* | 6.9 | *2.2* |
| | All (N=27) | 4.2 | *3.5* | 8 | *3.2* | 2.9 | *2.9* | 6 | *2.3* |

a. *Articulation rate and pausing*. The descriptive data for articulation rate and pausing are shown in Table 3. The final model for articulation rate included the fixed effects of condition, talker group and sex, the interaction between condition and sex, and participant as a random factor (see Table 2 in Supplementary Materials). Relative to NORM, articulation rate was lower for HLS (*p*<.001) but only marginally so for BAB2 (*p*=.060); resulting in a difference in articulation rate between the two challenging conditions (*t*=-5.67, *p*<.001). YA talkers spoke at a significantly faster rate (*M*=3.81 syllables/second) than both OANH (*M*=3.42) and OAHL talkers (*M*=3.47; both *p*<.001), who did not differ significantly (*t*=0.39, *p*=.694). Male talkers had a faster articulation rate than female talkers in both NORM (*t*=-2.66, *p*=.009) and HLS (*t*=-4.54, *p*<.001) conditions but not in BAB2 condition (*t*=-1.76, *p*=.082). As can be seen in Figure 2, OAHL men tended not to slow down their articulation rate when interacting with their interlocutor with a simulated hearing loss whereas OAHL women did, but this interaction did not reach significance.

In summary, as regards group effects which are the focus of this analysis, all talker groups adopted a strategy of slowing down their speech when speaking to an interlocutor with a simulated hearing loss although there was a tendency for men to slow down their speech less than women. When directly exposed to babble noise, talkers (in Talker A role) reduced their articulation rate to a lesser degree than in HLS; this was the case for both younger and older adults and may be the consequence of communication generally being less impaired in this condition, or by a desire to try and complete the task quickly as communication in noise was unpleasant.

For normalised pause frequency, the final model included the main effects of condition, talker group, and an interaction between talker group and condition, and participant as a random factor (see Table 4 in Supplementary materials). There were significantly higher proportions of within-talker silent pauses in OANH speech (*M*=0.10) than in OAHL speech (*M*=0.08; *t*=2.22, *p*=.028) but neither OANH speech (*M*=0.09; *p*=.083) nor OAHL (*M*=0.08, *p*=.667) differed significantly from YA speech. As to the effect of communicative condition, talkers produced overall more silent pauses in HLS (*M*=0.10; *p*=.026) than in NORM (*M*=0.09) but produced less pauses in the BAB2 condition (*M*=0.08; *p*=.006); the difference between HLS and BAB2 was also significant (t=4.93, *p*<.001). However, the interaction between talker group and condition showed that whereas YA and OAHL talkers paused more in HLS and reduced their pause rate for BAB2, OANH showed a different pattern: relative to the NORM condition, OANH *increased* their pausing significantly for BAB2 (*p*=.004). In summary, pausing seems primarily being used as a strategy when speaking with an interlocutor with a simulated hearing impairment whereas talkers tend to reduce their pause frequency when communicating in background noise, although this was not the case for older adults with normal hearing.

b. *Increasing vocal effort:* Data expressed as percentage change relative to the NORM condition for the measures of ME1-3 kHz, median F0 and F1 range are presented in Tables 4 and 5. The final model for ME1-3 kHz included the fixed effects of condition, talker group, sex and interaction between talker group and condition (see Table 4 in Supplementary Materials), and participant as a random factor. ME1-3 kHz differed significantly as a function of communicative condition (See Figure 3): it was higher for the adverse conditions (BAB2: *M*=67.34 and HLS *M*=65.32) than for NORM (63.32; both *p*<.001), and also higher in BAB2 than in HLS (*t*=-4.75, *p*<.001). ME1-3 kHz was higher in YA talkers (*M*=67.11) than OANH (*M*=65.01) and OAHL (*M*=63.99) talkers; both comparisons *p*<.001. However, the hearing status of the OA talkers also made a difference, with higher ME1-3 kHz for OANH than OAHL talkers (*t*=-2.25, *p*=.025). For talker sex, as expected, female talkers (*M*=66.31) had significantly higher ME1-3 kHz than male talkers (*M*=64.13; *p*<.001). For the talker group and condition interaction, there was lower ME1-3 kHz for OANH than YA talkers when they were communicating in background noise (BAB2) but not in HLS, whereas OAHL had lower ME1-3 kHz than YA talkers in both adverse conditions.

In summary, all talker groups used strategies that boosted the amount of energy in the mid-frequency region of the spectrum when communication was made more difficult, and especially when directly exposed to noise, but the voice of OAHL talkers had less relative mid-frequency energy overall than that of YA and OANH groups.

For median F0 (see Table 4), because the interaction between talker group, condition and sex was significant, the final model included all main effects and interactions (see Table 6 in supplemental materials). Relative to NORM, men increased their median F0 more than women when speaking in the BAB2 condition (increase from 83.2 to 88.7 semitones for men versus 90.5 to 94.4 semitones for women). YA women showed the smallest increase across the two conditions (91.4 to 93.8 semitones).

Next, we examined evidence for a strategy of increasing vocal effort, which would be marked by correlations between an increase in mid-frequency energy, increase in F1 range and increase in median F0 as was found in Liénard and di Benedetto (1999) and Graezter et al. (2017) for example. Note that no change in vowel F2 was found by Liénard and Benedetto (1999) when vocal effort was increased. Table 6 displays Pearson's correlations between changes in these acoustic characteristics in HLS or BAB2 relative to the value in NORM for each talker group. Significant correlations between changes in these three measures were obtained for the OAHL group for both the HLS and BAB2 conditions. This is therefore consistent with talkers in this group increasing their vocal effort when attempting to clarify their speech. This was the case even in the HLS condition when they themselves were hearing without any additional interference. No such consistent correlations are obtained for the OANH group, where the only significant but weak correlation was between changes in F1 range and in ME1-3 kHz in the HLS condition. For the YA talker group, ME1-3 kHz was correlated with F1 range in the HLS condition and F2 range in the BAB2 condition but, here again, there are no significant concomitant changes in F1 range, median F0 and mean energy which would be indicative of increased vocal effort.

c. *Hyperarticulation strategy.* For F1 range, the final model included fixed effects of condition, talker group and sex, and an interaction between sex and condition (see Table 6 in Supplementary Materials and Figure 4). Talkers had larger F1 ranges in both adverse conditions (BAB2 *M*= 5.27 and HLS *M*=5.26) than in NORM (*M*=4.77); the two adverse conditions did not differ significantly (*t*=-0.68, *p*=.498). However, the interaction between sex and condition revealed that only female talkers significantly increased their F1 range in adverse conditions (see Table 4). OAHL talkers (*M*=4.68) had smaller F1 range than YA talkers (*M*=5.48) and OANH (*M*=5.16, *t*=-2.67, *p*=.008). YA talkers had marginally larger F1 range than OANH.

