

# Linking dynamic-range compression across the ears can improve speech intelligibility in spatially separated noise

Ian M. Wiggins<sup>a)</sup> and Bernhard U. Seeber<sup>b)</sup>

MRC Institute of Hearing Research, University Park, Nottingham, NG7 2RD, United Kingdom

(Received 17 June 2012; revised 11 December 2012; accepted 13 December 2012)

Recently introduced hearing devices allow dynamic-range compression to be coordinated at the two ears through a wireless link. This study investigates how linking compression across the ears might improve speech intelligibility in the presence of a spatially separated steady noise. An analysis of the compressors' behavior shows how linked compression can preserve interaural level differences (ILDs) and, compared to compression operating independently at each ear, improve the long-term apparent speech-to-noise ratio (SNR) at the ear with the better SNR. Speech intelligibility for normal-hearing listeners was significantly better with linked than with unlinked compression. The performance with linked compression was similar to that without any compression. The benefit of linked over unlinked compression was the same for binaural listening and for monaural listening to the ear with the better SNR, indicating that the benefit was due to changes to the signal at this ear and not to the preservation of ILDs. Differences in performance across experimental conditions were qualitatively consistent with changes in apparent SNR at the better ear. Predictions made using a speech intelligibility model suggest that linked compression could potentially provide a user of bilateral hearing aids with an improvement in intelligibility of up to approximately ten percentage points.

© 2013 Acoustical Society of America. [http://dx.doi.org/10.1121/1.4773862]

PACS number(s): 43.66.Ts, 43.66.Pn, 43.71.Gv [EB]

Pages: 1004–1016

## I. INTRODUCTION

Dynamic-range compression is used in hearing aids and cochlear implants to compensate for the limited dynamic range of the impaired auditory system. When a person with bilateral hearing loss is provided with two hearing devices, one for each ear, the devices typically operate independently of one another. Thus, any compression will act independently at each ear. Recently, hearing devices have been introduced that incorporate a wireless link, allowing compression (and other processing) to be coordinated across the ears. It has been claimed that linking the devices in this way can improve speech understanding in noise, sound localization, and perceived naturalness of sound (Hansen, 2008; Sockalingam *et al.*, 2009; Kreisman *et al.*, 2010). The typical rationale for linking compression across the ears is that it allows interaural level differences (ILDs) to be preserved—ILDs are one of two types of binaural cue that underlie many of our spatial hearing abilities (Middlebrooks and Green, 1991), and so alterations to natural ILDs caused by compression acting independently at each ear are potentially deleterious (Byrne and Noble, 1998; Keidser *et al.*, 2006; Musa-Shufani *et al.*, 2006; Moore, 2008; Wiggins and Seeber, 2011, 2012). However, the preservation of ILDs is not the sole effect of linking compression across the ears: Changes in the behavior of the compression system as

a whole also affect the monaural signals at each ear. The present study investigates how linking compression across the ears might improve speech intelligibility in the presence of a spatially separated noise, whether through the preservation of ILDs or through changes to the monaural signals delivered to each ear.

There is a rich literature on the impact of compression on speech intelligibility, with reviews provided by Dillon (1996) and Souza (2002), among others. Despite much research, there remains a lack of consensus on how compression should be implemented in hearing devices. It is clear that compression can have both positive effects, such as increasing the audibility of low-level speech (Souza and Turner, 1999) and reducing discomfort for high-level speech (Kam and Wong, 1999), and negative effects, such as reducing temporal and spectral contrast (Plomp, 1988) and disturbing the speech envelope in a variety of other ways (Stone and Moore, 2007). Which of these effects will come to the fore in any particular study will depend on the choice of outcome measure, subject population, stimuli, implementation and parameter settings of the compression system, absolute sound levels, and a variety of other procedural factors. In addition to these sources of variability between studies, a recurring theme in the literature is the importance of individual differences: The best form of compression is unlikely to be the same for all hearing-impaired people, possibly even for those with a similar audiogram (Gatehouse *et al.*, 2006; Moore, 2008).

Speech is easier to understand when it is spatially separated from an interfering sound than when it is co-located with the interferer (Bronkhorst, 2000), a phenomenon known as spatial release from masking (SRM). Hansen (2008) and Kreisman *et al.* (2010) reported better intelligibility of speech

<sup>a)</sup>Author to whom correspondence should be addressed. Current address: NIHR Nottingham Hearing Biomedical Research Unit, Ropewalk House, 113 The Ropewalk, Nottingham, NG1 5DU, United Kingdom. Electronic mail: Ian.Wiggins@nottingham.ac.uk

<sup>b)</sup>Current address: Audio Information Processing, Technische Universität München, Arcisstr. 21, 80333 München, Germany.

in spatially separated noise for hearing-impaired listeners using wirelessly linked hearing aids compared to independently operating devices. However, in neither study could the improvement be attributed directly to linked compression because other aspects of the processing were also synchronized across the ears (e.g., noise reduction and microphone directionality mode), and the wirelessly linked devices were compared to reference devices which sometimes differed in bandwidth and receiver type. Kates (2008) reported the outcome of a study in which listener preference for different compression schemes was evaluated in ten hearing-impaired listeners in a variety of everyday listening situations: There was no clear preference for systems in which compression was linked across the ears.

While studies that have specifically investigated the effect of linking compression across the ears are few, it is important to note that several studies have found significant SRM in users of bilateral hearing aids performing independent compression, albeit smaller than that obtained by normal-hearing listeners (e.g., Marrone *et al.*, 2008; Neher *et al.*, 2009). Thus, it is clear that independent compression does not eradicate the benefit of spatial separation. Indeed, in an evaluation of a two-channel fast-acting compression system, Moore *et al.* (1992) concluded that independent compression at the two ears did not adversely affect the perceptual processing of binaural cues in a speech-recognition task. However, in that study the spatially separated condition comprised independent 12-talker babble maskers (at  $\pm 90^\circ$ ) on each side of a central target talker, and the symmetry of this arrangement may have limited the potential impact of independent compression. Kalluri and Edwards (2007) tested how independent compression affected SRM in a three-talker scenario designed to promote informational masking and therefore increased reliance on perceived location as a cue for extracting the target speech (cf. Freyman *et al.*, 1999). For normal-hearing listeners, independent compression reduced SRM, but only in a condition in which the sources were spatialized solely by ILD cues; when full binaural cues were available, compression caused little reduction in SRM.

Overall, it seems that the question of whether there is any advantage for speech intelligibility of linking compression across the ears remains open. This paper considers the effect of linked versus unlinked fast-acting compression on the intelligibility of speech in the presence of a spatially separated steady noise. Rather than testing in a clinical setting, the approach adopted here was to gain a detailed understanding of how the compression systems behave, and to test the potential consequences in a controlled experiment using a group of homogeneous normal-hearing listeners. Section II describes how each compression system affects the acoustic signals arriving at the listener's ears. The analysis covers both the effect on binaural cues (specifically ILDs) and changes to the monaural signals at each ear. Section III presents the results of a psychoacoustic experiment which tested if and how these effects of compression affect speech intelligibility. Section IV considers the extent to which the effects of compression observed in the experiment may apply to hearing-impaired listeners wearing bilateral hearing

aids, and whether similar effects can be expected in the case of a fluctuating interfering sound.

## II. UNDERSTANDING HOW COMPRESSION AFFECTS THE ACOUSTIC SIGNALS AT THE LISTENER'S EARS

### A. Simulated listening scenario

The simulated listening scenario comprised speech from directly in front and steady speech-shaped noise from an azimuth of  $60^\circ$  (in the right hemifield). The speech had a level of 65 dB sound pressure level (SPL) at the center position of an absent hypothetical listener's head. The level of the noise was set to give a desired speech-to-noise ratio (SNR) at this position. We refer to this as the "nominal source SNR" since the actual SNR at each ear when the head is present is different.

The speech and noise signals were filtered with anechoic head-related transfer functions (HRTFs) and then summed to produce realistic left- and right-ear signals. The HMS II acoustic manikin HRTFs from the AUDIS catalog were used (Blauert *et al.*, 1998), corrected for use with diffuse-field equalized headphones. The left- and right-ear signals thus generated were used as the inputs to the dynamic-range compression systems described below. A block diagram summarizing the main stages in the signal processing scheme is provided in Fig. 1.

### B. Details of the compression systems

All processing was performed independently in two frequency channels: A low-frequency channel (100 to 2000 Hz) and a high-frequency channel (2000 to 5000 Hz). The decision to use two frequency channels was motivated by a desire to capture the differences in behavior of a bilateral compression system at low and high frequencies while keeping the number of channels small so that the behavior could be analyzed on a within-channel basis. The differences in compression behavior at different frequencies arise because naturally occurring ILDs are larger at high frequencies than at low, meaning that the input level to the compressors will differ more between the ears at high frequencies. Filtering into frequency channels was achieved using 256-tap linear-phase finite impulse response filters to avoid phase distortion. The delay introduced by the filters was removed.

