Published in final edited form as:

Hear Res. 2014 October ; 0: 102–109. doi:10.1016/j.heares.2014.07.008.

Abnormal intelligibility of speech in competing speech and in noise in a frequency region where audiometric thresholds are near-normal for hearing-impaired listeners

Agnès C. Léger^{a,*}, David T. Ives^b, and Christian Lorenzi^b

^a Research Laboratory of Electronics, Massachusetts Institute of Technology, 77 Massachusetts avenue, Cambridge, MA 02139-4307, USA.

^b Laboratoire des systèmes perceptifs. UMR CNRS 8248. Département d'Etudes Cognitives, Institut d'Etude de la Cognition, École normale supérieure, Paris Sciences et Lettres, 29 rue d'Ulm, 75005 Paris, France.

Abstract

The ability to identify syllables in the presence of speech-shaped noise and a single-talker background was measured for 18 normal-hearing (NH) listeners, and for eight hearing-impaired (HI) listeners with near-normal audiometric thresholds for frequencies up to 1.5 kHz and a moderate to severe hearing loss above 2 kHz. The stimulus components were restricted to the low-frequency (< 1.5 kHz) region, where audiometric thresholds were classified clinically as normal or near normal for all listeners. Syllable identification in a speech background was measured as a function of the fundamental-frequency (F0) difference between competing voices (ranging from 1 semitone to ~1 octave). HI listeners had poorer syllable intelligibility than NH listeners in all conditions. Intelligibility decreased by about the same amount for both groups when the F0 difference between competing voices was reduced. The results suggest that the ability to identify speech against noise or an interfering talker was disrupted in frequency regions of near-normal hearing for HI listeners, but that the ability to benefit from the tested F0 differences was not disrupted. This deficit was not predicted by the elevated absolute thresholds for speech in speech, but it was for speech in noise. It may result from supra-threshold auditory deficits associated with ageing.

Keywords

Speech; interfering speaker; fundamental frequency; noise; cochlear damage; supra-threshold deficits; age

© 2014 Elsevier B.V. All rights reserved.

* Corresponding author: Research Laboratory of Electronics, Room 36-757, Massachusetts Institute of Technology, 77 Massachusetts avenue, Cambridge, MA 02139-4307, USA. Tel.: +1-617-253-8134. legeragnes@gmail.com.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. INTRODUCTION

Listeners with sensorineural hearing loss typically show substantial difficulties in identifying speech, especially in complex backgrounds such as fluctuating noise or single- and multi-talker backgrounds (Baer and Moore, 1993, 1994; Desloge et al., 2010; Dubno et al., 2003; Duquesnoy, 1983; Festen and Plomp, 1990; George et al., 2006; Gustafsson and Arlinger, 1994; Peters et al., 1998; Summers and Molis, 2004; ter Keurs et al., 1993; for a review, see Moore, 2007). The reduction in the ability of hearing-impaired (HI) listeners to understand speech in the presence of competing backgrounds has been traditionally explained in terms of elevated audiometric thresholds and the associated reduced audibility, and supra-threshold auditory deficits in the spectral and temporal analysis of incoming sounds (e.g., Bernstein et al., 2013; Jepsen and Dau, 2011; Plomp, 1978; for reviews, see George et al., 2006; Moore, 2007; Rhebergen et al., 2010a, 2010b). Several supra-threshold auditory deficits have been identified, such as reduced frequency selectivity, reduced temporal-envelope sensitivity and/or reduced temporal resolution, and reduced ability to process temporal fine structure cues (TFS, the rapid fluctuations in amplitude close to the center frequency of a narrow-band signal; Moore, 2007, 2014).

Over the last decade, a wealth of psychophysical and modeling studies have attempted to tease apart the contribution of reduced audibility and supra-threshold auditory deficits to the intelligibility of speech in complex backgrounds for HI listeners (e.g., Bernstein and Grant, 2009; Christiansen and Dau, 2012; Léger et al., 2012b, 2012c; Rhebergen et al., 2006, 2010b; Strelcyk and Dau, 2009). In most studies, reduced audibility, as measured by audiometric hearing loss, was not sufficient to explain the poorer intelligibility of speech, especially in the presence of fluctuating noise or interfering speech (e.g., Bernstein and Grant, 2009; Hopkins and Moore, 2011; Lorenzi et al., 2006; Neher et al., 2012; Sheft et al., 2012; Summers and Molis, 2004; Summers et al., 2013). However, a recent attempt to fully compensate for the reduction in speech audibility was successful at restoring normal speech intelligibility in several conditions where speech was masked by fluctuating noise (Phatak and Grant, 2012). Moreover, models based on speech audibility such as the extended speech intelligibility index model (ESII; Rhebergen et al., 2006, 2010a, 2010b) are reasonably successful at predicting the intelligibility of speech in non-stationary noise for HI listeners. Thus, it is still unclear to what extent supra-threshold auditory deficits contribute to the differences in speech intelligibility between NH and HI listeners. The goal of the present study was to investigate the potential contribution of supra-threshold auditory deficits to speech intelligibility in a background sound in conditions limiting the influence of reduced audibility.

Several studies (e.g., Horwitz et al., 2002; Léger et al., 2012b, 2012c; Lorenzi et al., 2009; Papakonstantinou et al., 2011; Strelcyk and Dau, 2009) have shown mild-to-severe speech identification deficits for HI listeners using stimuli filtered into a low-frequency region (lower than 1-3 kHz) where audiometric thresholds were classified as normal (lower than 20 dB HL) or near-normal (lower than 25-30 dB HL) according to audiological conventions (i.e., ANSI, 1996).