For F2 range, the final model included fixed effects of condition, talker group and sex, and an interaction between talker group and condition and sex and condition (see Table 7 in Supplementary Materials). YA talkers (*M*=8.00) had significantly smaller F2 ranges than OANH (*M*=8.64) and OAHL (*M*=8.38) talkers; the two older groups did not differ from each other (*t*=-0.33, *p*=.745). However, the significant interaction between talker group and condition shows that whereas YA talkers showed increased F2 range in HLS (NORM: *M*=7.87, HLS: *M*=8.51, BAB2: *M*=7.66), OAHL (NORM: *M*=8.87, HLS *M*=8.46, BAB2 *M*=7.87) and OANH (NORM: *M*=9.02, HLS: *M*=8.77, BAB2: *M*=8.13) *reduced* their F2 range for the adverse conditions. For condition and sex interaction, male talkers increased the F2 range in HLS condition (NORM: *M*=8.16, HLS: *M*=8.35, BAB2: *M*=7.72), whereas female talkers did not (NORM: *M*=9.02, HLS: *M*=8.79, BAB2: *M*=8.00).

It should be noted that because OANH and YA groups include more women than men whereas the OAHL group is balanced in talker sex, and because the normalisation method for vowel formant ranges may not totally account for physiological differences across speakers, some group effects may be influenced by group profiles in terms of talker sex.

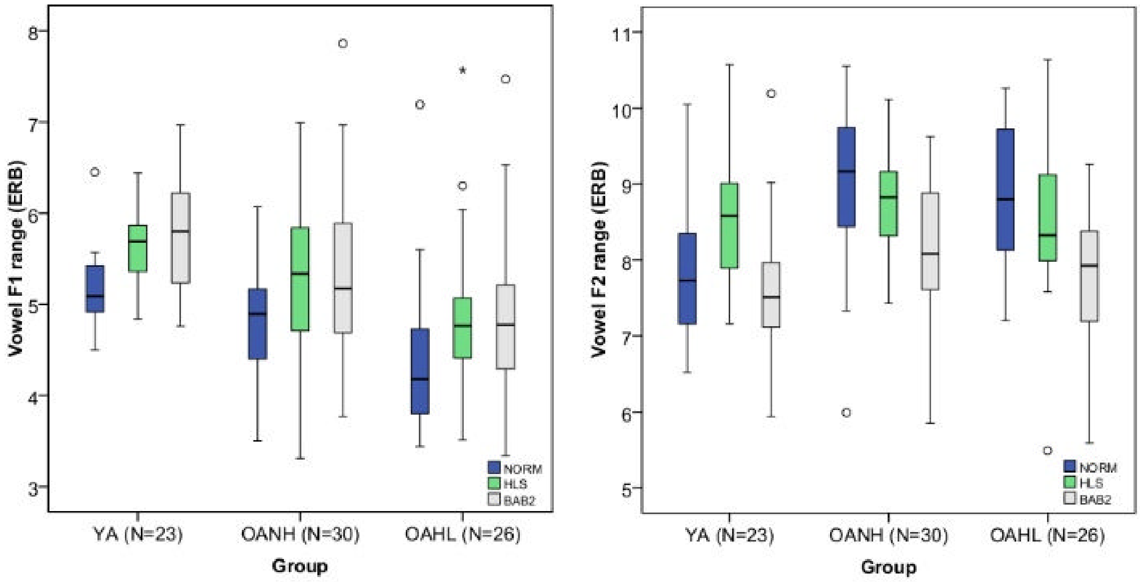

**FIGURE 4.** Measures of vowel F1 (left panel) and F2 (right panel) frequency range in the NORM, HLS and BAB2 conditions for the three talker groups.

**TABLE V***:* Data for F1 range measures expressed as mean percentage change (and standard deviation) in the HLS and BAB2 conditions relative to the measures in the NORM condition, calculated per individual speaker.

| | | F1 range | | | |
|---|---|---|---|---|---|
| | % change | HLS | | BAB2 | |
| | | Mean | S.D. | Mean | S.D. |
| YA | F (N=15) | 14.5 | *10.6* | 19.3 | *16.7* |
| | M (N=11) | 4.3 | *13.6* | 0.7 | *7.2* |
| | All (N=26) | 10 | *12.8* | 12 | *16.4* |
| OANH | F (N=17) | 14.2 | *9.6* | 14.2 | *12.1* |
| | M (N=13) | 3.4 | *8.6* | 2.2 | *11* |
| | All (N=30) | 9.5 | *10.6* | 9 | *13* |
| OAHL | F (N=13) | 16.8 | *14.1* | 14.7 | *14.6* |
| | M (N=14) | 9.3 | *16.6* | 8.5 | *15.5* |
| | All (N=27) | 13.1 | *15.6* | 11.5 | *15.1* |

**TABLE VI:** Pearson's correlations for the measures of percentage of change in the HLS (left panel) and BAB2 (right panel) conditions relative to the NORM condition for the following measures: vowel F1 range, vowel F2 range, mean energy 1-3 kHz (ME 1-3 kHz) and median F0 measures. Correlations are presented separately for each of the three listener groups: YA, OANH and OAHL. Correlations that are statistically significant at *p*=0.05 or better are highlighted in bold (*=*p*<0.05, **=*p*<0.01).

| **HLS** | | | | | **BAB2** | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | YA (N=25) | | | | | YA (N=25) | | | |
| | F1 range | F2 range | ME13kHz | Median F0 | | F1 range | F2 range | ME13kHz | Median F0 |
| F1 range | 1 | 0.146 | **.433*** | 0.077 | F1 range | 1 | 0.192 | 0.129 | -0.312 |
| F2 range | | 1 | -0.351 | -0.008 | F2 range | | 1 | **-.528**** | 0.092 |
| ME13kHz | | | 1 | -0.061 | ME13kHz | | | 1 | -0.157 |
| | OANH (N=30) | | | | | OANH (N=30) | | | |
| | F1 range | F2 range | ME13kHz | Median F0 | | F1 range | F2 range | ME13kHz | Median F0 |
| F1 range | 1 | -0.027 | **.389*** | .228 | F1 range | 1 | -0.017 | 0.117 | 0.046 |
| F2 range | | 1 | -0.04 | **0.368*** | F2 range | | 1 | 0.007 | -0.14 |
| ME13kHz | | | 1 | 0.215 | ME13kHz | | | 1 | 0.134 |
| | OAHL (N=26) | | | | | OAHL (N=26) | | | |
| | F1 range | F2 range | ME13kHz | Median F0 | | F1 range | F2 range | ME13kHz | Median F0 |
| F1 range | 1 | -0.284 | **.708**** | **.617**** | F1 range | 1 | -0.138 | **.395*** | **.456*** |
| F2 range | | 1 | -0.283 | -0.228 | F2 range | | 1 | -0.225 | -0.285 |
| ME13kHz | | | 1 | **.747**** | ME13kHz | | | 1 | **.727**** |

## IV DISCUSSION

This study focused on age and hearing-status related differences in the adaptations made by adult talkers when communication with a conversational partner was made more challenging. It also examined how these talker-related factors affected the acoustic characteristics of conversational speech produced in an interactive setting in good listening conditions.