Two compression systems were compared: Unlinked and linked compression. In the unlinked case, wide-dynamic-range compression was performed independently at each ear, as would occur in a traditional bilateral hearing-device fitting without communication between the devices. For linked compression, identical gain was applied at both ears at all instants. The amount of gain at any moment was equal to the minimum gain that would have been applied by compressors operating independently at each ear, i.e., maximal compression was applied at both ears at all times, regardless of which ear drove the compressors.

The compression was fast-acting with an ANSI S3.22 attack time of 5 ms and a release time of 60 ms in both frequency channels (ANSI, 2003). The compression ratio was 3:1 in both channels. The compression threshold was 50 dB SPL in the low-frequency channel and 35 dB SPL in the high-frequency channel: These values were approximately equal to

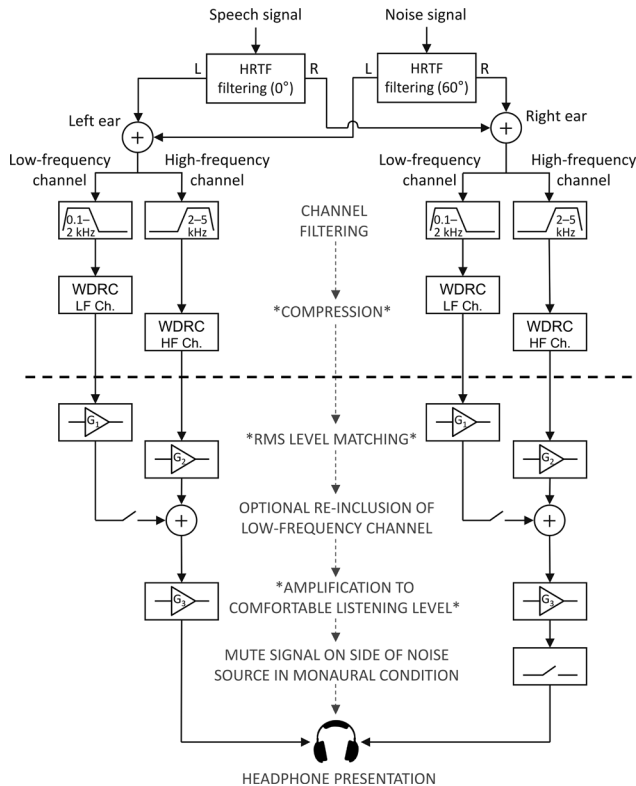


FIG. 1. Signal processing block diagram. Speech and noise signals were filtered with HRTFs and summed to give left- and right-ear signals. Subsequent processing stages are labeled down the center; those highlighted with asterisks involved joint (i.e., bilaterally linked) processing at the two ears. CHANNEL FILTERING: The signal at each ear was filtered into low- (0.1 to 2 kHz) and high- (2 to 5 kHz) frequency channels; \*COMPRESSION\*: Wide dynamic-range compression was applied separately in each frequency channel, either independently at each ear (unlinked condition) or linked across the ears (see main text for details). In the “uncompressed condition,” the compression stage was bypassed. \*RMS LEVEL MATCHING\*: (This and all subsequent stages are specific to the speech intelligibility experiment described in Sec. III.) Long-term levels were matched before and after compression, separately in each frequency channel but with identical gain at both ears to preserve ILDs. OPTIONAL RE-INCLUSION OF LOW-FREQUENCY CHANNEL: The low-frequency channel was added back in to the signal at each ear in the both-channels condition only. \*AMPLIFICATION TO COMFORTABLE LISTENING LEVEL\*: Identical gain was applied at both ears to preserve ILDs. MUTE SIGNAL ON SIDE OF NOISE SOURCE IN MONAURAL CONDITION: Only the “better ear” signal was presented in the monaural condition.

the levels evoked by a steady speech-shaped noise (described below) at an overall level of 50 dB SPL. Overall, these compression parameters are representative of those used in hearing aids providing “syllabic compression,” with the amount of compression being at the upper end of the range prescribed by modern fitting procedures (e.g., Moore *et al.*, 2010). The input–output function was a simple broken-stick shape with a linear response below the compression threshold and a fixed compression ratio above threshold. The audio signal was delayed by 2.5 ms (equal to one-half of the attack time) relative to the gain-control signals to reduce overshoot and undershoot effects (Robinson and Huntington, 1973).

### C. Stimuli

The speech used for the analysis of the compressors’ behavior comprised 16 concatenated Institute of Electrical

and Electronics Engineers (IEEE) sentences. The sentences were taken from a corpus recorded by Stacey and Summerfield (2007) and included at least 1 sentence from each of 12 talkers (6 female, 6 male). The total duration of the speech was 37 s. The steady speech-shaped noise was generated by repeated addition (100 000 repetitions) of randomly selected and time-delayed sentences from the entire corpus. The noise therefore had a spectrum similar to the long-term average of the speech corpus.

### D. Changes in apparent SNR at each ear

Previous studies have shown that the “apparent SNR,” the long-term SNR at the output of a compression system, can differ from the SNR at its input (Souza *et al.*, 2006; Naylor and Johannesson, 2009; Rhebergen *et al.*, 2009). Figure 2 shows the apparent SNR at each ear as a function of nominal source SNR without compression and after unlinked or linked compression. The apparent SNR was determined by storing the gain-control signals generated when the mixture of speech and noise was compressed, applying those gain-control signals to the original speech and noise signals arriving at each ear separately, and then calculating the long-term root-mean-square (rms) levels of the speech and noise at each ear (cf. Stone and Moore, 2007; Rhebergen *et al.*, 2009).

When the noise is substantially higher in level than the speech (i.e., at strongly negative nominal source SNRs), the apparent SNR is similar with or without compression. However, as the noise level is decreased (increasing the nominal source SNR), compression begins to reduce the apparent SNR compared to the uncompressed reference. This is consistent with the results of Souza *et al.* (2006) and Naylor and Johannesson (2009): For speech in steady noise, the apparent

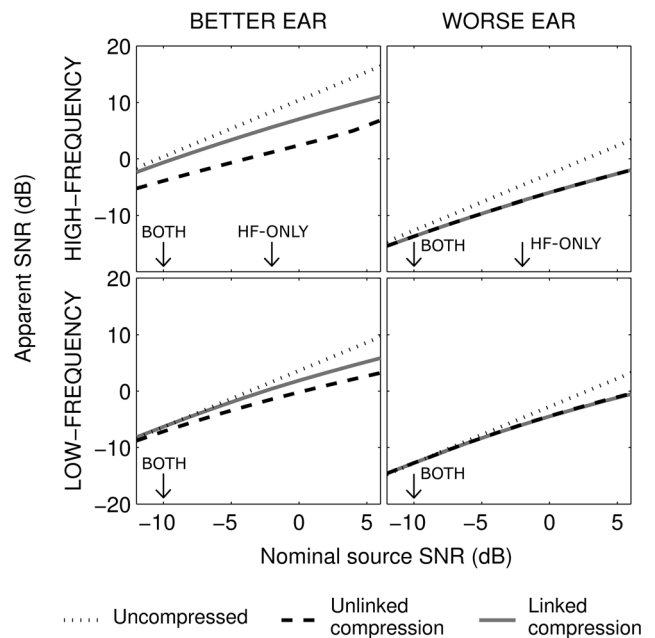


FIG. 2. Apparent long-term SNR at the better ear (left) and worse ear (right) in the high-frequency (upper panels) and low-frequency (lower panels) channel. The vertical arrows indicate the nominal source SNR tested in the both-channels (“BOTH”) and high-frequency-channel-only (“HF-ONLY”) conditions of the speech intelligibility experiment, respectively (this experiment is described in Sec. III).

SNR after compression is the same as or worse than the input SNR. However, at the better ear (the ear with the more favorable SNR), there is a difference between the two types of compression: The apparent SNR is less adversely affected by linked than by unlinked compression. This “benefit” of linked over unlinked compression varies with nominal source SNR and here is a maximum of 2.6 and 4.8 dB in the low- and high-frequency channels, respectively. At the worse ear, the apparent SNR is always identical after either linked or unlinked compression: With speech from directly in front and noise from the right-hand side, the instantaneous level at the right (worse) ear is always at least as great as at the left ear, and so the gain-control signal for the worse ear is the same regardless of whether the compression is linked or not. On average, the SNR at the better ear is 5.8 and 11.7 dB more favorable than at the worse ear in the low- and high-frequency channels, respectively.