Strelcyk and Dau (2009) measured speech-reception thresholds (SRTs) for speech in noise for sentences lowpass filtered at 1 kHz for normal-hearing (NH) listeners and HI listeners with elevated audiometric thresholds above 1.5 kHz, but normal (≤ 20 dB HL) audiometric thresholds up to 1 kHz. The SRTs were measured for speech in an unmodulated speech-shaped noise and in amplitude-modulated speech-shaped noise. Both noises were lowpass filtered at 1 kHz. All HI listeners had higher SRTs in noise than the NH listeners, but the difference between HI and NH listeners was significant only for the modulated noise. In the latter condition, SRTs were poorer for the HI listeners than for the NH listeners by about 3 dB. Léger et al. (2012b) measured consonant identification in quiet and in the presence of unmodulated, amplitude-modulated, and spectrally modulated speech-shaped noise at several signal-to-noise ratios for stimuli restricted to a low-frequency region (<1.5 kHz) for NH listeners and HI listeners with near-normal hearing (≤ 25 dB HL) in the tested low-frequency region. Consonant identification was poorer for the HI listeners than for the NH listeners in all conditions. However, the deficits for the HI listeners relative to the NH listeners were much greater for speech in noise (deficit of about 10 percentage points) than in quiet (deficit of about 5 percentage points). Inconsistent with the results of Strelcyk and Dau (2009), the deficit shown by HI listeners was comparable for unmodulated and amplitude-modulated noise. The results of these two studies, amongst others, suggest that supra-threshold auditory deficits affect speech intelligibility when speech is presented in background sounds. However, the results also indicate that the supra-threshold deficits associated with mild forms of hearing loss do not necessarily affect the ability to make use of spectral or temporal gaps in the masker.

Strelcyk and Dau (2009) also measured the ability of their NH and HI listeners to identify sentences in a two-talker background (running male and female speech mixed at equal level and time reversed). SRTs were poorer for the HI than for the NH listeners by about 4 dB. In contrast to the conditions described above, the stimuli (target sentences and two-talker background) were not restricted to the low-frequency region where audiometric thresholds were normal (≤ 20 dB HL). The HI listeners showed elevated audiometric thresholds above 1 kHz, and stimuli were not spectrally shaped to compensate for their reduced audibility in the high-frequency region. Although SRTs were not significantly correlated with audiometric thresholds at any frequency, it is unclear whether supra-threshold deficits in the low-frequency region were entirely responsible for the reduced ability of the HI listeners to identify the target speech. Audibility and/or supra-threshold deficits in the high-frequency region may also have limited speech intelligibility, because the HI listeners might not have benefited as much as NH listeners from speech cues available in the high-frequency region.

The goal of the present study was to assess whether supra-threshold auditory deficits are associated with a reduced ability to identify speech in background speech for HI listeners, using an approach similar to that developed by Horwitz et al. (2002), Léger et al. (2012a, 2012b, 2012c), Lorenzi et al. (2009) and Strelcyk and Dau (2009). Identification scores were measured for nonsense syllables filtered into a “low” frequency region (≤ 1.5 kHz) for listeners with normal (≤ 20 dB HL) or near-normal (≤ 30 dB HL) audiometric thresholds below 1.5 kHz and a moderate to severe hearing loss above 1.5 kHz (the across-frequency average of the audiometric thresholds above 1.5 kHz ranged from 35 to 73 dB HL). Identification scores were measured for syllables uttered by a male speaker presented in

quiet, in unmodulated speech-shaped noise, and in a single-talker background (nonsense syllables uttered by a female speaker). In the condition where target syllables were presented against competing syllables, consonant identification was measured as a function of the fundamental frequency (F0) difference between the two voices (this was achieved by shifting down the F0 of the competing syllables). Based on previous work (e.g., Assmann and Summerfield, 1990, 1994; Brokx and Nootebom, 1982; Chalikia and Bregman, 1993; Oxenham and Simonson, 2009; Scheffers, 1983), it was expected that consonant identification would decrease as the F0 difference between concurrent voices was reduced. Supra-threshold auditory deficits (e.g., reduced frequency selectivity and/or degraded TFS coding) may degrade the ability to perceive F0 differences and in turn the ability to separate competing talkers. Therefore, it was also expected that HI listeners would be less affected than NH listeners by changes in F0 difference between competing syllables, as suggested by previous work conducted with broadband stimuli (e.g., Arehart et al., 1997; Strelcyk and Dau, 2009; Stubbs and Summerfield, 1988; Summers and Leek, 1998).

Several studies have demonstrated the influence of age on speech intelligibility for listeners with near-normal audibility in the tested frequency regions. A deleterious effect of age was reported by Grose et al. (2009), Dubno et al. (2002) and Moore (2012) for the identification of broadband speech in noise, and by Vongpaisal and Pichora-Fuller (2007) and Arehart et al. (2011) for the identification of vowels presented in a competing speech background. Furthermore, several studies showed an effect of age on various psychophysical tasks testing frequency discrimination (Harris et al., 2008; He et al., 1998; Konig, 1957), temporal-envelope processing (Grose et al., 2009; Harris et al., 2010; He et al., 2008) and TFS sensitivity (Füllgrabe, 2013; He et al., 2007; Moore et al., 2012; Strelcyk and Dau, 2009). However, in some studies (e.g., Moore et al., 2012), the effect of age disappeared when the effect of audiometric thresholds was controlled for. Overall, these results suggest that ageing may lead to supra-threshold deficits, which may in turn affect speech intelligibility. In the present study, the effects of age on speech intelligibility was assessed by using a wide range of age for NH listeners.

2. METHOD

2.1. Listeners

All listeners were fully informed about the goal of the study and provided written consent before their participation. The study was approved by the French “Regional Ethics Committee” (CPP Ile de France; 07018 - ID RCB: 2007-A00343-50). All listeners were native French speakers and they had no history of cognitive impairment or psychiatric disorders. Audiometric thresholds and ages are reported in Table 1.

Eighteen listeners had normal (≤ 20 dB HL) audiometric thresholds for octave frequencies between 0.125 and 8 kHz. The NH listeners were aged 21 to 60 years (mean=40 years). Eight listeners had normal (≤ 20 dB HL) or near-normal (≤ 30 dB HL) audiometric thresholds for octave frequencies between 0.125 and 1.5 kHz and a hearing loss at higher frequencies. All hearing losses were of sensory origin, as confirmed by the absence of air-bone gaps in the audiometric thresholds. The HI listeners did not suffer from tinnitus and their age range

was 47 to 74 years (mean=62 years). Despite attempts to match the ages of the two groups, the HI listeners were older than the NH listeners.