In their conversational speech produced in good communicative conditions, YA and OA talkers varied in terms of their articulation rate, mid-frequency energy and F0 characteristics. The finding of a slower speech rate in older adults mirrors previous findings (Smith et al, 1987; Sadagopan and Smith, 2013; Smiljanić and Gilbert, 2017a). However, the finding of an age-related difference in mid-frequency energy does not concur with the acoustic analyses of the speech produced by younger and older adults in Smiljanić and Gilbert (2017a), as shown in their Table 1 for the 'conversational quiet' condition. This may potentially be due to the use of different elicitation approaches (spontaneous vs read speech) used in these studies. These differences in speech characteristics between older and younger adults in their habitual modes of communication have primarily been ascribed to age-related changes in the physiology of the vocal tract and vocal folds (Xue and Hao, 2003; Ramig et al., 2001).

The presence of a mild hearing loss in OA talkers did not lead to significant differences in the acoustic characteristics of their spontaneous speech produced in good communicative conditions relative to peers with normal hearing. This was the case even though the OAHL talkers, overall, took longer to do the task and so were less efficient at completing this simple task than their peers. However, these groups may have differed in other aspects of the interaction, such as the rate of disfluencies, for which further examination is warranted.

In the HLS condition, the adaptations made by Talker A, who was hearing normally, were purely interlocutor-oriented. In this condition, all three talker groups slowed down their articulation rate for their 'impaired' interlocutor; the use of this clear speech strategy did therefore not vary as a function of age and hearing status. All groups showed increased amount of relative energy in the mid-frequency range, which contains important acoustic cue information for speech perception, relative to their speech in the NORM condition, but only OAHL adults showed concomitant increases in F0, F1 and mid-frequency energy, which have been shown to be a marker of increased vocal effort (e.g. Liénard and di Benedetto, 1999; Graetzer, Bottalico and Hunter, 2017). Also, OAHL talkers showed a greater increase in median F0 in HLS relative to NORM than the OANH group which patterned with young adults.

The use by OAHL talkers of an increase in vocal effort as a clear speech strategy may be an effective one, even though, due to the use of AGC in the transmission channel, it would not have led to a louder signal for Talker B. This is because an increase in vocal effort would still result in a change in the spectral balance of the voice and boost the relative amount of energy in the mid-frequency region of the spectrum which is rich in acoustic cues. High energy in this region of the spectrum has been shown to be linked to increased speech intelligibility for speech presented in noise (e.g., Lu and Cooke, 2009; Hazan and Markham, 2004; Cooke et al., 2014). However, increasing vocal effort is also a strategy that may lead to vocal strain (Sundarrajan et al, 2017). Studies of the effect for older adults of using louder voice when communicating in noise have shown that they make greater abdominal muscle effort (Huber and Spruill, 2008). Over time, this could lead to fatigue and make it difficult to sustain such a strategy over a long time period. Whilst YA talkers used hyperarticulation, as shown by increased vowel F2 formant ranges, as an interlocutor-oriented clear speech strategy in this communicative situation, OA talkers as a group did not. In summary, while OANH adults generally patterned with YA adults in the interlocutor-oriented communication strategies that they used, OAHL adults varied from both young adults and their peers with normal hearing in the use of some strategies; their use of increased vocal effort, while effective, might lead to greater strain and fatigue if sustained.

As suggested by the Framework for Understanding Effortful Listening (FUEL) model (Pichora-Fuller et al., 2016), given a finite level of cognitive capacity, listening effort is affected by the cognitive demand of the task being undertaken as well as the motivation to complete the task. Even though this framework was formulated within the context of speech understanding, it also seems appropriate when considering speaking effort in the context of a communicative task. The motivation shown by OA groups was not evaluated directly but these participants had volunteered because they were interested in scientific studies and in finding out more about speech communication. As regards cognitive demands, several factors might have contributed to these being greater for OA groups relative to younger adults. The problem-solving task itself was cognitively demanding as it involved searching for differences, planning a strategy and formulating questions. The difficulties experienced by Talker B required Talker A to formulate repair strategies, which would also involve additional cognitive processing. Also, speech perception is known to degrade in older adults even for speech in quiet (e.g., Dubno, 2015) so it was likely to be affected for older adults in this HLS condition, regardless of their hearing status.

There are factors that could contribute to an even greater cognitive load being experienced by OAHL talkers, relative to their OANH peers, while interacting with their 'impaired' conversational partner. In trying to understand their conversational partner, OAHL talkers may have been affected by the lack of visual cues as they have been shown to be more reliant than OANH peers on visual information for speech perception (Davis et al., 2017). Also, the word-association test results, which showed that OAHL talkers had significantly greater difficulty for lexical retrieval under time pressure than OANH or YA talkers, suggest that they may have found it more difficult to formulate repair strategies such as rephrasing or expanding under time pressure. Note also that verbal fluency has been shown to have an association with measures of social loneliness (Schnittger et al., 2012). The combination of these additional cognitive demands may account for the more limited clear speech strategies used by OAHL talkers relative to their hearing peers and younger adults, although OAHL have been shown to expend more mental effort in listening tasks (e.g. Hornsby et al, 2016; Peele and Wingfield, 2016; Wayne and Johnsrude, 2015). The Ease of Language Understanding (ELU) model (Rönnberg et al., 2013) suggests that older adults with mild HL are sometimes unable to automatically access phonological representations and so require explicit and deliberative executive and working memory processes, such as inference-making, semantic integration, switching of attention, storing information or inhibition of irrelevant information, to do so. The fact that we did not find any relation between hearing status and our background cognitive measures may have been due to our choice of working memory measures. Rőnnberg et al. (2013) report that digit spans mainly tap into storage functions for short-memory and are not good predictors of language comprehension; they suggest that a reading span test is a more reliable measure of working memory capacity.