To understand how compression changes the apparent SNR, Fig. 3 plots the envelopes of an extract of the speech and noise signals in the high-frequency channel at the better ear. Without compression, the noise envelope is almost flat and the peaks in the speech clearly rise above the noise. The same is true for linked compression because, in this case, the gain-control signal at the better ear is heavily influenced by the *steady* noise signal at the worse ear. Thus, although linked compression reduces the overall gain at the better ear (not visible in Fig. 3 because the data were normalized for plotting), it does so in a relatively steady manner that causes little disturbance to the signal envelopes. Unlinked compression, in contrast, responds more selectively to the peaks in the speech signal at the better ear, causing the speech peaks

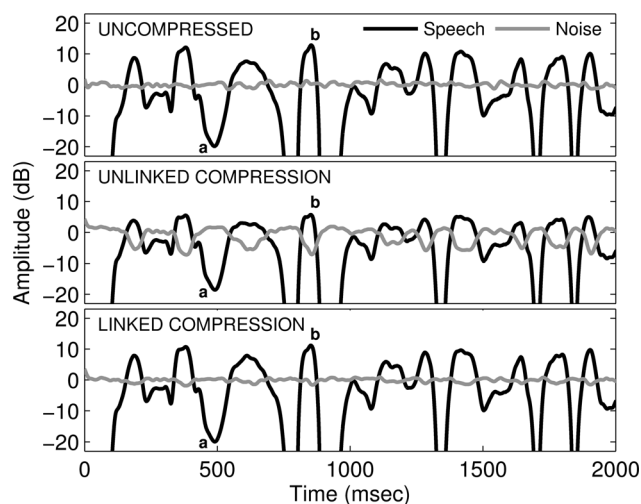


FIG. 3. Envelopes of an extract of the speech and noise signals in the high-frequency channel (2 to 5 kHz) at the better ear. The nominal source SNR was  $-2$  dB. The data in each panel were normalized so that the overall rms level of the noise was 0 dB. At point a, which marks a dip in the speech envelope, the level of the speech is the same (to within 1 dB) for all three processing conditions because the compressors’ behavior is dominated by the steady noise at such moments. At point b, marking a peak in the speech envelope, linked compression reduces the speech level by 2 dB compared to the uncompressed condition, and unlinked compression by 7 dB. This “penalizing” of the speech peaks by compression causes a reduction in the long-term apparent SNR, even though the instantaneous SNR is at all times unaffected by the processing.

to be “penalized” such that they barely rise above the average noise level (Naylor and Johannesson, 2009). Thus, although the instantaneous SNR is unaffected (because the compressor must apply the same gain to both speech and noise at any one instant), the apparent long-term SNR is reduced.

It is important to note that the changes in apparent SNR introduced by compression are not equivalent to changes in SNR brought about by adjusting the level of speech or noise at the source. However, there is some evidence, reviewed by Naylor and Johannesson (2009), to suggest that the changes in apparent SNR introduced by compression are perceptually relevant for intelligibility. The above analysis suggests that linked compression may offer an advantage over unlinked compression in this particular listening scenario through an improvement in the apparent long-term SNR at the better ear. The improvement in apparent SNR is larger at high frequencies than at low, and is dependent on the relative level of the speech and noise signals arriving at the listener’s ears.

### E. Disturbance to ILDs

One of the primary advantages claimed for linking compression across the ears is that it allows natural ILD cues to be preserved (Kreisman *et al.*, 2010; Wiggins and Seeber, 2011). The linked compression system tested here preserves ILDs because it forces identical gain to be applied at both ears at all moments. In contrast, the unlinked compression system can apply a different amount of gain at the two ears at any moment. Figure 4 plots the interaural gain difference (IGD, the momentary difference between the gain applied at each ear) for the unlinked compression system in the high-frequency channel. Results are shown for nominal source SNRs of  $+4$  and  $-10$  dB. The IGD is generally not equal to zero, meaning that the unlinked compression system will not preserve ILDs. During pauses in the speech, the IGD tends toward a value of about  $-8$  dB. At such moments, the behavior of both compressors is determined by the steady noise: Because the level of the noise is greater at the right ear than the left, the compressor at the right ear applies less gain, and

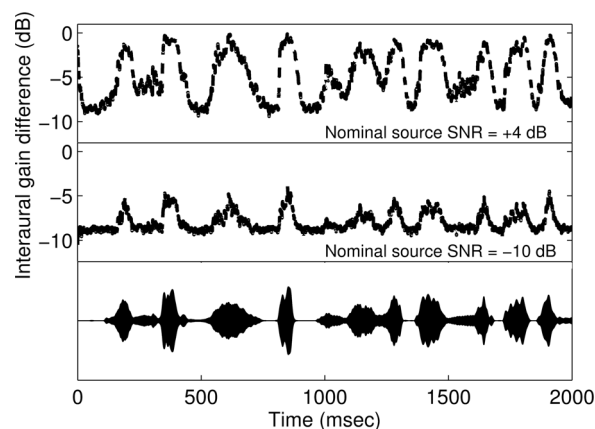


FIG. 4. IGD (momentary difference in the gain applied at the right and left ears) plotted against time in the high-frequency channel following unlinked compression. The nominal source SNR was  $+4$  dB for the top panel and  $-10$  dB for the middle panel. The bottom panel shows the envelope of the original speech signal in the high-frequency channel for reference.

hence the IGD is negative. During bursts of speech energy, the IGD tends back toward zero because the speech, which comes from directly in front, arrives with an equal level at both ears. At a nominal source SNR of +4 dB, the speech dominates the compressors' behavior at such moments, and so the IGD fully returns to zero. At a nominal source SNR of -10 dB, the noise still exerts some influence on the compressors' behavior even during peaks in the speech, and so the IGD only partially returns toward zero. Consequently, the overall magnitude of the fluctuations in IGD across time is greater for a nominal source SNR of +4 than -10 dB.

Wiggins and Seeber (2011, 2012) found that while a static offset in ILD introduced by unlinked compression could sometimes affect the perceived lateral position of a sound, it was dynamic fluctuations in ILD that were particularly salient and detrimental to accurate spatial perception. In the present study, we therefore sought to quantify not the overall offset in gain between the ears but rather the size of the fluctuations in IGD across time. To do this, we calculated the standard deviation of the IGD trace. Figure 5 shows the IGD standard deviation plotted against nominal source SNR. Consistent with the observations made above, fluctuations in IGD generally increased with increasing nominal source SNR. In the high-frequency channel, IGD fluctuations peaked at a nominal source SNR of +4 dB. IGD fluctuations were larger in the high- than in the low-frequency channel, consistent with earlier statements about the unlinked left- and right-ear compressors behaving more similarly in the low-frequency channel because of generally smaller ILDs at low frequencies.

The above analysis demonstrates that, for this particular listening scenario, the unlinked compression system causes substantial disturbance to ILDs. Therefore, to the extent that preservation of ILDs is important for speech intelligibility in this situation, unlinked compression would be expected to have a deleterious effect compared to both the uncompressed and linked compression conditions. Any disadvantage of unlinked compression would be expected to be most apparent at small positive nominal source SNRs.

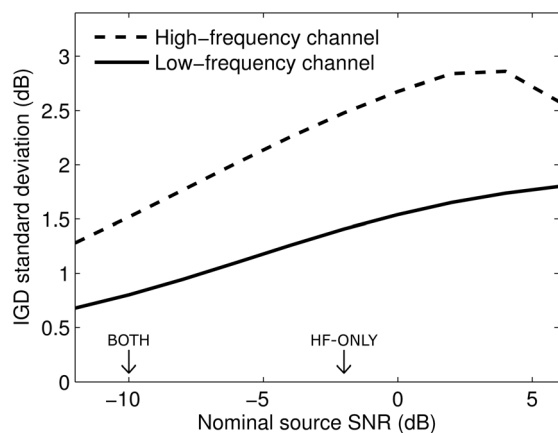


FIG. 5. Standard deviation of the IGD plotted against nominal source SNR in the low-frequency (solid line) and high-frequency (dashed line) channels. This provides a measure of the magnitude of the dynamic changes to ILDs introduced by unlinked compression. The vertical arrows indicate the nominal source SNRs tested in the speech intelligibility experiment (cf. Fig. 2 caption).

### III. SPEECH INTELLIGIBILITY EXPERIMENT

#### A. Overview

Speech intelligibility was measured to determine if and how the physical effects of compression on the signals arriving at the listener's ears described above affect perception. Normal-hearing participants listened to virtual-acoustic stimuli presented over headphones. Intelligibility was measured at fixed nominal source SNRs, rather than adaptively varying the SNR to determine a speech reception threshold because we wanted the behavior of the compressors to be identical for all participants. To assess the importance of the availability of low-frequency timing cues, listeners heard either both frequency channels together or the high-frequency channel on its own. Participants were tested in conditions of binaural listening and monaural listening to the better-ear signal only to distinguish any effects related specifically to binaural processing from those associated with changes to the signal at the better ear.

#### B. Methods

##### 1. Participants

Ten participants took part in the experiment (mean age 21 yrs, range 18 to 26 yrs, 6 females), all having audiometric thresholds  $\leq 20$  dB hearing level (HL) at octave frequencies between 125 Hz and 8 kHz, except for 2 participants who had a threshold of 25 dB HL at 8 kHz at one ear. All participants were native speakers of British English and none of them had previously heard the speech material used for testing. Approval for the study was granted by the Ethics Committee of the School of Psychology at the University of Nottingham. Participants were paid and typically attended a single session lasting 3 h.

##### 2. Equipment

Stimuli were prepared and the experiment administered using MATLAB (The MathWorks, Natick, MA). Signals were transferred digitally from a computer soundcard to a custom-built headphone amplifier with 24-bit digital-to-analog converters and presented through Sennheiser HD 600 headphones. Participants sat in a double-walled sound-attenuating booth and verbally submitted their responses to the experimenter through an intercom system.