The low-frequency average of the audiometric thresholds (LFA; 0.125 to 1.5 kHz) was comparable for NH (12 dB HL) and HI (16 dB HL) listeners. A two-way ANOVA with group (NH and HI) and frequency (in the low frequency region, 0.125 to 1.5 kHz) as within-subject factors showed that there was no significant difference between groups [$F(1,24)=2$, $p=0.15$]. However, there was a significant interaction between group and frequency [$F(5,120)=9$, $p<0.001$]. A post-hoc analysis (Tukey test) showed that the audiometric thresholds of the HI listeners were significantly higher than those of NH listeners at 1.5 kHz ($p<0.01$, $p>0.14$ for all the other comparisons). The resulting difference in audibility at 1.5 kHz between NH and HI listeners might have influenced the results. There was no significant correlation between age and LFA [$r=0.3$, $p=0.13$], but there was a significant correlation between age and audiometric threshold at 1.5 kHz [$r=0.41$, $p<0.05$].

2.2. Material

A method similar to that described by Léger et al. (2012c) was used to measure speech intelligibility. The reader is referred to that paper for details. The most significant features and methodological differences are described below.

2.2.1. Speech material—A set of 48 Vowel-Consonant-Vowel (VCV) stimuli spoken by a male native French speaker was used. The stimuli were composed of 16 consonants, / p, t, k, b, d, g, f, s, ʃ, v, z, ʒ, l, r, m, n /, combined with three different vowels, /i, a, u/. The set of 48 VCV stimuli was recorded twice, producing a complete set of 96 VCV items. The mean F0 of the male voice was estimated as 115 Hz using the YIN algorithm (de Cheveigné and Kawahara, 2002; standard deviation = 14 Hz). A speech-shaped noise (SSN) with a spectrum that matched the average spectrum of the whole set of VCV stimuli was generated.

The speech stimuli were low-pass filtered to restrict them to the tested region (< 1.5 kHz, Butterworth filter, 216 dB/oct slope). The speech stimuli were always presented with a highpass filtered SSN (>1.5 kHz, Butterworth filter, 108 dB/oct slope) at a signal-to-noise ratio (SNR) of +12 dB, to prevent the use of information from the transition band (Warren et al., 2004).

The speech stimuli were presented at 65 dB(A), except for one HI listener whose LFA was above 20 dB HL (27 dB HL). For this listener, speech stimuli were presented at 75 dB(A). Therefore, the speech signals were presented at an average sensation level ranging from about 40 to 50 dB SL for all except one listener. For one NH listener who had a negative LFA (-5 dB HL), the speech signals were presented at about 60 dB SL.

2.2.2. Maskers—All maskers were low-pass filtered to restrict them to the tested region (< 1.5 kHz). Maskers were all presented at a SNR of 0 dB. This SNR was chosen based on related research (Léger et al., 2012c) and piloting to avoid floor and ceiling effects for both NH and HI listeners.

In one condition, the masker was a steady SSN with a new sample of the SSN used for each presentation. In the other conditions, the masker was a VCV. Two sets of 48 VCV stimuli spoken by a female speaker were used, with the same vowel and consonant combination as before. The mean F0 of the female voice was estimated as 226 Hz using the YIN algorithm (standard deviation = 23 Hz). The ratio between the F0 of the male (115 Hz) and female (226 Hz) was close to an octave. A new VCV was used for each presentation, without replacement. The interfering VCV was randomly selected, so that both the vowel and the consonant could either be the same or different across the two speakers. The F0 of the interfering speaker was processed using the Praat software (Boersma and Weenink, 2010) to estimate the influence of F0 separation between the target and interfering speakers. Three conditions were tested, referred to as:

1. Interfering speaker, Octave (IS-Oct). The F0 of the female voice was not processed, that is its mean value was left as 226 Hz. For simplicity, this condition is referred to as corresponding to an octave difference between the F0s of the target and interfering speakers, although this is an approximation. Note that in this condition, the stimuli were not processed via Praat, which might have introduced uncontrolled differences between the speech signals used in the different masking conditions.
2. Interfering speaker, 3 semitones (IS-3st). The mean F0 of the interfering speaker was lowered to 136 Hz so that the mean separation between the F0s of the target and interfering speakers was 3 semitones.
3. Interfering speaker, 1 semitone (IS-1st). The mean F0 of the interfering speaker was lowered to 122 Hz so that the mean separation between the F0s of the target and interfering speakers was 1 semitone.

2.3 Procedure

Listeners were tested individually using a single-interval, sixteen-alternative procedure without feedback. Each condition was tested in one run, during which the two sets of 48 VCV utterances were presented in random order. The listener was instructed to identify the consonant spoken by the target male after each presentation. In the IS-1st condition, the gender of interfering talker was ambiguous and the listeners were asked to report the consonant spoken by the voice with the lowest pitch. The 16 possible responses were displayed orthographically on a computer screen, and the listener entered his/her choice by selecting one response using a computer mouse. Consonant identification scores were recorded.

Prior to data collection, listeners were trained using a run with unfiltered VCV stimuli presented in quiet, to familiarize them with the speech material and the task. This was followed by a run using low-pass filtered VCV stimuli presented in quiet, and finally by the four conditions with maskers in a random order that varied across listeners.

3. RESULTS

3.1. Comprehension of the task

All listeners had scores higher than 78% correct when tested with broadband stimuli. This suggests that all listeners understood the task and that the HI listeners had reasonably good speech intelligibility with unprocessed speech.

3.2. Intelligibility of low-pass filtered speech

All identification scores were converted into rationalized arcsine units (RAU, Howell, 1997; Studebaker, 1985). RAU scores are similar to percent scores for values from 15% to 85%, but RAU scores are “stretched” outside that range, reducing the effects of the bounded percent-correct scale and making the data more suitable for ANOVA. Identification scores are referred to as “scores,” for brevity.