In the condition in which both conversational partners were communicating in noise, age- and hearing-status effects were expected in the speech adaptations made by Talker A. Background noise is known to have a particularly deleterious effect on speech perception for older adults (e.g., Helfer and Wilber,1990; Helfer and Huntley, 1991; Divenyi, Stark and Haupt, 2005), and both OA groups showed elevated speech perception in noise thresholds in background tests. Our expectation was that, for our older adults, the presence of noise in the background would lead to changes more consistent with Lombard speech although the noise levels were not as severe as in studies such as Garnier and Heinrich (2014). Studies of Lombard speech have typically shown acoustic changes consistent with an increase in vocal effort: a raised F0, F1 and shift in the spectral balance of the long-term spectrum (Titze and Sundberg, 1992; Sundberg and Nordenberg, 2006). In the BAB2 condition, all talker groups used strategies that boosted the amount of energy present in the mid-frequency region of the spectrum and all groups raised their F0 but only the OAHL group showed the significant correlation between increases in F0, F1 and mid-frequency energy that are consistent with the increase in vocal effort typically seen in Lombard speech. Neither of the OA groups showed evidence of vowel hyperarticulation, unlike the YA group, and all talker groups showed only a marginal reduction in articulation rate in this condition. In this communicative condition too, therefore, although many of the strategies were common to all talker groups, age- and hearing-status related differences occurred, with OAHL talkers producing changes which were again consistent with an increase in vocal effort and with experiencing greater difficulty with the task.

In conclusion, older adults used many clear speech strategies previously documented for young adults. However, hearing status had an effect on the use of some clear speech strategies even if it did not change the acoustic characteristics of speech when produced in favourable communicative conditions. It is worth noting that this effect of hearing status was found for healthy and active adults with a mild degree of presbycusis. Background questionnaire information revealed that although some of our older participants owned hearing aids, all but one reported either not using them at all or only using them very occasionally in specific situations. Despite this, there were measurable effects of this mild hearing loss on clear speech strategies, as shown above, and also on speech perception in noise, as shown by elevated SNRs in the WiNics test for the OAHL group relative to their OANH peers. The OAHL group also performed more poorly than OANH peers when participants were put under time pressure as shown by lower word association scores and higher task transaction times, even in the NORM condition when there were no communication barriers. This confirms that a mild age-related hearing loss has consequences for effective speech communication even when the older adult is not directly exposed to noisy conditions. It is well known that the individuals with hearing loss need to allocate greater cognitive resources to maintain good communication than individuals with normal hearing (e.g. McCoy et al., 2005). This can lead to mental fatigue for sustained tasks (Hornsby, 2013) and the strategy of increased vocal effort that was particularly adopted by the OAHL adults in our talker group are likely to further increase vocal strain and mental fatigue. Our findings confirm the view (e.g. Peelle and Wingfield, 2016) that amplification of older adults' hearing, even for those with mild age-related hearing loss, may be beneficial in reducing at least some of the sources of additional cognitive load for this population. An increased understanding of communication strategies that are effective at making speech clearer in certain environments might also help in advising older adults on how to be effective in a way that avoids vocal strain. For example, it can be suggested that slowing down their speech, introducing pauses and back-channelling to signal understanding is preferable to shouting. This study also showed that the spectral balance of the voice has a big effect on how well one can be understood in noisy conditions; this spectral balance is to a degree dependent on how strongly and regularly vocal chords are vibrating, so instructing older adults to maintain the health of their voice by not smoking, keeping well hydrated and avoiding shouting could also be beneficial.

## V ACKNOWLEDGMENTS

This work was supported by the UK Economic and Social Research Council (grant number ES/L007002/1). We acknowledge the contribution of Jay Desloges (Sensimetrics Corporation) to the implementation of the hearing loss simulation procedure.

## VI ENDNOTE

1. The approach of converting F0 to semitones relative to 1 Hz does not fully normalise for sex differences but was used here so that our data would be more easily comparable to related studies with child data (Hazan et al., 2016). To check that the use of a F0 scaling approach which was more successful in normalising for sex differences would not significantly change these findings, F0 values in Hz were converted to values in semitones relative to a base frequency

for each talker ($7^{th}$ percentile of the F0 distribution in the NORM condition) using the formula in Yuan and Liberman (2014). For the main statistical effects of group on median F0 in NORM, using this scaling only an effect of group remained, ($\chi2(2)=6.52$, *p*=.039): YA talkers had higher F0 than both older groups (*t*=-2.52, *p*=.014) which did not differ (*p*=.417). When statistics were carried out to look at the effect of condition, all interactions remained in the final model except for the sex*group interaction (*p*=.863) In terms of the correlations reported in Table 6, significant correlations remained the same for the HLS condition; for BAB2, the only difference is that changes in vowel F1 range and in median F0 were not correlated using the measure of F0 in semitones re base frequency whereas they were when F0 was scaled in semitones relative to 1 Hz.

See supplementary material at [URL] for full statistical results and model summaries.

2. In order to evaluate whether any difference between the adaptation strategies used by the OANH and OAHL groups may be due to a difference in mean age between the groups rather than their hearing status, statistics were also run after the age difference had been removed. The only statistical effect that differed when OA groups were age-matched was the correlation between changes in median F0 and change in F1 in the BAB2 condition was weakened for the OAHL group (*p*=.065).

## VII REFERENCES

Anderson, A. H., Bader, M., Bard, E. G. , Boyle, E. H. , Doherty, G. M., Garrod, S. C., Isard, S. D., Kowtko, J. C., McAllister, J. M., Miller, J. , Sotillo, C. F. , Thompson, H. S. , Weinert, R. (1991). The HCRC Map Task Corpus. *Language and Speech, 34*, 351-366.

Baken, R. J. (2005). The aged voice: A new hypothesis. *Journal of Voice*, *19*, 317-325.