##### 3. Test material

A total of 480 IEEE sentences from the corpus recorded by Stacey and Summerfield (2007) were used as the target speech. We divided the material into 30 lists of 16 sentences, with each list containing the same number of sentences from each talker (1 sentence from 8 of the 12 talkers and 2 sentences from the remaining 4). Each sentence contained five keywords. A 60-s steady speech-shaped noise with a spectrum equal to the long-term average spectrum of the speech material was generated as described in Sec. II C. For each trial, a 4.5 s portion of the noise was extracted at random. The test sentence was temporally aligned in the center of the noise. The maximum sentence duration was 3.5 s, meaning that the

noise began at least 500 ms before the sentence and continued for at least 500 ms afterwards.

#### 4. Signal processing

The simulated configuration of speech and noise sources, HRTF filtering, and compression systems were all as described in Sec. II (see Fig. 1). To summarize: The speech came from directly in front and the noise from 60° azimuth. The HMS II acoustic manikin HRTFs from the AUDIS catalog were used (Blauert *et al.*, 1998), equalized for use with headphones having a diffuse-field response. All processing was performed independently in a low- (100 to 2000 Hz) and a high-frequency channel (2000 to 5000 Hz). In an uncompressed reference condition, there was no further processing, although the same filtering into two frequency channels was performed and so the speech was band-limited to the range 100 to 5000 Hz. Two fast-acting compression systems were tested: Unlinked and linked compression. The compression threshold was 50 and 35 dB SPL in the low- and high-frequency channels, respectively; the compression ratio was 3:1 in both channels; attack and release times were 5 and 60 ms, respectively (ANSI, 2003).

Because the compression reduced overall levels, we matched the long-term rms level (summed across the ears) before and after compression. This was done separately within each frequency channel. Because the level matching was performed based on the summed level across the ears, identical gain was applied at both ears and ILDs were not affected. Subjectively, this level-matching procedure removed any clearly noticeable differences in the overall spectral shape or loudness between processing conditions.

In the “both-channels” condition the two frequency channels were recombined, while in the “high-frequency-channel-only” condition the high-frequency channel was presented in isolation. All stimuli were amplified to an overall level of approximately 65 dB SPL for presentation, with the constraint that identical amplification was applied at both ears to avoid disrupting ILDs. Stimuli were faded in and out with 500 ms raised-cosine ramps, which affected the leading and trailing noise but not the target speech. For the monaural listening condition, only the signal at the better ear (the ear on the opposite side to the noise source) was presented; the signal at the worse ear was muted.

#### 5. Experimental conditions and procedure

All participants were tested in two bandwidth conditions (both-channels and high-frequency-channel-only), two listening modes (binaural and monaural), and three processing conditions (uncompressed, unlinked compression, and linked compression), giving a total of 12 experimental conditions. Two 16-sentence lists were randomly assigned to each experimental condition, independently for each participant. The order in which each participant heard the two bandwidth conditions and the two listening modes was randomized. The experiment was split into eight runs, with bandwidth condition and listening mode fixed within a run. Each run comprised one sentence list for each of the three processing conditions. The order in which the processing conditions

were tested was randomized for each run, although the processing condition was fixed for the duration of each list to allow participants to employ a consistent listening strategy from one trial to the next.

On each trial, the listener heard the test sentence once and then verbally repeated what they heard to the experimenter. No repeats were allowed. The experimenter scored the number of keywords correctly identified. An additional practice trial was included at the start of each run to acclimatize the participant to the upcoming condition. Practice trials were not scored and used sentences from lists that were not used for testing.

The nominal source SNR was  $-10$  dB for the both-channels condition and  $-2$  dB for the high-frequency-channel-only condition. These values were chosen on the basis of pilot testing to give scores centered on 50% correct and to avoid ceiling or floor effects. Odd-numbered participants were tested with the simulated noise source at an azimuth of  $-60^\circ$ , and even-numbered participants with the noise source at  $+60^\circ$ .

#### 6. Training

All participants undertook a training session lasting approximately 1 h before testing began. Training comprised four runs with one sentence list per run. None of the training lists were used for testing. For all participants, training runs followed this order: Both-channels binaural, both-channels monaural, high-frequency-channel-only binaural, and high-frequency-channel-only monaural. Within each run, the processing condition was varied randomly from trial to trial to ensure that participants heard a representative sample of all test conditions. For the first 2 trials in each run the nominal source SNR was elevated by 8 dB to make the task easier; for the next 2 trials it was elevated by 4 dB, and for the final 12 trials it was set to the same value as used in testing. During training, if the participant gave an incorrect response on the first presentation, the stimulus was replayed. If the response was still incorrect after a second presentation, the experimenter told the participant the correct answer and the stimulus was replayed one final time.

#### C. Results

The percentage of keywords correctly identified was calculated across the 16 sentences in each list. Because we did not know in advance whether all sentence lists would be equally difficult, we checked for any consistent across-list differences in performance as follows: A mean percent-correct score was calculated across all participants for each experimental condition (combination of bandwidth condition, listening mode, and processing condition). Then, the mean score for the relevant experimental condition was subtracted from each individual sentence-list score. This removed the mean experimental effects so that relative sentence-list scores obtained in different experimental conditions could be compared to assess overall list difficulty. Figure 6 shows the relative score for each list. A repeated-measures analysis of variance (RM-ANOVA) confirmed that there was a significant effect of list,  $F(23, 207) = 5.57$ ,  $p < 0.001$ . Thus, although each list contained the same number of sentences from each

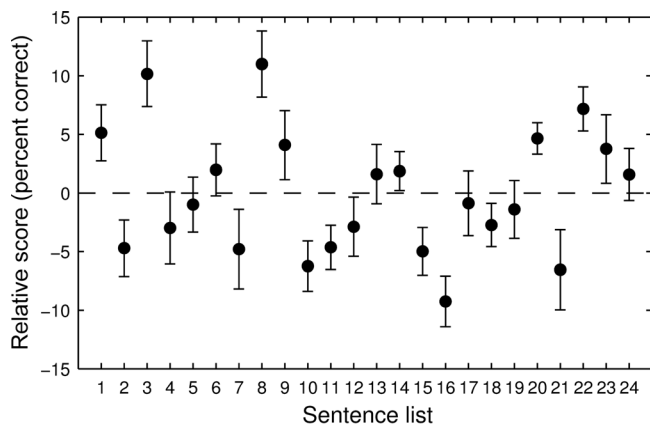


FIG. 6. Overall differences in performance across individual sentence lists after removing the mean experimental effects. The relative percent-correct score is plotted for each list (mean  $\pm$  1 standard error). Positive (negative) values indicate that a particular list was harder (easier) than the average.

talker, some lists were easier than others. To correct for this, the mean relative score for the relevant list was subtracted from each original sentence-list score. This removed “noise” in the data set arising from unintended differences in difficulty across lists: The procedure achieved a small but consistent reduction in variance while having a minimal effect on the mean score for each experimental condition.

To check for a possible training effect during testing, a four-way RM-ANOVA was conducted with factors bandwidth condition, listening mode, processing condition, and repetition. The main effect of repetition was significant,  $F(1,9) = 5.53$ ,  $p = 0.043$ , indicating that there was an overall difference in scores between the first and second sentence lists tested in each experimental condition. However, the mean improvement from the first to the second list was small (1.7 percentage points), and so scores were averaged across the two lists for subsequent analysis. A further ANOVA was run to check for any effect of which side the simulated noise source was on, and therefore which ear was tested monaurally. Bandwidth condition, listening mode, and processing condition were included as within-subject factors, and side-of-noise-source as a between-subjects factor. The main effect of side-of-noise-source was non-significant,  $F(1,8) = 0.09$ ,  $p = 0.779$ , and the mean overall score for noise-on-the-left versus noise-on-the-right differed by only 0.7 percentage points. The side of the noise source was therefore not considered further.

The percent-correct scores (corrected for list difficulty as described above) for each experimental condition are plotted in Fig. 7. As different nominal source SNRs were tested in the both-channels and high-frequency-channel-only conditions, it is not appropriate to directly compare scores between them. Therefore, separate RM-ANOVAs were conducted, each including factors listening mode and processing condition. Considering first the both-channels condition, the main effect of listening mode was significant,  $F(1,9) = 249.81$ ,  $p < 0.001$ . The mean score was 28.7 percentage points higher for binaural than for monaural better-ear listening. The main effect of the processing condition was also significant,  $F(2,18) = 21.09$ ,  $p < 0.001$ . Bonferroni-corrected pairwise comparisons revealed a significant difference between