Figure 1 shows the averaged and individual scores for the NH and HI listeners. HI listeners generally had lower scores than NH listeners. Scores were best in the quiet condition and poorest in the noise condition for all groups. For the speech masker, decreasing the difference in F0 between voices generally led to a decrease in scores. However, decreasing the F0 difference below 3 semitones did not affect the scores for the NH listeners, whereas it did lead to a small decrease in scores for the HI listeners.

An ANOVA was conducted on the scores with group (NH and HI) as a between-subject factor and condition (Quiet, IS-Oct, IS-3st, IS-1st, Noise) as a between-subject factor. The main effect of group was significant [$F(1,24)=8, p<0.05$]. The main effect of condition was significant [$F(4,96)=396, p<0.001$]. A post-hoc analysis showed that mean scores for the conditions with maskers (noise or interfering speaker) were lower than the mean score for the Quiet condition (all $p<0.001$). Mean scores for the interfering speaker conditions were higher than the mean score for the Noise condition (all $p<0.001$). The mean score for the IS-Oct condition was higher than for the IS-3st and IS-1st conditions (both $p<0.001$), but the scores for the IS-3st and the IS-1st conditions were not significantly different ($p=0.87$). The interaction between group and condition was not significant [$F(4,96)=1, p=0.26$]. This suggests that HI listeners did not have a significantly different deficit when the background was noise or speech. It also suggests that the effect of changing the F0 difference between the two speakers was similar for NH and HI listeners, even though the HI listeners had lower scores in all conditions.

3.3. Influence of age and audiometric thresholds on the intelligibility of low-pass filtered speech

Pearson correlations and partial correlations were calculated between scores for each of the five conditions, and age and audiometric thresholds in the low-frequency region (LFA and audiometric threshold at 1.5 kHz) for all listeners, to estimate if factors related to ageing and/or influencing audiometric thresholds also influenced speech intelligibility. For HI listeners only, a correlation was also calculated between scores and audiometric thresholds in the high-frequency region (HFA, between 2 and 8 kHz) to estimate if speech intelligibility deficits in the low-frequency region were related to the severity of the hearing loss. For each

analysis, the criterion for significance ($p < 0.05$) was adjusted using a Bonferroni correction which divides the criterion by the number of comparisons (5 comparisons). Therefore, significance was considered as achieved for $p < 0.01$. The results of these analyses are reported in Table 2.

Figure 2 shows the individual scores as a function of age. HI listeners were on average older than NH listeners, and this may have been a confounding factor in the analysis. There were significant negative correlations between age and score for the Quiet and the IS-1st conditions, but not for the other conditions, although all correlations were of similar magnitude. When the effect of the LFA (between 0.125 and 1.5 kHz) was partialled out, the correlation between age and score was significant for the Quiet, IS-Oct and IS-1st conditions. Taken together, these results suggest that age might have contributed to the deficit demonstrated by the HI listeners.

Figure 3 shows the individual scores as a function of the LFA. There was a significant correlation between LFA and scores for the Noise condition, but not for the other conditions. Note that this significant correlation might have been driven by two outliers, one NH listener and one HI listener, who had the lowest (−5 dB HL) and highest (27 dB HL) LFA, respectively. Also the HI listener that had the poorest scores in noise had the best or second best scores amongst the HI listeners in all other conditions. When the effect of age was controlled using a partial correlation analysis, none of the correlations between LFA and score were significant. These results suggest that elevated audiometric thresholds (as estimated by the LFA) may not have driven the intelligibility deficit demonstrated by the HI listeners for speech in speech, but may have contributed to their deficits for speech in noise.

The HI listeners had significantly higher audiometric thresholds than the NH listeners at 1.5 kHz. It could be the case that elevated audiometric thresholds at 1.5 kHz had an influence on speech intelligibility. Therefore, a correlation analysis between audiometric thresholds at 1.5 kHz and scores was computed. There was a significant correlation between audiometric thresholds at 1.5 kHz and scores for the Noise condition, but not for the other conditions. The individual scores in the Noise condition are plotted as a function of their audiometric threshold at 1.5 kHz in the bottom right panel of Figure 3. The correlation remained significant when the effect of age was controlled using a partial correlation analysis. This result suggests that hearing loss in the highest frequencies of the tested frequency region might have affected speech intelligibility in a noise background.

Correlations between HFA and scores were computed for the HI listeners only. There were no significant correlations between HFA and scores. This suggests that the severity of the hearing loss in the high-frequency region did not influence speech intelligibility for the HI listeners when tested in the low-frequency region.

4. DISCUSSION

4.1. Speech intelligibility deficits for HI listeners in frequency regions of clinically normal or near-normal hearing

Consistent with previous work, HI listeners showed speech-intelligibility deficits in frequency regions where hearing sensitivity was considered as clinically normal or near-normal for speech presented in quiet (as reported by Léger et al., 2012b, 2012c; and Lorenzi et al., 2009) and in a (notionally) steady noise (as reported by Horwitz et al., 2002; and Léger et al., 2012b, 2012c). These results support the idea that HI listeners have supra-threshold auditory deficits in frequency regions where audibility is normal or near-normal. Furthermore, the HI listeners demonstrated a poorer ability than NH listeners to understand speech in the presence of a competing speaker when both target and masker were restricted to the low-frequency region where audiometric thresholds were normal or near normal. This confirms and extends the results obtained by Strelcyk and Dau (2009) for HI listeners using target and interfering speech that were not restricted to the low-frequency region of normal hearing.

The deficits demonstrated by the HI listeners were modest (about 5 RAU), and similar in the different conditions tested in the current study (quiet, noise, competing speaker). Modest deficits were also reported by Horwitz et al. (2002), Léger et al. (2012c) and Lorenzi et al. (2009) for HI listeners with similar audiograms and tested with similar methods in quiet and in noise.