Baker, R., & Hazan, V. (2011). DiapixUK: task materials for the elicitation of multiple spontaneous speech dialogs. *Behavior Research Methods*, *43*(3), 761-770. https://doi.org/10.3758/s13428-011-0075-y

Barr, D.J., Levy, R., Scheepers, C. and Tily, H.J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68* (3), 255-278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1 – 48. https://doi.org/10.18637/jss.v067.i01

Bilodeau-Mercure, M., & Tremblay, P. (2016). Age Differences in Sequential Speech Production: Articulatory and Physiological Factors. *Journal of the American Geriatrics Society, 64*(11), 177-182. https://doi.org/10.1111/jgs.14491

Bóna, J. (2014). Temporal characteristics of speech: The effect of age and speech style. *The Journal of the Acoustical Society of America, 136*, EL 116-121. https://doi.org/10.1121/1.4885482

Cooke, M., King, S., Garnier, M., Aubanel, V. (2014). The listening talker: a review of human and algorithmic context-induced modifications of speech. *Computer Speech and Language, 28*(2), 543–571. https://doi.org/10.1016/j.csl.2013.08.003

Cooke, M., Lu, Y. (2010). Spectral and temporal changes to speech produced in the presence of energetic and informational maskers. *The Journal of the Acoustical Society of America*, 128, 2059-2069. https://doi.org/10.1121/1.3478775

Davis, C., Kim, J., Tuomainen, O., & Hazan, V. (2017). The effect of age and hearing loss on partner-directed gaze in a communicative task. *Proceedings of AVSP 2017*, Stockholm, Sweden, 25-26 August 2017.

De Looze, C. & Hirst, D.J. (2008). Detecting Key and Range for the Automatic Modelling and Coding of Intonation, Actes de Speech Prosody 2008 Conference, Campinas, Brazil, 135-138.

Divenyi, P. L., Stark, P. B., & Haupt, K. M., (2005). Decline of speech understanding and auditory thresholds in the elderly. *The Journal of the Acoustical Society of America, 118*, 1089-1100.

Dubno, J. (2015) Speech recognition across the life span: longitudinal changes from middle-age to older adults**.** *American Journal of Audiology, 24*(2), 84 -87.

Duchin, S. W., and Mysak, E. D. (1987). Disfluency and rate characteristics of young adult, middle-aged, and older males. *Journal of Communication Disorders, 20*, 245–257.

Ferguson, S. H., & Kewley-Port, D. (2007). Talker differences in clear and conversational speech: Acoustic characteristics of vowels. *Journal of Speech, Language, and Hearing Research, 50*, 1241-1255. https://doi.org/10.1044/1092-4388

Flynn, N. E. J., & Foulkes, P. (2011). Comparing vowel formant normalization methods. In Lee, W. S. & Zee, E. (eds.), *Proceedings of the 17th ICPhS*. Hong Kong: The City University of Hong Kong, 683–686.

Garnier, M., & Heinrich, N. (2014). Speaking in noise: How does the Lombard effect improve acoustic contrasts between speech and ambient noise? *Computer Speech and Language, 28*(2), 580-597. https://doi.org/10.1016/j.csl.2013.07.005

Glasberg, B. R., & Moore, B. C. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research*, *47*(1–2), 103-138.

Glave, R. D., and Rietveld, A. C. M. (1975). Is the effort dependence of speech loudness explicable on the basis of acoustical cues? *The Journal of the Acoustical Society of America*, *58*, 875–879.

Gordon-Salant, S. (2005). Hearing loss and aging: new research findings and clinical implications. *Journal of Rehabilitation Research and Development, 42*, 9–24.

Gordon-Salant, S. (2014). Aging, hearing loss and speech understanding: Stop shouting, I can't understand you. In Popper, A. N., and Fay, R. R. (Eds.). Perspectives on Auditory Research. Springer-Verlag, New York, pp. 211-228.

Gould, O.N. and Shaleen, L. (1999). Collaboration with diverse partners: how older women adapt their speech. Journal of Language and Social Psychology, 18, 395-418.

Goy, H., Fernandes, D. N., Pichora-Fuller, M. K., & Van Lieshout, P. (2013). Normative voice data for younger and older adults. *Journal of Voice, 27(*5*),* 545-555. https://doi.org/10.1016/j.jvoice.2013.03.002

Graetzer, S., Bottalico, P. and Hunter, E.J. (2017). Speech produced in noise: Relationship between listening difficulty and acoustic and durational parameters. *The Journal of the Acoustical Society of America, 142(*2*),* 974-983. https://doi.org/10.1121/1.4997906

Hazan, V., & Baker, R. (2011). Acoustic–phonetic characteristics of speech produced with communicative intent to counter adverse listening conditions. *The Journal of the Acoustical Society of America, 130(*4), 2139–2152. https://doi.org/10.1121/1.3623753

Hazan, V. and Markham, D (2004) Acoustic-phonetic correlates of talker intelligibility for adults and children. *The Journal of the Acoustical Society of America, 116*, 3108-3118.

Hazan, V., Messaoud-Galusi, S., Rosen, S., Nouwens, S., Shakespeare, B. (2009). Speech perception abilities of adults with dyslexia: is there any evidence for a true deficit? *Journal of Speech, Hearing and Language Research, 52*(6), 1510-1529. https://doi.org/10.1044/1092-4388(2009/08-0220)

Hazan, V. L., Tuomainen, O., & Pettinato, M. (2016). Suprasegmental Characteristics of Spontaneous Speech produced in Good and Challenging Communicative Conditions by Talkers aged 9 to 14 years old. *Journal of Speech, Language, and Hearing Research, 59*, S1596-S1607. https://doi.org/10.1044/2016_JSLHR-S-15-0046.

Helfer, K., & Huntley, R. (1991). Aging and consonant errors in reverberation and noise. *The Journal of the Acoustical Society of America, 90*, 1786–1796.

Helfer, K. S., & Wilber, L. A. (1990). Hearing loss, aging, and speech perception in reverberation and noise. *Journal of Speech and Hearing Research, 33,* 149-155.

Hollien, H., & Shipp, T. (1972). Speaking fundamental frequency and chronologic age in males. *Journal of Speech and Hearing Research, 15,* 155-159.

Hooper, C. R., & Cralidis, A. (2009). Normal changes in the speech of older adults. You've still got what it takes, it just takes a little longer! *Perspectives on Gerontology, 14,* 47-56. https://doi.org/10.1044/gero14.2.47

Hornsby, B. W. Y. (2013). The Effects of Hearing Aid Use on Listening Effort and Mental Fatigue Associated with Sustained Speech Processing Demands. *Ear and Hearing, 34(*5*),* 523-534. https://doi.org/10.1097/AUD.0b013e31828003d8

Hornsby, B. W., Naylor, G., & Bess, F. H. (2016). A taxonomy of fatigue concepts and their relation to hearing loss. *Ear and Hearing, 37*, 136S-144S. https://10.1097/AUD.0000000000000289

Huber, J. E., & Spruill, J. (2008). Age-Related Changes to Speech Breathing with Increased Vocal Loudness. *Journal of Speech, Language, and Hearing Research, 51*(3), 651-668. https://10.1044/1092-4388(2008/047)