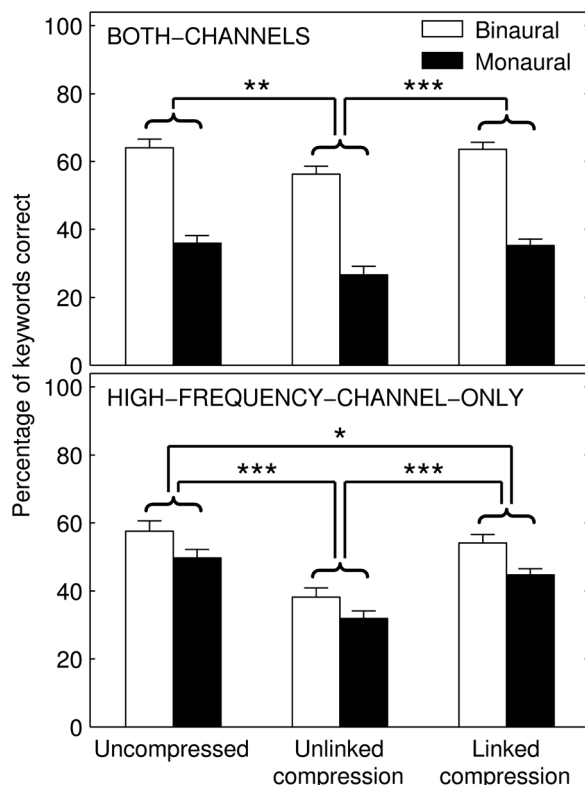


FIG. 7. Mean percent-correct score across the ten participants for each experimental condition. Error bars indicate one standard error and asterisks indicate significant differences between processing conditions (\*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ ).

the uncompressed and unlinked compression conditions ( $p = 0.004$ ), and between unlinked and linked compression ( $p = 0.001$ ) but not between the uncompressed and linked compression conditions ( $p = 1.0$ ). Performance was about 8 percentage points worse in the unlinked compression condition than in both the uncompressed and linked compression conditions. The interaction between listening mode and processing condition was non-significant,  $F(2,18) = 0.15$ ,  $p = 0.859$ .

In the high-frequency-channel-only condition, there was a significant main effect of listening mode,  $F(1,9) = 47.05$ ,  $p < 0.001$ , with scores being on average 7.8 percentage points higher for binaural than for monaural listening. The main effect of processing condition was also significant,  $F(2,18) = 86.38$ ,  $p < 0.001$ , and Bonferroni-corrected pairwise comparisons confirmed that all three processing conditions differed significantly from one another (all at  $p < 0.05$ ). The mean score was highest in the uncompressed condition, 4.2 percentage points lower in the linked compression condition, and a further 14.4 percentage points lower in the unlinked compression condition. There was no significant interaction between listening mode and processing condition,  $F(2,18) = 0.67$ ,  $p = 0.527$ .

The absence of any significant interaction between listening mode and processing condition suggests that the binaural benefit obtained from adding the ear with the less favorable SNR was not affected by the processing. This can be seen more clearly in Fig. 8 in which binaural squelch (binaural performance minus monaural performance) is plotted. A RM-ANOVA with factors bandwidth condition and

processing condition confirmed that there was no significant main effect of processing condition on binaural squelch,  $F(2,18)=0.13$ ,  $p=0.880$ , nor a significant interaction between processing condition and bandwidth condition,  $F(2,18)=0.70$ ,  $p=0.509$ . There was, as expected, a significant main effect of bandwidth condition,  $F(1,9)=115.04$ ,  $p<0.001$ . Binaural squelch was on average 20.9 percentage points greater in the both-channels condition than in the high-frequency-channel-only condition, the difference presumably resulting from the availability of low-frequency interaural timing information in the both-channels condition (Bronkhorst, 2000).

## IV. DISCUSSION

### A. The benefit of linking compression across the ears

Speech intelligibility for normal-hearing listeners was tested in the presence of a spatially separated steady noise with no compression and with unlinked or linked bilateral compression. Performance was best with no compression but was as good (both-channels condition) or nearly as good (within approximately 4 percentage points, high-frequency-channel-only condition) with linked compression. Performance was 8 to 14 percentage points better with linked than with unlinked compression. The advantage of linked compression was the same for monaural and binaural listening. This indicates that linked compression had its beneficial effect through changes to the signal at the better ear, rather than through the preservation of ILDs. Indeed, in this task, disruption to ILD cues did not harm intelligibility: The benefit of binaural over monaural listening was no smaller with unlinked compression (which disturbed ILDs) than with linked or no compression (both of which preserved ILDs).

How did linked compression change the signal at the better ear to improve intelligibility compared to unlinked compression? Figure 2 shows that linked compression gave a more favorable apparent SNR (the long-term ratio of speech-to-noise energy at the output of the compressor) at the better ear. This is because the bilateral link allowed the strong noise signal at the worse ear to influence the behavior of the compressor at the opposite ear, producing a “noise suppression” effect at the better ear (cf. Kates, 2008). The pattern of results in the experiment was qualitatively consistent with the changes in apparent SNR at the better ear: Unlinked compression reduced the apparent SNR substan-

tially compared to the uncompressed reference (cf. Fig. 2), and measured intelligibility was substantially worse with unlinked compression than with no compression; linked compression, in contrast, reduced the apparent SNR to only a small degree, and measured intelligibility was nearly as good with linked compression as with no compression. Thus, it seems that the apparent SNR at the better ear is a perceptually relevant measure. However, an alternative way of viewing the results is that linked compression disturbed the signal envelopes at the better ear less than unlinked compression (cf. Fig. 3), and so deleterious effects associated with changes to the speech envelope may therefore have been reduced by linking the compressors (Plomp, 1988; Stone and Moore, 2007). In this respect, linked compression was beneficial in this speech-in-steady-noise scenario because it behaved more linearly in practice. Note that, in principle, a monaural device might achieve the same effect either by linearizing the input-output function or by switching to slower compression time constants in the presence of a steady background noise.

Why did the disturbance to ILDs caused by unlinked compression not affect performance? Particularly in the high-frequency-channel-only condition an effect might have been expected, given the increased salience of the ILD as a localizing cue for high-pass sounds (Macpherson and Middlebrooks, 2002). At the nominal source SNR tested in the high-frequency-channel-only condition ( $-2$  dB), the IGD standard deviation in the high-frequency channel was 2.5 dB. While this is slightly smaller than the peak value of 2.9 dB, which occurred at a higher nominal source SNR (cf. Fig. 5), it nonetheless demonstrates that ILD cues were disturbed by unlinked compression in the specific condition tested. The most probable explanation for why ILD disturbance had no discernible effect is that the binaural interaction benefit obtained in this type of speech-in-steady-noise scenario depends on binaural decorrelation cues that are conveyed primarily through ITDs (Bronkhorst, 2000): Compression does not directly affect ITDs (Musa-Shufani *et al.*, 2006; Kalluri and Edwards, 2007). It remains possible that by disturbing ILDs unlinked compression may have affected the perceived location of the sources, especially in the high-frequency-channel-only condition where low-frequency ITD cues were unavailable (Wiggins and Seeber, 2011, 2012). However, even if such effects did occur, there is evidence that perceived location does not play a critical role in speech understanding (Bronkhorst, 2000; Edmonds and Culling, 2005), unless there is a paucity of other cues by which the target speech can be segregated from the interfering sound (Freyman *et al.*, 1999). This was not the case for the steady noise masker used here.

### B. Modeling speech intelligibility for the normal-hearing listeners

The measured intelligibility differences between processing conditions were qualitatively consistent with changes in apparent SNR at the better ear. However, it is not immediately clear how apparent SNR should be combined across frequency channels to predict overall intelligibility. We investigated whether the experimental results could be modeled

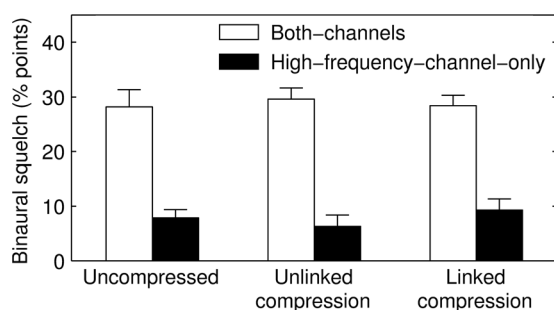


FIG. 8. Mean binaural squelch (binaural performance minus monaural performance) across the ten participants for each experimental condition. Error bars indicate one standard error.



using the  $I_3$  measure (Kates and Arehart, 2005).  $I_3$  is a coherence-based extension to the Speech Intelligibility Index (SII: ANSI, 1997), making it applicable to speech disrupted not only by additive noise but also by nonlinear distortions such as those generated by compression. In calculating  $I_3$ , the original speech is segmented into high-, mid-, and low-level amplitude regions and the coherence-based SII is calculated separately for each region. Figure 3 shows that the change in apparent SNR caused by compression primarily results from a reduction in the relative level of the speech peaks. However, the peaks typically correspond to steady vowel sounds occurring in syllable nuclei (Greenberg, 2006), which may carry less information than the weaker consonants and vowel-consonant transitions (Yoo *et al.*, 2005). By considering intelligibility in three amplitude regions,  $I_3$  is in principle able to account for the effects of compression on all the information-carrying components of the speech signal, not just the peaks.