4.2. Effects of varying the F0 difference between competing speakers on speech identification

There was a clear effect of varying the F0 of the interfering speaker on the ability of both NH and HI listeners to identify speech in a speech background filtered into the low-frequency region. Intelligibility worsened by about 8-10 RAU when the F0 separation between the two competing speakers was decreased from 1 octave to 1 semitone. This is consistent with the results of Oxenham and Simonson (2009) showing that, for lowpass filtered (<1.2 kHz) sentences, decreasing F0 separation between two competing speakers (from about 4.5 to about 0.5 semitones) led to poorer speech intelligibility (decrease of about 20%) at 0-dB SNR.

Speech intelligibility in a speech background was also impaired for HI listeners in frequency regions of normal or near-normal audiometric thresholds. However, this deficit was modest, and smaller than that observed in noise, and the data for NH and HI listeners strongly overlapped for the largest F0 differences (further testing is necessary to confirm this deficit). Furthermore, despite tendencies, the HI listeners did not show significantly different changes in performance when the F0 separation between the two speakers was varied (in contrast with previous studies that showed a smaller influence of changes in F0 difference for HI listeners than for NH listeners; e.g., Arehart et al., 1997; Strelcyk and Dau, 2009; Stubbs and Summerfield, 1988; Summers and Leek, 1998). Therefore, it is unclear whether HI listeners had an impaired ability to use F0 cues to understand speech in speech. For NH listeners, most of the improvement in speech identification (e.g., vowel identification)

associated with F0 differences between competing speech signals occurs between 0- and 1-semitone F0-differences, with little additional benefit as differences in F0 increase beyond 1 semitone (e.g., Scheffers, 1983; Assmann & Summerfield, 1990; Summers and Leek, 1998). In the present study, the smaller magnitude of F0 differences between concurrent syllables was equal to 1 semitone. This may explain why, in comparison with Summers and Leek (1998) who used much lower magnitudes in F0 difference (i.e., 0, 0.25, 0.5 and 1 semitone, etc.) when assessing concurrent-vowel identification for their HI listeners, the present NH listeners did not benefit more than HI listeners from F0 differences between competing speech signals.

4.3. Effect of elevated audiometric thresholds

Audiometric thresholds in the tested frequency region, as estimated by the LFA, were not significantly higher for HI listeners than for NH listeners (younger and older), despite being slightly elevated. No relationship was found between scores for speech in quiet and in speech and audiometric thresholds. However, there was a significant correlation between scores for speech in noise and LFA and audiometric thresholds at 1.5 kHz (the cut-off frequency of the filtered speech). At this frequency, the HI listeners had higher audiometric thresholds than the NH listeners. It could be the case that a reduction of audibility at this frequency contributed to the intelligibility deficits demonstrated by the HI listeners for speech in noise.

There was no relationship between speech intelligibility scores and hearing loss in the high-frequency region for HI listeners. This suggests that the severity of the high-frequency hearing loss did not drive the speech intelligibility deficit demonstrated by HI listeners in the low-frequency region.

4.4. Effect of age

No significant correlation between age and scores in noise was found. This is consistent with previous work conducted by Léger et al. (2012) with the same material and methods. This is however inconsistent with previous work showing a modest effect of age on the identification of nonsense syllables or sentences presented against a steady-state noise masker at SNRs close to 0 dB (e.g., Takahashi and Bacon, 1992; Dubno et al., 2002; Moore, 2012). This discrepancy may be related to differences in listening bandwidth: here, the speech target and background were lowpass filtered below 1.5 kHz, whereas they were not filtered in Takahashi and Bacon (1992), Dubno et al. (2002) and Moore (2012).

Significant correlations were found between age and scores in quiet and in the presence of an interfering speaker. This is again inconsistent with the results of Moore (2012) in which young and elderly listeners with normal hearing had comparable identification scores for nonsense syllables presented in quiet. However, speech identification in quiet was close to ceiling in both young and elderly listeners in Moore (2012), precluding the observation of ageing effects. The present results are in line with the results of Strelcyk and Dau (2009) and Moore (2012), showing reduced intelligibility of sentences in a single- or two-talker background in elderly listeners with normal hearing below 1 kHz (Strelcyk and Dau, 2009) or 6 kHz (Moore, 2012). They are also consistent with the results of Summers and Leek

(1998) who found that the ability to benefit from F0 differences between competing speech signals (e.g., concurrent vowels) degrades with age.

In conclusion, the effect of age remains unclear, but the results suggest that the specific effects of background sounds on speech processing - especially those reported for competing speech signals - might be influenced by age.

5. CONCLUSIONS

The intelligibility of speech in speech was investigated for listeners with sensorineural hearing loss using stimuli that were filtered into a low-frequency region where audibility was normal or near normal. Speech identification was also measured for speech presented in quiet or in a steady noise masker in this low-frequency region. The results showed that:

1. HI listeners had modest speech intelligibility deficits in quiet, in noise and in the presence of a competing speaker for various F0 differences between competing voices.
2. HI listeners had speech identification deficits that were similar for speech in noise and in speech. Their deficit was also similar for all the tested F0 differences.
3. Elevated audiometric thresholds might have influenced the intelligibility of speech for speech noise, but not for speech in quiet or in speech.
4. The age of the listener influenced intelligibility for speech in quiet and in speech, but not for speech in noise.

The results suggest that HI listeners tested in frequency regions of near-normal hearing show speech intelligibility deficits that can only partly be explained by audiometric threshold elevation and age. This implies that other factors related to suprathreshold deficits also contribute to their reduced understanding of speech in noise and in speech. However, these perceptual distortions may not affect the ability to use F0 differences (1 semitone) to understand speech in speech.

Acknowledgments

We are very grateful to A Stephan and S Garnier (Entendre, France) for their help in testing the listeners. AC Léger was supported by a grant from National Institutes of Health (NIH/NIDCD) grant R01-DC000117. DT Ives was supported by a grant from Starkey France. C Lorenzi was supported by ANR-11-0001-02 PSL* and ANR-10-LABX-0087, as well as by a grant (HEARFIN Project) from ANR. We would like to thank Brian CJ Moore and two anonymous reviewers for their very helpful comments on previous versions of this manuscript.