Jacewicz, E., Fox, R. A., and Wei, L. (2010). Between-speaker and within-speaker variation in speech tempo of American English. *The Journal of the Acoustical Society of America, 128,* 839-850. https://doi.org/10.1121/1.3459842

Kemper, S., Vandeputte, D., Rice, K., Cheung, H. and Gubarchuk, J. (1995). Speech adjustments to aging during a referential communication task. *Journal of Language and Social Psychology 14*(1-2), 40-59. https://doi.org/10.1177/0261927X95141003

Krause, J. C. & Braida, L. D. (2004). Acoustic properties of naturally produced clear speech at normal speaking rates. *The Journal of the Acoustical Society of America, 115,* 362-378. https://doi.org/10.1121/1.1635842

Liénard J. S., & Di Benedetto, M. G. (1999). Effect of vocal effort on spectral properties of vowels. *The Journal of the Acoustical Society of America, 106,* 411-422.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modelling* (pp. 403–439). Dordrecht, the Netherlands: Kluwer Academic.

Linville, S. E., & Rens, J. (2001). Vocal Tract Resonance Analysis of Aging Voice Using Long-Term Average Spectra. *Journal of Voice, 15*(3), 323-330. https://doi.org/10.1016/S0892-1997(01)00034-0

Lu, Y., Cooke, M. (2009). The contribution of changes in F0 and spectral tilt to increased intelligibility of speech produced in noise. *Speech Communication, 51,* 1253-1262. https://doi.org/10.1016/j.specom.2009.07.002

McCloy, D. (2015*). phonR: Tools for phoneticians and phonologists*. R package version 1.0-0. ⟨http://drammock.github.io/phonR/⟩.

McCoy, S.L., Tun, P.A., Cox, L.C., Colangelo, M., Stewart, R.A., Wingfield, A. (2005). Hearing loss and perceptual effort: Downstream effects on older adults' memory for speech. *Quarterly Journal of Experimental Psychology, 58,* 22–33. https://doi.org/10.1080/02724980443000151

McInerney, M., & Walden, P. (2013). Evaluating the use of an assistive listening device for communication efficiency using the Diapix task: A pilot study. *Folia Phoniatrica Logopedia, 65,* 25-31. https://doi.org/10.1159/000350490.

Moore, T. J. (1981). Voice communication jamming research. AGARD Conference Proceedings 311: Aural Communication in Aviation AGARD, Neuilly-Sur-Seine, France, 2:1–2:6.

Peelle, J. E., & Wingfield, A. (2016). The neural consequences of age-related hearing loss. *Trends in neurosciences, 39,* 486-497. https://doi.org/10.1016/j.tins.2016.05.001

Picheny M. A., Durlach N. I., & Braida L. D. (1986). Speaking clearly for the hard-of-hearing II: Acoustic characteristics of clear and conversational speech. *Journal of Speech and Hearing Research. 29*, 434-446. https://doi.org/10.1044/jshr.2904.434

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., Lemke, U., Lunner, T., Matthen, M., Mackersie, C. L., Naylor, G., Phillips, N. A., Richter, M., Rudner, M., Sommers, M. S., Tremblay, K. L., & Wingfield, A. (2016). Hearing Impairment and Cognitive Energy: The Framework for Understanding Effortful Listening (FUEL). *Ear and Hearing, 37,* 5S-27S. https://doi.org/10.1097/AUD.0000000000000312

Ramig, L. O., Gray, S., Baker, K., Corbin-Lewis, K., Buder, E., Luschei, E., Coon, H., & Smith, M. (2001). The aging voice: a review, treatment data and familial and genetic perspectives. *Folia Phoniatrica et Logopaedica, 53,* 252-265. https://doi.org/10.1159/000052680

Rinker, T. W. (2013). qdap: Quantitative discourse analysis package. version 1.3.1. Buffalo, NY: University at Buffalo. http://github.com/trinker/qdap

Reubold, U., Harrington, J., & Kleber, F. (2010). Vocal aging effects on F0 and the first formant: A longitudinal analysis in adult speakers. *Speech Communication, 52,* 638-651. https://doi.org/10.1016/j.specom.2010.02.012

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., Dahlström, ö., Signoret, C., Stenfelt, S., Pichora-Fuller, M.K. & Rudner, M. (2013). The Ease of Language Understanding (ELU) model: Theoretical, empirical, and clinical advances. Frontiers in Systems Neuroscience, 7, 31. https://doi.org/10.3389/fnsys.2013.00031

Sadagopan, N., & Smith, A. (2013). Novel nonword learning in older speakers. *Journal of Speech, Language, and Hearing Research, 56*, 1552-1566.

Sataloff, R. T., Rosen, D. C., Hawkshaw, M., & Spiegel, J. R. (1997). The Three Ages of Voice: The Aging Adult Voice. *Journal of Voice, 11,* 156-160.

Scarborough, R. & Zellou, G. (2013). Clarity in communication: "Clear" speech authenticity and lexical neighborhood density effects in speech production and perception. *The Journal of the Acoustical Society of America, 134,* 3793-3807. https://doi.org/10.1016/10.1121/1.4824120

Schoof, T., and Rosen, S. (2014). The role of auditory and cognitive factors in understanding speech in noise by normal-hearing older adults. *Frontiers in Aging Neuroscience, 6,*307. https://doi.org/10.3389/fnagi.2014.00307

Schum, D. J. (1996). Intelligibility of Clear and Conversational Speech of Young and Elderly Talkers. *Journal of the American Academy of Audiology, 7*, 212-218.

Schnittger, R. I., Wherton, J., Prendergast, D., & Lawlor, B. A. (2012). Risk factors and mediating pathways of loneliness and social support in community-dwelling older adults. *Aging & Mental Health, 16*(3), 335-346. https://doi.org/10.1080/13607863.2011.629092

Semel, E., Wiig, E. H., & Secord, W. A. (2006). Clinical evaluation of language fundamentals, fourth edition—Screening test (CELF-4 screening test). Toronto, Canada: The Psychological Corporation/A Harcourt Assessment Company.