Intelligibility was predicted using  $I_3$  for the experimental conditions based on the signals arriving at the better ear only. We predicted monaural intelligibility in this way because the SII procedure currently cannot be applied to binaural listening with spatially separated sources. However, the absence of any effect of processing condition on binaural squelch implies that, in this particular listening scenario, binaural performance can be considered equal to monaural performance plus a fixed value that does not depend on processing condition. Evaluating monaural intelligibility is therefore sufficient to assess any differences between processing conditions. The speech material was a concatenation of 16 IEEE sentences as described in Sec. II C. The  $I_3$  calculation was made in accordance with Kates and Arehart (2005) and the standard SII procedure (ANSI, 1997). We used Kates and Arehart's (2005) Eq. (14) without modification to combine the coherence-based SII values for the three amplitude regions: This formula applies the greatest weight to the mid-level region, a moderate weight to the low-level region, and zero weight to the high-level region, and so changes solely to the peaks of the speech signal would have had minimal influence on predicted intelligibility.

Figure 9 compares the  $I_3$  predictions of monaural intelligibility (lines, left axis) to measured intelligibility (symbols, right axis). The  $I_3$  model captures the measured differences between processing conditions well, although it does predict slightly larger differences in the high-frequency-channel-only condition than were measured experimentally (at least at the particular nominal source SNR tested). Note that we applied a small correction to the  $I_3$  predictions before plotting (+0.08 in the both-channels condition; -0.06 in the high-frequency-channel-only condition) to calibrate the model to the overall level of performance measured in each bandwidth condition: Such discrepancies in the prediction of absolute performance might result from differences in task, speech material, and listeners between the present study and that to which the  $I_3$  model was originally fitted (Kates and Arehart, 2005).

Note that the maximum predicted benefit of linked over unlinked compression in Fig. 9 is similar in the both-channels and high-frequency-channel-only conditions (corresponding to a difference in  $I_3$  of about 0.17) but it occurs at

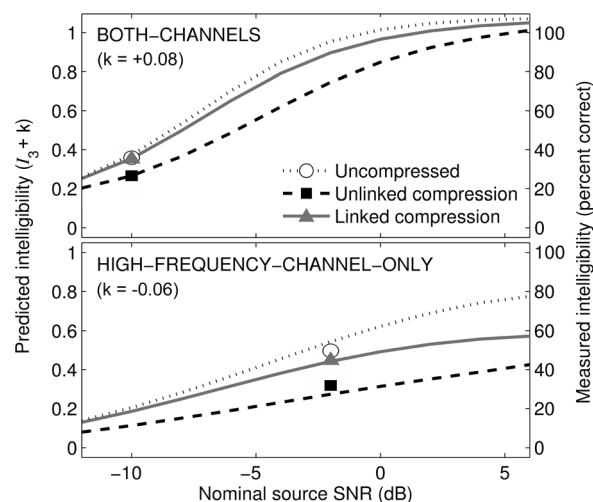


FIG. 9. Comparison of predicted (lines, left axis) and measured (symbols, right axis) speech intelligibility for monaural listening to the ear with the better SNR. A correction was applied to the  $I_3$  values (+0.08 in the both-channels condition; -0.06 in the high-frequency-channel-only condition) to calibrate the model to the overall level of performance measured in each bandwidth condition.

different nominal source SNRs (-4 and +2 dB, respectively). Thus, at a more positive nominal source SNR than was tested in the experiment, linked compression would be predicted to offer a greater benefit in the both-channels condition than was measured here. This is significant in that the both-channels condition is more representative of real-world listening than the high-frequency-channel-only condition.

### C. Speech intelligibility predictions for a hypothetical bilateral hearing-aid user

The present experiment demonstrates a benefit of linked over unlinked compression for normal-hearing listeners. However, since the study is motivated by the use of compression in hearing devices, it is important to consider how intelligibility might be affected for hearing-impaired listeners using such devices. In the present study, audibility was not an issue, both because of the participants' good hearing sensitivity and the way in which rms levels were equated before and after compression. Linked compression was advantageous because it resulted in less disruption to the speech and noise envelopes at the better ear, and, in turn, provided a better apparent SNR at this ear compared to unlinked compression. In a real device, without the offline level-matching procedure used here, this increase in envelope fidelity would be achieved at the cost of reduced overall gain at the better ear. Thus, it is possible that the benefit of linked compression may not be realized by a hearing-impaired listener because of a counteracting reduction in audibility.

In Sec. IV B, it was shown that the  $I_3$  model was able to capture the main patterns in the normal-hearing listeners' results. The  $I_3$  procedure was developed with the aim of predicting intelligibility for hearing-aid users: Kates and Arehart (2005) validated  $I_3$  predictions against measured intelligibility for nine hearing-impaired listeners for speech corrupted by additive noise, peak-clipping distortion, or

center-clipping distortion, and Kates (2010) used  $I_3$  to model the effects of dynamic-range compression in hearing aids. The strength of the model is that, in principle, it can simultaneously account for the two main effects of compression: The beneficial improvement in audibility of low-level speech components and the potentially detrimental effects of envelope distortions.

We used  $I_3$  to predict speech intelligibility for a hypothetical hearing-impaired listener fitted bilaterally with compressive hearing aids that were either unlinked or linked. As before, intelligibility was predicted for monaural listening to the ear with the better SNR but, for the reasons discussed above, any differences between compression systems seen monaurally would be expected to occur for binaural listening also. A symmetrical hearing loss with no conductive element was assumed with an audiogram corresponding to the  $N_3$  “Moderate” loss defined in Bisgaard *et al.* (2010): This was the standard audiogram that characterized the highest proportion of 28 244 measured ears. The CAMEQ method (Moore *et al.*, 1999a) was used to prescribe compression parameters and insertion gains for a hypothetical 12-channel hearing aid with an overall bandwidth of 125 to 5750 Hz. Compression thresholds were set equal to the level evoked in each channel by a steady noise with the long-term average speech spectrum defined in Byrne *et al.* (1994) presented from directly in front at an equivalent free-field level of 50 dB SPL. Attack and release times in all channels were 5 and 60 ms, respectively (ANSI, 2003). The linked compression system operated as described in Sec. II B, but in 12 frequency channels rather than 2. The audiogram, channel edge frequencies, compression thresholds, and prescribed compression ratios and insertion gains are given in Table I.

A simulation similar to that described in Sec. II was run but including this hypothetical hearing loss and amplification prescription. The use of the HMS II acoustic manikin HRTFs

from the AUDIS catalog (Blauert *et al.*, 1998) was assumed to give signals representative of those that might be received at the microphones of in-the-canal hearing aids. The CAMEQ prescription was implemented in such a way that the target levels were achieved at the eardrums of the hypothetical listener. In calculating  $I_3$ , the methods of Sec. 5.3 of ANSI S3.5 were used to convert these levels at the eardrum to the equivalent free-field levels required in the SII calculation procedure (ANSI, 1997).

Figure 10 plots predicted intelligibility against nominal source SNR for the hypothetical hearing-aid user listening monaurally to the ear with the better SNR, for unlinked and linked compression. No “uncompressed” condition was included because a different type of gain prescription would have been required for linear amplification, complicating a comparison between linear and compressive amplification. Predicted intelligibility is higher for linked than for unlinked compression across all nominal source SNRs. The greatest benefit of linked over unlinked compression occurs at a nominal source SNR of  $-4$  dB and is a difference in  $I_3$  of 0.11. On the assumption that the absence of any effect of processing condition on binaural squelch observed for normal-hearing listeners in the experiment would hold for hearing-impaired listeners too, a similar benefit of linked over unlinked compression would be predicted for binaural listening. Thus, in this particular case at least, the benefit of linked compression is predicted to not be counteracted by a reduction in audibility associated with the lower overall gain applied by the linked system.

These results suggest that linked compression may offer bilateral hearing-aid users a small to moderate advantage over unlinked compression for understanding speech in the presence of a spatially separated noise. However, further work is needed to confirm that the predicted benefit of linked compression exists for real hearing-impaired listeners, including listeners with audiograms and gain prescriptions differing

TABLE I. Details of the hearing loss, hearing-aid configuration, and gain prescription used to predict speech intelligibility for a hypothetical hearing-impaired listener.

Audiogram										
Frequency (Hz)	250	375	500	750	1000	1500	2000	3000	4000	6000
dB HL	35	35	35	35	40	45	50	55	60	65
Hearing-aid configuration and gain prescription										
Channel no.	Geometric center frequency (Hz)		Lower edge frequency (Hz)		Upper edge frequency (Hz)		CT	CR	IG50	IG80
1	177		125		250		46	1.53	10.2	0.0
2	433		250		750		47	1.11	10.3	7.4
3	968		750		1250		40	1.99	20.8	5.9
4	1479		1250		1750		30	1.89	20.5	6.4
5	1984		1750		2250		27	1.98	22.9	8.0
6	2488		2250		2750		28	2.01	21.7	6.6
7	2990		2750		3250		27	1.99	20.7	5.8
8	3491		3250		3750		24	1.99	21.4	6.5
9	3992		3750		4250		20	2.01	22.2	7.1
10	4493		4250		4750		22	2.22	21.6	5.1
11	4994		4750		5250		21	2.51	20.8	2.8
12	5494		5250		5750		22	2.80	20.2	1.0

CT, compression threshold (dB SPL); CR, compression ratio; IGxx, insertion gain (dB) for a sinusoid at the channel center frequency with an input level of 50 or 80 dB SPL.