REFERENCES

- ANSI. ANSI S3-1996, American National Standards Specification for Audiometers. American National Standards Institute; New York: 1996.
- Arehart KH, King CA, McLean-Mudgett KS. Role of fundamental frequency differences in the perceptual separation of competing vowel sounds by listeners with normal hearing and listeners with hearing loss. *J. Speech Lang. Hear. Res.* 1997; 40:1434–1444. [PubMed: 9430762]
- Arehart KH, Souza PE, Muralimanohar RK, Miller CW. Effects of age on concurrent vowel perception in acoustic and simulated electroacoustic hearing. *J. Speech Lang. Hear. Res.* 2011; 54:190–210. [PubMed: 20689036]

- Assmann PF, Summerfield Q. Modeling the perception of concurrent vowels: vowels with different fundamental frequencies. *J. Acoust. Soc. Am.* 1990; 88:680–697. [PubMed: 2212292]
- Assmann PF, Summerfield Q. The contribution of waveform interactions to the perception of concurrent vowels. *J. Acoust. Soc. Am.* 1994; 95:471–484. [PubMed: 8120258]
- Baer T, Moore BCJ. Effects of spectral smearing on the intelligibility of sentences in noise. *J. Acoust. Soc. Am.* 1993; 94:1229–1241.
- Baer T, Moore BCJ. Effects of spectral smearing on the intelligibility of sentences in the presence of interfering speech. *J. Acoust. Soc. Am.* 1994; 95:2277–2280. [PubMed: 8201124]
- Bernstein JG, Grant KW. Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.* 2009; 125:3358–3372. [PubMed: 19425676]
- Bernstein JGW, Summers V, Grassi E, Grant KW. Auditory models of suprathreshold distortion and speech intelligibility in persons with impaired hearing. *J. Am. Acad. Audiol.* 2013; 24:307–328. [PubMed: 23636211]
- Boersma, P.; Weenink. Praat: doing phonetics by computer. Version 5.1.32. 2010.
- Brokx JPL, Nootbom SG. Intonation and the perceptual separation of simultaneous voices. *J. Phon.* 1982; 10:23–36.
- Chalikia MH, Bregman AS. The perceptual segregation of simultaneous vowels with harmonic, shifted, or random components. *Percept. Psychophys.* 1993; 53:125–133. [PubMed: 8433910]
- Christiansen C, Dau T. Relationship between masking release in fluctuating maskers and speech reception thresholds in stationary noise. *J. Acoust. Soc. Am.* 2012; 132:1655–1666. [PubMed: 22978894]
- De Cheveigné A, Kawahara H. YIN, a fundamental frequency estimator for speech and music. *J. Acoust. Soc. Am.* 2002; 111:1917–30. [PubMed: 12002874]
- Desloge JG, Reed CM, Braida LD, Perez ZD, Delhorne LA. Speech reception by listeners with real and simulated hearing impairment: effects of continuous and interrupted noise. *J. Acoust. Soc. Am.* 2010; 128:342–359. [PubMed: 20649229]
- Dubno JR, Horwitz AR, Ahlstrom JB. Benefit of modulated maskers for speech recognition by younger and older adults with normal hearing. *J. Acoust. Soc. Am.* 2002; 111:2897–2907. [PubMed: 12083223]
- Dubno JR, Horwitz AR, Ahlstrom JB. Recovery from prior stimulation: masking of speech by interrupted noise for younger and older adults with normal hearing. *J. Acoust. Soc. Am.* 2003; 113:2084–2094. [PubMed: 12703719]
- Duquesnoy AJ. Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons. *J. Acoust. Soc. Am.* 1983; 74:739–43. [PubMed: 6630729]
- Festen JM, Plomp R. Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *J. Acoust. Soc. Am.* 1990; 88:1725–1736. [PubMed: 2262629]
- Füllgrabe C. Age-dependent changes in temporal-fine-structure processing in the absence of peripheral hearing loss. *Am. J. Audiol.* 2013; 22:313–315. [PubMed: 23975124]
- George EL, Festen JM, Houtgast T. Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.* 2006; 120:2295–2311. [PubMed: 17069325]
- Große JH, Mamo SK, Hall JW. Age effects in temporal envelope processing: speech unmasking and auditory steady state responses. *Ear Hear.* 2009; 30:568–575. [PubMed: 19633565]
- Gustafsson HA, Arlinger SD. Masking of speech by amplitude-modulated noise. *J. Acoust. Soc. Am.* 1994; 95:518–529. [PubMed: 8120263]
- Harris KC, Eckert MA, Ahlstrom JB, Dubno JR. Age-related differences in gap detection: effects of task difficulty and cognitive ability. *Hear. Res.* 2010; 264:21–29. [PubMed: 19800958]
- Harris KC, Mills JH, He N-J, Dubno JR. Age-related differences in sensitivity to small changes in frequency assessed with cortical evoked potentials. *Hear. Res.* 2008; 243:47–56. [PubMed: 18597958]