Sluijter, A. M. C., and Heuven, V. J. van (1996). Spectral balance as an acoustic correlate of linguistic stress. *The Journal of the Acoustical Society of America*, 100, 2471–2485. https://doi.org/10.1121/1.417994

Smiljanić, R. (2013). Can older adults enhance the intelligibility of their speech? *The Journal of the Acoustical Society of America, 133*(2), EL 129-135. https://doi.org/10.1121/1.4776774

Smiljanić, R., & Bradlow. A. R. (2005). Production and perception of clear speech in Croatian and English. *The Journal of the Acoustical Society of America, 118,* 1677-1688. https://doi.org/10.1121/1.2000788

Smiljanić, R., & Bradlow, A. R. (2009). Speaking and hearing clearly: Talker and listener factors in speaking style changes. *Language and Linguistics Compass, 3*, 236-264. https://doi.org/10.1111/j.1749-818X.2008.00112.x

Smiljanić, R., & Gilbert, R. C. (2017a). Acoustics of clear and noise-adapted speech in children, young, and older adults. *Journal of Speech, Language, and Hearing Research, 60*, 3081-3096. https://doi.org/10.1044/2017_JSLHR-S-16-0130

Smiljanić, R., & Gilbert, R. C. (2017b). Intelligibility of noise-adapted and clear speech in child, young adult, and older adult talker. Journal of Speech, Language, and Hearing Research, 60, 3069-3080. https://doi.org/10.1044/2017_JSLHR-S-16-0165

Smith, B. L., & Wasowicz, J., & Preston, J. (1987). Temporal characteristics of the speech of normal elderly adults. *Journal of Speech, Language, and Hearing Research, 30*, 522-529. https://doi.org/10.1044/jshr.3004.522

Sørensen, J., Fereczkowski, M., & MacDonald, E. N. (2017). The effect of noise and second language on turn taking in task-oriented dialog. *The Journal of the Acoustical Society of America 141*, 3520. https://doi.org/10.1121/1.4987409

Spaniol, J., Voss, A., Bowen, H. J., & Grady, C. L. (2011). Motivational incentives modulate age differences in visual perception. *Psychology and Aging, 26,* 932-939. https://doi.org/10.1037/a0023297

Sundarrajan, A., Fujiki, R. B., Loerch, S. E, Venkatraman, A., Sivasankar, M., & Preeti, S. (2017). Vocal loading and environmental humidity effects in older adults. *Journal of Voice, 31*, 707-713. https://doi.org/10.1016/j.jvoice.2017.02.002

Sundberg, J., Nordenberg, M. (2006). Effects of vocal loudness variation on spectrum balance as reflected by the alpha measure of long-term-average spectra of speech. *The Journal of the Acoustical Society of America, 120*, 453–457.

Titze I. R., and Sundberg J. (1992). Vocal intensity in speakers and singers. *The Journal of the Acoustical Society of America, 91*, 2936-2946.

Torre, P., & Barlow, J. A. (2009). Age-related changes in acoustic characteristics of adult speech. *Journal of Communication Disorders, 42*, 324-333. https://doi.org/10.1016/j.jcomdis.2009.03.001

Traunmuller, H., & Eriksson, A. (2000). Acoustic effects of variation in vocal effort by men, women, and children. *The Journal of the Acoustical Society of America, 107,* 3438-3451. https://doi.org/10.1121/1.429414

Uchanski, R. M., Choi, S., Braida, L. D., Reed, C. M., & Durlach, N. I. (1996). Speaking clearly for the hard of hearing IV: Further studies of the role of speaking rate. *Journal of Speech, Language, and Hearing Research. 39,* 494-509. https://doi.org/10.1044/jshr.3903.494

Van Engen, K.J., Baese-Berk, M., Baker, R.E., Choi, A., Kim, M., Bradlow, A.R. (2010). The Wildcat Corpus of Native- and Foreign-Accented English: communicative efficiency across conversational dyads with varying language alignment profiles. *Language and Speech 53,* 510-540. https://doi.org/10.1177/0023830910372495

Wayne, R. V., & Johnsrude, I. S. (2015). A review of causal mechanisms underlying the link between age-related hearing loss and cognitive decline. *Ageing research reviews, 23,* 154-166. https://doi.org/10.1016/j.arr.2015.06.002

Xue, S. A., & Hao, G. J. (2003). Changes in the human vocal tract due to aging and the acoustic correlates of speech production: A pilot study. *Journal of Speech, Language, and Hearing Research 46*, 689-701. https://doi.org/10.1044/1092-4388(2003/054)

Yuan, J., & Liberman, M. (2014). F0 declination in English and Mandarin Broadcast News Speech. *Speech Communication 65*, 67-74. https://doi.org/10.1016/j.specom.2014.06.001

Zhang, X., Fung, H. H., Stanley, J. T., Isaacowitz, D. M., & Ho, M. Y. (2013). Perspective taking in older age revisited: A motivational perspective. *Developmental Psychology, 49*, 1848-1858. https://doi.org/10.1037/a0031211

Zurek, P. M., & Desloge, J. G. (2007). Hearing loss and prosthesis simulation in audiology. *Hearing Journal, 60*, 32-33. https://doi.org /10.1097/01.HJ.0000281789.77088.

## Supplementary materials (statistical results and model summaries).

***Table 1: Time8***

| **Fixed effects** | | **Estimate** | **SE** | **t** | **Estimated *p*** |
|---|---|---|---|---|---|
| (intercept) | | 4.52 | 0.27 | 16.79 | **<.001** |
| Condition | BAB2 | -0.17 | 0.17 | -0.97 | .334 |
| | HLS | 1.51 | 0.18 | 8.42 | **<.001** |
| Talker Group | OANH | 0.33 | 0.34 | 0.95 | .341 |
| | OAHL | 1.32 | 0.35 | 3.72 | **<.001** |
| Random effects | | Variance | | | |
| Participant | (intercept) | 1.22 | | | |
| Residual | | 1.19 | | | |

Number of observations=234, Participants (N)=82

***Table 2: Articulation rate***

| **Fixed effects** | | **Estimate** | **SE** | **t** | **Estimated *p*** |
|---|---|---|---|---|---|
| (intercept) | | 3.86 | 0.08 | 49.88 | **<.001** |
| Condition | BAB2 | -0.10 | 0.06 | -1.88 | .060 |
| | HLS | 0.43 | 0.06 | -7.55 | **<.001** |
| Sex | Male | 0.28 | 0.08 | 3.26 | **.001** |
| Talker Group | OANH | -0.39 | 0.09 | -4.54 | **<.001** |
| | OAHL | -0.39 | 0.09 | -4.04 | **<.001** |
| Condition*Sex | BAB2, M | -0.11 | 0.08 | -1.38 | .168 |
| | HLS, M | 0.11 | 0.08 | 1.36 | .173 |
| Random effects | | Variance | | | |
| Participant | (intercept) | 0.08 | | | |
| Residual | | 0.07 | | | |