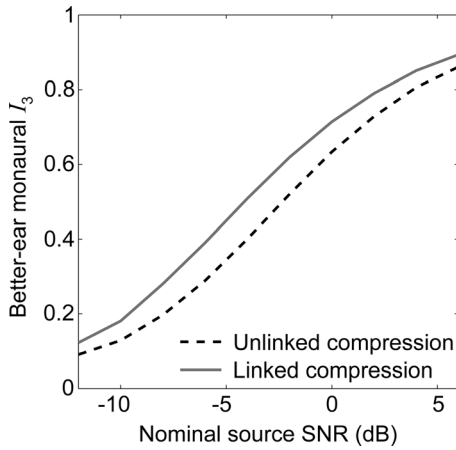


FIG. 10. Predicted intelligibility ( $I_3$ ) for monaural listening to the ear with the better SNR for a hypothetical hearing-impaired listener wearing bilateral in-the-canal hearing aids performing either unlinked (dashed line) or linked (solid line) compression.

from those assumed here. Further points worth noting are that (1) a considerable variation may be expected in how individual hearing-impaired listeners respond to fast-acting compression (Gatehouse *et al.*, 2006; Moore, 2008); (2) the benefit of linked compression may be smaller for behind-the-ear aids because the attenuation of a lateral source at the contralateral ear measured through such devices can be smaller than is measured through microphones placed in or close to the ear canals (Festen and Plomp, 1986)—this would result in more similar input levels to the left- and right-ear compressors, and therefore less difference between linked and unlinked compression; (3) the linked compression system tested in the present study assumed zero delay in transmitting information from one ear to the other—in a real wireless system, a certain amount of latency would be unavoidable.

#### D. Extension to fluctuating maskers

The present experiment demonstrates a benefit of linked compression for listening to speech in the presence of a spatially separated steady noise. However, performance with linked compression was never better than without compression: Linked compression was better than unlinked compression because it behaved more like linear amplification in a scenario in which compression could only reduce the apparent SNR. When the interfering sound has a fluctuating rather than steady envelope, fast-acting compression can potentially improve the apparent SNR by applying more gain to the speech during temporal dips in the masker (Olsen *et al.*, 2005; Naylor and Johannesson, 2009). We re-ran the simulation described in Sec. II but with the steady noise replaced by background speech. Like the target speech, the background speech comprised 16 concatenated IEEE sentences, although different sentences were used for target and background. Since there are now two speech sources, we refer to “target-to-background ratio” (TBR) rather than SNR. Figure 11 plots apparent TBR at each ear against nominal source TBR. At negative nominal source TBRs, where the target speech is lower in level than the background speech,

compression generally improves the apparent TBR at both ears. At the better ear, the apparent TBR is consistently higher for linked than for unlinked compression across the entire range of nominal source TBRs considered, with the improvement being greater in the high- than in the low-frequency channel.

To the extent that an improvement in apparent long-term TBR at the better ear corresponds to better intelligibility, these results suggest that linked compression will be advantageous for understanding speech in the presence of a spatially separated fluctuating interferer, such as competing speech. It is possible, however, that apparent TBR might not be the most perceptually relevant measure in a speech-in-speech scenario, where informational masking may play an important role (Kidd *et al.*, 2008). Linking compression across the ears will mean that, if the target speech is dominant at one ear and the background speech at the other, the background speech will exert a greater influence on the gain-control signal for the better ear. This may exacerbate a potentially deleterious effect associated with compressing a mixture of target and background speech, which is that the two sources acquire a common component of modulation after compression, making them harder to segregate (Stone and Moore, 2007). It should also be noted that some studies (Moore *et al.*, 1999b; Olsen *et al.*, 2004) suggest that hearing-impaired listeners are less able than normal-hearing listeners to benefit from SNR improvements provided by compressive relative to linear amplification in fluctuating backgrounds, possibly because of supra-threshold deficits such as a reduced ability to make use of temporal fine structure information (Moore, 2008).

In a speech-in-speech scenario, it is also possible that disruption to ILDs caused by unlinked compression may have a non-negligible impact on intelligibility: If the target and background speech are similar and therefore easily

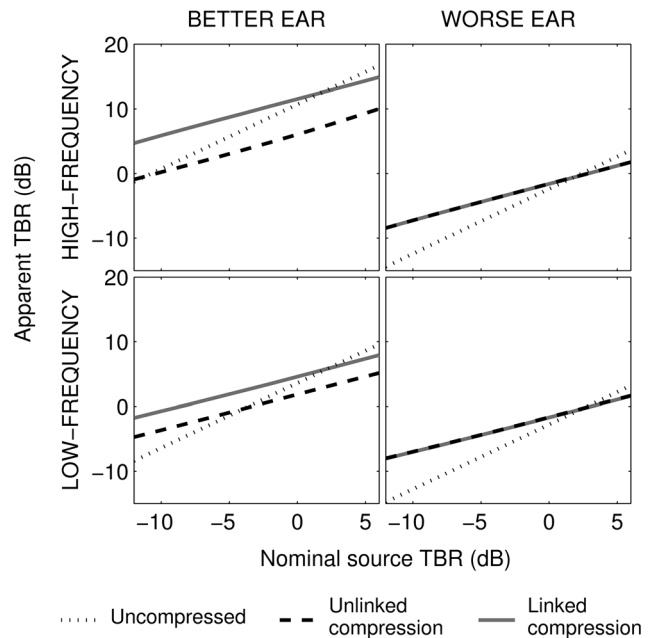


FIG. 11. As in Fig. 2 but for a two-talker scenario where target speech comes from directly in front and background speech (rather than noise) from an azimuth of  $60^\circ$ . The apparent long-term TBR after compression is plotted against the nominal source TBR.

confusable, perceived spatial separation may play a role in allowing the target speech to be understood (Freyman *et al.*, 1999). If the disruption to ILDs caused by unlinked compression reduces the perceived spatial separation between target and background, then linked compression may confer an additional benefit through the preservation of ILDs. It does seem, however, that if accurate low-frequency ITD cues are available then the impact of unlinked compression on perceived location may be small (Musa-Shufani *et al.*, 2006; Keidser *et al.*, 2006; Wiggins and Seeber, 2011).

## V. CONCLUSIONS

Speech intelligibility in steady noise was tested in normal-hearing listeners for three processing conditions: No compression, unlinked bilateral compression, and linked compression. The speech came from directly in front and the noise from an azimuth of either  $-60^\circ$  or  $+60^\circ$ . Stimuli were spatialized by filtering with HRTFs and presented over headphones. Compression was applied offline, independently in low- and high-frequency channels, and was fast-acting with a 3:1 ratio.

Mean performance was 8 to 14 percentage points better with linked than with unlinked compression. Performance was best with no compression but was either as good or nearly as good (within approximately 4 percentage points) with linked compression. The differences between processing conditions were similar regardless of whether participants listened binaurally or monaurally to the ear with the better SNR. This implies that the differences between processing conditions were driven by the way in which the speech and noise signals at this ear were altered, not by the preservation of binaural cues. Indeed, disturbance to ILDs caused by unlinked compression did not harm performance, probably because the binaural benefit obtained in this type of speech-in-steady-noise scenario depends on binaural decorrelation cues conveyed primarily by ITDs.

The differences between processing conditions were qualitatively consistent with the effects of compression on the apparent long-term SNR at the better ear: Both types of compression tended to reduce the apparent SNR but linked compression reduced it less than unlinked compression. The benefit of linked over unlinked compression can be expected to depend strongly on the relative level of the speech and noise signals arriving at the listener's ears. Linked compression can provide a similar improvement in apparent SNR in the case of a fluctuating interferer, such as competing speech, although it is yet to be tested whether this leads to improved intelligibility. Speech intelligibility predictions made using the  $I_3$  model for a hypothetical listener with a moderate hearing loss using bilateral compressive hearing aids suggest that linked compression could potentially provide an improvement in intelligibility of up to about 10 percentage points in the case of a spatially separated steady noise.

## ACKNOWLEDGMENTS

This work was funded through the intramural program of the Medical Research Council (UK). We thank Katharina Egger who assisted us with preliminary work on the topic

while on a student placement sponsored by Action on Hearing Loss, and Paula Stacey and Quentin Summerfield for permission to use the speech material. We also thank the two anonymous reviewers for their helpful suggestions.