- He NJ, Dubno JR, Mills JH. Frequency and intensity discrimination measured in a maximum-likelihood procedure from young and aged normal-hearing subjects. *J. Acoust. Soc. Am.* 1998; 103:553–565. [PubMed: 9440340]
- He NJ, Mills JH, Ahlstrom JB, Dubno JR. Age-related differences in the temporal modulation transfer function with pure-tone carriers. *J. Acoust. Soc. Am.* 2008; 124:3841–3849. [PubMed: 19206810]
- He NJ, Mills JH, Dubno JR. Frequency modulation detection: effects of age, psychophysical method, and modulation waveform. *J. Acoust. Soc. Am.* 2007; 122:467–477. [PubMed: 17614504]
- Hopkins K, Moore BCJ. The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. *J. Acoust. Soc. Am.* 2011; 130:334–349. [PubMed: 21786903]
- Horwitz AR, Dubno JR, Ahlstrom JB. Recognition of low-pass-filtered consonants in noise with normal and impaired high-frequency hearing. *J. Acoust. Soc. Am.* 2002; 111:409–416. [PubMed: 11831814]
- Howell, DC. *Statistical Methods for Psychology*. Duxbury; Belmont, CA: 1997.
- Jepsen ML, Dau T. Characterizing auditory processing and perception in individual listeners with sensorineural hearing loss. *J. Acoust. Soc. Am.* 2011; 129:262–281. [PubMed: 21303008]
- Konig E. Pitch discrimination and age. *Acta Otolaryngol. (Stockh.)*. 1957; 48:475–489. [PubMed: 13508204]
- Léger AC, Moore BCJ, Gnansia D, Lorenzi C. Effect of spectral smearing on temporal and spectral masking release in low- and mid-frequency regions. *J. Acoust. Soc. Am.* 2012a; 131:4114–4123. [PubMed: 22559383]
- Léger AC, Moore BCJ, Lorenzi C. Abnormal speech processing in frequency regions where absolute thresholds are normal for listeners with high-frequency hearing loss. *Hear. Res.* 2012b; 294:95–103. [PubMed: 23104012]
- Léger AC, Moore BCJ, Lorenzi C. Temporal and spectral masking release in low- and mid-frequency regions for normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.* 2012c; 131:1502–1514. [PubMed: 22352520]
- Lorenzi C, Debruille L, Garnier S, Fleuriot P, Moore BCJ. Abnormal processing of temporal fine structure in speech for frequencies where absolute thresholds are normal (L). *J. Acoust. Soc. Am.* 2009; 125:27–30. [PubMed: 19173391]
- Lorenzi C, Gilbert G, Carn H, Garnier S, Moore BCJ. Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proc. Natl. Acad. Sci. U. S. A.* 2006; 103:18866–18869. [PubMed: 17116863]
- Moore, BCJ. *Cochlear Hearing Loss: Physiological, Psychological and Technical Issues*. Wiley-Interscience; Chichester, England: 2007.
- Moore, BCJ. The importance of temporal fine structure for the intelligibility of speech in complex backgrounds. In: Dau, T.; Jepsen, ML.; Cristensen-Dalsgaard, J.-C.; Poulsen, T., editors. *Proceedings of ISAAR 2011: “Speech Perception and Auditory Disorders”*. Centertryk A/S; Denmark: 2012.
- Moore, BCJ. *Auditory processing of temporal fine structure: effects of age and hearing loss*. World Scientific; Hackensack, NJ, NJ: 2014.
- Moore BCJ, Glasberg BR, Stoev M, Füllgrabe C, Hopkins K. The influence of age and high-frequency hearing loss on sensitivity to temporal fine structure at low frequencies (L). *J. Acoust. Soc. Am.* 2012; 131:1003–1006. [PubMed: 22352474]
- Neher T, Lunner T, Hopkins K, Moore BCJ. Binaural temporal fine structure sensitivity, cognitive function, and spatial speech recognition of hearing-impaired listeners (L). *J. Acoust. Soc. Am.* 2012; 131:2561–2564. [PubMed: 22501036]
- Oxenham AJ, Simonson AM. Masking release for low- and high-pass-filtered speech in the presence of noise and single-talker interference. *J. Acoust. Soc. Am.* 2009; 125:457–68. [PubMed: 19173431]
- Papakonstantinou A, Strelcyk O, Dau T. Relations between perceptual measures of temporal processing, auditory-evoked brainstem responses and speech intelligibility in noise. *Hear. Res.* 2011; 280:30–37. [PubMed: 21354285]

- Peters RW, Moore BCJ, Baer T. Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people. *J. Acoust. Soc. Am.* 1998; 103:577–587. [PubMed: 9440343]
- Phatak SA, Grant KW. Phoneme recognition in modulated maskers by normal-hearing and aided hearing-impaired listeners. *J. Acoust. Soc. Am.* 2012; 132:1646–1654. [PubMed: 22978893]
- Plomp R. Auditory handicap of hearing impairment and the limited benefit of hearing aids. *J. Acoust. Soc. Am.* 1978; 63:533–549. [PubMed: 670550]
- Rhebergen KS, Lyzenga J, Dreschler WA, Festen JM. Modeling speech intelligibility in quiet and noise in listeners with normal and impaired hearing. *J. Acoust. Soc. Am.* 2010a; 127:1570–1583. [PubMed: 20329857]
- Rhebergen KS, Versfeld NJ, de Laat JAPM, Dreschler WA. Modelling the speech reception threshold in non-stationary noise in hearing-impaired listeners as a function of level. *Int. J. Audiol.* 2010b; 49:856–865. [PubMed: 20936997]
- Rhebergen KS, Versfeld NJ, Dreschler WA. Extended speech intelligibility index for the prediction of the speech reception threshold in fluctuating noise. *J. Acoust. Soc. Am.* 2006; 120:3988–3997. [PubMed: 17225425]
- Scheffers MT. Simulation of auditory analysis of pitch: an elaboration on the DWS pitch meter. *J. Acoust. Soc. Am.* 1983; 74:1716–1725. [PubMed: 6655129]
- Scheffers, MTM. Ph.D. thesis. University of Groningen; The Netherlands: 1983. Sifting vowels: Auditory pitch analysis and sound segregation..
- Sheft S, Shafiro V, Lorenzi C, McMullen R, Farrell C. Effects of age and hearing loss on the relationship between discrimination of stochastic frequency modulation and speech perception. *Ear Hear.* 2012; 33:709–720. [PubMed: 22790319]
- Strelcyk O, Dau T. Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing. *J. Acoust. Soc. Am.* 2009; 125:3328–3345. [PubMed: 19425674]
- Stubbs RJ, Summerfield Q. Evaluation of two voice-separation algorithms using normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.* 1988; 84:1236–1249. [PubMed: 3198862]
- Studebaker GA. A “rationalized” arcsine transform. *J. Speech Hear. Res.* 1985; 28:455–462. [PubMed: 4046587]
- Summers V, Leek MR. F0 processing and the separation of competing speech signals by listeners with normal hearing and with hearing loss. *J. Speech Lang. Hear. Res.* 1998; 41:1294–1306. [PubMed: 9859885]
- Summers V, Makashay MJ, Theodoroff SM, Leek MR. Suprathreshold auditory processing and speech perception in noise: hearing-impaired and normal-hearing listeners. *J. Am. Acad. Audiol.* 2013; 24:274–292. [PubMed: 23636209]
- Summers V, Molis MR. Speech recognition in fluctuating and continuous maskers: effects of hearing loss and presentation level. *J. Speech Lang. Hear. Res.* 2004; 47:245–256. [PubMed: 15157127]
- Takahashi GA, Bacon SP. Modulation detection, modulation masking, and speech understanding in noise in the elderly. *J. Speech. Hear. Res.* 1992; 35:1410–1421. [PubMed: 1494284]
- Ter Keurs M, Festen JM, Plomp R. Limited resolution of spectral contrast and hearing loss for speech in noise. *J. Acoust. Soc. Am.* 1993; 94:1307–14. [PubMed: 8408972]
- Vongpaisal T, Pichora-Fuller MK. Effect of age on F0 difference limen and concurrent vowel identification. *J. Speech Lang. Hear. Res.* 2007; 50:1139–1156. [PubMed: 17905901]
- Warren RM, Bashford JA, Lenz PW. Intelligibility of bandpass filtered speech: steepness of slopes required to eliminate transition band contributions. *J. Acoust. Soc. Am.* 2004; 115:1292–1295. [PubMed: 15058351]