Number of observations=244, Participants (N)=82

***Table 3: Pause frequency***

| **Fixed effects** | | **Estimate** | **SE** | **t** | **Estimated *p*** |
|---|---|---|---|---|---|
| (intercept) | | 0.09 | 0.01 | 12.67 | **<.001** |
| Condition | BAB2 | -0.01 | 0.01 | -2.73 | **.006** |
| | HLS | 0.01 | 0.01 | 2.23 | **.026** |
| Talker Group | OANH | -0.02 | 0.01 | -1.73 | .083 |
| | OAHL | <.0.01 | 0.01 | 0.43 | .667 |
| Condition*Talker Group | BAB2, OANH | 0.02 | 0.01 | 2.90 | **.004** |
| | HLS, OANH | 0.01 | 0.01 | 1.93 | .053 |
| | BAB2, OAHL | 0.01 | 0.01 | 0.75 | .450 |
| | HLS, OAHL | 0.01 | 0.01 | 0.98 | .326 |
| Random effects | | Variance | | | |
| Participant | (intercept) | <.0.01 | | | |
| Residual | | <.0.01 | | | |

Number of observations=244, Participants (N)=82

***Table 4: ME1-3kHz***

| **Fixed effects** | | **Estimate** | **SE** | **t** | **Estimated *p*** |
|---|---|---|---|---|---|
| (intercept) | | 66.46 | 0.53 | 124.29 | **<.001** |
| Condition | BAB2 | 3.19 | 0.34 | 9.32 | **<.001** |
| | HLS | 1.57 | 0.34 | 4.63 | **<.001** |
| Talker Group | OANH | -2.55 | 0.66 | -3.87 | **<.001** |
| | OAHL | -4.02 | 0.68 | -5.91 | **<.001** |
| Sex | M | -2.12 | 0.50 | -4.27 | **<.001** |
| Condition*Talker Group | BAB2, OANH | 0.98 | 0.46 | 2.11 | **.034** |
| | HLS, OANH | 0.34 | 0.46 | 0.74 | 0.462 |
| | BAB2, OAHL | 1.65 | 0.48 | 3.48 | **<.001** |
| | HLS, OAHL | 1.08 | 0.48 | 2.27 | **.023** |
| Random effects | | Variance | | | |
| Participant | (intercept) | 4.52 | | | |
| Residual | | 1.43 | | | |

Number of observations=242, Participants (N)=82

***Table 5: Median F0***

| **Fixed effects** | | **Estimate** | **SE** | **t** | **Estimated *p*** |
|---|---|---|---|---|---|
| (intercept) | | 91.40 | 0.65 | 140.91 | **<.001** |
| Condition | BAB2 | 2.37 | 0.47 | 5.02 | **<.001** |
| | HLS | 1.26 | 0.42 | 3.03 | **.002** |
| Talker Group | OANH | -2.06 | 0.88 | -2.35 | **.019** |
| | OAHL | -0.33 | 0.93 | -0.35 | .726 |
| Sex | M | -9.17 | 1.01 | -9.04 | **<.001** |
| Condition*Talker Group | BAB2, OANH | 2.21 | 0.64 | 3.47 | **<.001** |
| | HLS, OANH | 0.18 | 0.56 | 0.32 | .748 |
| | BAB2, OAHL | 2.31 | 0.68 | 3.39 | **<.001** |
| | HLS, OAHL | 0.86 | 0.60 | 1.43 | .154 |
| Condition*Sex | BAB2, M | 2.97 | 0.71 | 4.17 | **<.001** |
| | HLS, M | 1.14 | 0.63 | 1.81 | .070 |
| Talker group *Sex | NH, M | 3.66 | 1.33 | 2.76 | **.006** |
| | HL, M | 1.47 | 1.35 | 1.09 | .276 |
| Condition*Talker Group* Sex | BAB2, OANH, M | -2.24 | 0.96 | -2.32 | **.020** |
| | HIS, OANH, M | -0.36 | 0.85 | -0.42 | .673 |
| | BAB2, OAHL, M | -1.95 | 0.99 | -1.98 | **.048** |
| | HIS, OAHL, M | -0.33 | 0.86 | -0.38 | .704 |
| Random effects | | Variance | | | |
| Participant | (intercept) | 4.67 | | | |
| Residual | | 1.22 | | | |

Number of observations=245, Participants (N)=82

***Table 6: vowel F1 range***

| **Fixed effects** | | **Estimate** | **SE** | **t** | **Estimated *p*** |
|---|---|---|---|---|---|
| (intercept) | | 5.19 | 0.14 | 36.17 | **<.001** |
| Condition | BAB2 | 0.76 | 0.81 | 9.27 | **<.001** |
| | HLS | 0.70 | 0.81 | 8.58 | **<.001** |
| Talker Group | OANH | -0.28 | 0.17 | -1.71 | .087 |
| | OAHL | -0.72 | 0.17 | -4.23 | **<.001** |
| Sex | M | -0.18 | 0.15 | -1.18 | .240 |
| Condition*Sex | BAB2, M | -0.59 | 0.12 | -4.84 | **<.001** |
| | HLS, M | -0.47 | 0.12 | -3.94 | **<.001** |
| Random effects | | Variance | | | |
| Participant | (intercept) | 0.33 | | | |
| Residual | | 0.15 | | | |

Number of observations=244, Participants (N)=83

***Table 7: Vowel F2 range***

| **Fixed effects** | | **Estimate** | **SE** | **t** | **Estimated *p*** |
|---|---|---|---|---|---|
| (intercept) | | 8.26 | 0.190 | 43.39 | **<.001** |
| Condition | BAB2 | -0.53 | 0.21 | -2.55 | **.011** |
| | HLS | 0.44 | 0.21 | 2.09 | **.036** |
| Sex | Male | -0.89 | 0.19 | -4.67 | **<.001** |
| Talker Group | OANH | 1.15 | 0.23 | 4.93 | **<.001** |
| | OAHL | 1.07 | 0.24 | 4.49 | **<.001** |
| Condition*Sex | BAB2, M | 0.58 | 0.21 | 2.78 | **.005** |
| | HLS, M | 0.47 | 0.21 | 2.24 | **.025** |
| Condition*Talker Group | BAB2, OANH | -0.62 | 0.26 | -2.40 | **.016** |
| | HLS, OANH | -0.89 | 0.25 | -3.52 | **<.001** |
| | BAB2, OAHL | -0.82 | 0.26 | -3.13 | **.002** |
| | HLS, OAHL | -1.08 | 0.26 | -4.13 | **<.001** |
| Random effects | | Variance | | | |
| Participant | (intercept) | 0.30 | | | |
| Residual | | 0.44 | | | |

Number of observations=244, Participants (N)=83