- ANSI (1997). ANSI S3.5-1997, *Methods for the Calculation of the Speech Intelligibility Index* (American National Standards Institute, New York).
- ANSI (2003). ANSI S3.22-2003, *Specification of Hearing Aid Characteristics* (American National Standards Institute, New York).
- Bisgaard, N., Vlaming, M. S. M. G., and Dahlquist, M. (2010). "Standard audiograms for the IEC 60118-15 measurement procedure," *Trends Amplif.* **14**, 113–120.
- Blauert, J., Brueggen, M., Bronkhorst, A. W., Drullman, R., Reynaud, G., Pellioux, L., Krebber, W., and Sottek, R. (1998). "The AUDIS catalog of human HRTFs," *J. Acoust. Soc. Am.* **103**, 3082.
- Bronkhorst, A. W. (2000). "The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions," *Acta. Acust. Acust.* **86**, 117–128.
- Byrne, D., Dillon, H., Tran, K., Arlinger, S., Wilbraham, K., Cox, R., Hagerman, B., Hetu, R., Kei, H., Lui, C., Kiessling, J., Nasser Kotby, M., Nasser, N., El Kholy, W. A. E., Nakanishi, Y., Oyer, H., Powell, R., Stephens, D., Meredith, R., Sirimanna, T., Tavartkiladze, G., Frolenkov, G. I., Westerman, S., and Ludvigsen, C. (1994). "An international comparison of long-term average speech spectra," *J. Acoust. Soc. Am.* **96**, 2108–2120.
- Byrne, D., and Noble, W. (1998). "Optimizing sound localization with hearing aids," *Trends Amplif.* **3**, 51–73.
- Dillon, H. (1996). "Compression? Yes, but for low or high frequencies, for low or high intensities, and with what response times?" *Ear Hear.* **17**, 287–307.
- Edmonds, B. A., and Culling, J. F. (2005). "The role of head-related time and level cues in the unmasking of speech in noise and competing speech," *Acta. Acust. Acust.* **91**, 546–553.
- Festen, J. M., and Plomp, R. (1986). "Speech-reception threshold in noise with one and two hearing aids," *J. Acoust. Soc. Am.* **79**, 465–471.
- Freyman, R. L., Helfer, K. S., McCall, D. D., and Clifton, R. K. (1999). "The role of perceived spatial separation in the unmasking of speech," *J. Acoust. Soc. Am.* **106**, 3578–3588.
- Gatehouse, S., Naylor, G., and Elberling, C. (2006). "Linear and nonlinear hearing aid fittings—2. Patterns of candidature," *Int. J. Audiol.* **45**, 153–171.
- Greenberg, S. (2006). "A multi-tier framework for understanding spoken language," in *Listening to Speech: An Auditory Perspective*, edited by S. Greenberg and W. A. Ainsworth (Lawrence Erlbaum Associates, Mahwah, NJ), Chap. 25, pp. 411–434.
- Hansen, L. B. (2008). "Epoq study measures user benefits," *Hear. J.* **61**, 47–49.
- Kalluri, S., and Edwards, B. (2007). "Impact of hearing impairment and hearing aids on benefits due to binaural hearing," in *Proceedings of the 19th International Congresses on Acoustics*, September 2007, Madrid, Spain, pp. 4100–4105.
- Kam, A. C., and Wong, L. L. (1999). "Comparison of performance with wide dynamic range compression and linear amplification," *J. Am. Acad. Audiol.* **10**, 445–457.
- Kates, J. M. (2008). *Digital Hearing Aids* (Plural Publishing, San Diego, CA), Chap. 13, pp. 406–416.
- Kates, J. M. (2010). "Understanding compression: Modeling the effects of dynamic-range compression in hearing aids," *Int. J. Audiol.* **49**, 395–409.
- Kates, J. M., and Arehart, K. H. (2005). "Coherence and the speech intelligibility index," *J. Acoust. Soc. Am.* **117**, 2224–2237.
- Keidser, G., Rohrseitz, K., Dillon, H., Hamacher, V., Carter, L., Rass, U., and Convery, E. (2006). "The effect of multi-channel wide dynamic range compression, noise reduction, and the directional microphone on horizontal localization performance in hearing aid wearers," *Int. J. Audiol.* **45**, 563–579.
- Kidd, G., Jr., Mason, C. R., Richards, V. M., Gallun, F. J., and Durlach, N. I. (2008). "Informational masking," in *Auditory Perception of Sound Sources*, edited by W. A. Yost, A. N. Popper, and R. R. Fay (Springer, New York), Chap. 6, pp. 143–190.
- Kreisman, B. M., Mazeveski, A. G., Schum, D. J., and Sockalingam, R. (2010). "Improvements in speech understanding with wireless binaural broadband digital hearing instruments in adults with sensorineural hearing loss," *Trends Amplif.* **14**, 3–11.

- Macpherson, E. A., and Middlebrooks, J. C. (2002). "Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited," *J. Acoust. Soc. Am.* **111**, 2219–2236.
- Marrone, N., Mason, C. R., and Kidd, G., Jr. (2008). "Evaluating the benefit of hearing aids in solving the cocktail party problem," *Trends Amplif.* **12**, 300–315.
- Middlebrooks, J. C., and Green, D. M. (1991). "Sound localization by human listeners," *Annu. Rev. Psychol.* **42**, 135–159.
- Moore, B. C. J. (2008). "The choice of compression speed in hearing aids: Theoretical and practical considerations and the role of individual differences," *Trends Amplif.* **12**, 103–112.
- Moore, B. C. J., Glasberg, B. R., and Stone, M. A. (1999a). "Use of a loudness model for hearing aid fitting: III. A general method for deriving initial fittings for hearing aids with multi-channel compression," *Br. J. Audiol.* **33**, 241–258.
- Moore, B. C. J., Glasberg, B. R., and Stone, M. A. (2010). "Development of a new method for deriving initial fittings for hearing aids with multi-channel compression: CAMEQ2-HF," *Int. J. Audiol.* **49**, 216–227.
- Moore, B. C. J., Johnson, J. S., Clark, T. M., and Pluinage, V. (1992). "Evaluation of a dual-channel full dynamic range compression system for people with sensorineural hearing loss," *Ear Hear.* **13**, 349–370.
- Moore, B. C. J., Peters, R. W., and Stone, M. A. (1999b). "Benefits of linear amplification and multichannel compression for speech comprehension in backgrounds with spectral and temporal dips," *J. Acoust. Soc. Am.* **105**, 400–411.
- Musa-Shufani, S., Walger, M., von Wedel, H., and Meister, H. (2006). "Influence of dynamic compression on directional hearing in the horizontal plane," *Ear Hear.* **27**, 279–285.
- Naylor, G., and Johannesson, R. B. (2009). "Long-term signal-to-noise ratio at the input and output of amplitude-compression systems," *J. Am. Acad. Audiol.* **20**, 161–171.
- Neher, T., Behrens, T., Carlile, S., Jin, C., Kragelund, L., Petersen, A., and Schaik, A. (2009). "Benefit from spatial separation of multiple talkers in bilateral hearing-aid users: Effects of hearing loss, age, and cognition," *Int. J. Audiol.* **48**, 758–774.
- Olsen, H. L., Olofsson, Å., and Hagerman, B. (2004). "The effect of presentation level and compression characteristics on sentence recognition in modulated noise," *Int. J. Audiol.* **43**, 283–294.
- Olsen, H. L., Olofsson, Å., and Hagerman, B. (2005). "The effect of audibility, signal-to-noise ratio, and temporal speech cues on the benefit from fast-acting compression in modulated noise," *Int. J. Audiol.* **44**, 421–433.
- Plomp, R. (1988). "The negative effect of amplitude compression in multi-channel hearing aids in light of the modulation-transfer function," *J. Acoust. Soc. Am.* **83**, 2322–2327.
- Rhebergen, K. S., Versfeld, N. J., and Dreschler, W. A. (2009). "The dynamic range of speech, compression, and its effect on the speech reception threshold in stationary and interrupted noise," *J. Acoust. Soc. Am.* **126**, 3236–3245.
- Robinson, C. E., and Huntington, D. A. (1973). "The intelligibility of speech processed by delayed long-term averaged compression amplification," *J. Acoust. Soc. Am.* **54**, 314.
- Sockalingam, R., Holmberg, M., Eneroth, K., and Shulte, M. (2009). "Binaural hearing aid communication shown to improve sound quality and localization," *Hear. J.* **62**, 46–47.
- Souza, P. E. (2002). "Effects of compression on speech acoustics, intelligibility, and sound quality," *Trends Amplif.* **6**, 131–165.
- Souza, P. E., Jenstad, L. M., and Boike, K. T. (2006). "Measuring the acoustic effects of compression amplification on speech in noise," *J. Acoust. Soc. Am.* **119**, 41–44.
- Souza, P. E., and Turner, C. W. (1999). "Quantifying the contribution of audibility to recognition of compression-amplified speech," *Ear Hear.* **20**, 12–20.
- Stacey, P. C., and Summerfield, A. Q. (2007). "Effectiveness of computer-based auditory training in improving the perception of noise-vocoded speech," *J. Acoust. Soc. Am.* **121**, 2923–2935.
- Stone, M. A., and Moore, B. C. J. (2007). "Quantifying the effects of fast-acting compression on the envelope of speech," *J. Acoust. Soc. Am.* **121**, 1654–1664.
- Wiggins, I. M., and Seeber, B. U. (2011). "Dynamic-range compression affects the lateral position of sounds," *J. Acoust. Soc. Am.* **130**, 3939–3953.
- Wiggins, I. M., and Seeber, B. U. (2012). "Effects of dynamic-range compression on the spatial attributes of sounds in normal-hearing listeners," *Ear Hear.* **33**, 399–410.
- Yoo, S., Boston, J. R., Durrant, J. D., Kovacyk, K., Karn, S., Shaiman, S., El-Jaroudi, A., and Li, C.-C. (2005). "Relative energy and intelligibility of transient speech information," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, March 2005, Philadelphia, PA, pp. 69–72.