Intelligibility of speech in noise and in speech can be impaired for speech filtered into frequency regions of near-normal hearing.

The impairment is not fully explained by reduced audibility.

The impairment may result from supra-threshold auditory deficits associated with mild hearing loss and aging.

Figure 1.
Average scores (in RAU) for each condition for the two groups of listeners (NH and HI).
The error bars show ± 1 standard error.

Figure 2.

Individual scores (in RAU) as a function of age (in years) for younger NH (21 to 34 years old), older NH (42 to 60 years old) and HI listeners. Each panel shows the scores for one condition. A star next to the name of the condition indicates that the Person correlation between score and age was significant (corrected significance threshold: $p < 0.01$, see Table 2), p and r values are reported for each comparison.

Figure 3.

Individual scores (in RAU) as a function of LFA (in dB HL) for NH and HI listeners. Results (p and r values) of Pearson correlations between score and LFA are reported. The bottom right panel shows scores in the Noise condition as a function of the audiometric threshold (AT, in dB HL) at 1.5 kHz. Otherwise as Figure 2.

Table 1

Age (in years) and pure-tone audiometric thresholds (in dB HL) for NH and HI listeners. The average (in bold) and standard deviation (SD) are shown for each group. Individual data are shown for HI listeners. The right-most column (LFA) shows the low-frequency average of the audiometric thresholds (across the audiometric frequencies of 0.125 to 1.5 kHz).

Age			Audiometric thresholds										LFA		
			125	250	500	750	1000	1500	2000	3000	4000	6000	8000		
NH	Mean	40	17	14	14	13	11	11	11	10	11	10	17	14	12
	SD	12	4	7	6	4	4	5	5	5	5	7	6	9	5
HI	Mean	62	17	14	13	14	16	23	31	43	50	66	71	16	
	SD	9	6	6	7	8	7	9	10	13	13	16	18	6	
	HI1	47	25	20	15	10	15	25	25	35	35	60	60	18	
	HI2	55	10	10	10	10	15	30	50	65	75	75	75	14	
	HI3	56	20	20	25	30	30	35	40	45	40	60	60	27	
	HI4	63	10	5	5	5	5	15	20	25	40	55	>95	8	
	HI5	66	20	10	5	10	20	30	35	50	55	100	105	16	
	HI6	67	20	15	10	15	20	20	25	40	55	65	65	17	
	HI7	69	20	20	20	20	15	15	30	30	45	45	60	18	
HI8	74	10	10	10	10	10	10	25	50	55	70	>95	10		

Table 2

Correlations between scores in each condition and: age (panel A), LFA (panel B), audiometric threshold at 1.5 kHz (panel C) or HFA (panel D). Partial correlations were computed controlling for the effect of: Panel A, the LFA; Panel B and C: age. Note that the correlations between HFA and scores were calculated for HI listeners only. Pearson's coefficient: r ; level of significance: p ; bold characters: $p < 0.01$ (corrected significance threshold).

A. Age				
Condition	Correlation		Partial correlation	
	r	p	r	p
Quiet	-0.54	0.004	-0.64	0.000
IS-Oct	-0.49	0.011	-0.55	0.002
IS-3st	-0.46	0.018	-0.44	0.019
IS-1st	-0.58	0.002	-0.60	0.000
Noise	-0.37	0.059	-0.27	0.180

B. LFA				
Condition	Correlation		Partial correlation	
	r	p	r	p
Quiet	0.18	0.388	0.43	0.023
IS-Oct	0.09	0.663	0.29	0.148
IS-3st	-0.15	0.456	-0.01	0.943
IS-1st	-0.04	0.865	0.18	0.373
Noise	-0.50	0.009	-0.44	0.019

C. Audiometric threshold at 1.5 kHz				
Condition	Correlation		Partial correlation	
	r	p	r	p
Quiet	-0.09	0.660	0.17	0.410
IS-Oct	0.12	0.560	0.40	0.037
IS-3st	-0.21	0.310	-0.02	0.906
IS-1st	-0.28	0.170	-0.05	0.779
Noise	-0.71	<0.001	-0.66	<0.001

D. HFA [for HI listeners only]

Condition	Correlation	
	r	p
Quiet	-0.30	0.469
IS-Oct	-0.08	0.856
IS-3st	-0.33	0.428
IS-1st	-0.47	0.237
Noise	-0.34	0.409