Effects of envelope discontinuities on perceptual restoration of amplitude-compressed speech

Deniz Başkent,[a)] Cheryl Eiler, and Brent Edwards
Starkey Hearing Research Center, 2150 Shattuck Avenue, Suite 408, Berkeley, California 94704

(Received 19 January 2008; revised 23 March 2009; accepted 5 April 2009)

An interrupted signal may be perceptually restored and, as a result, perceived as continuous, when the interruptions are filled with loud noise bursts. Additionally, when the signal is speech, an improvement in intelligibility may be observed. The perceived continuity of interrupted tones is reduced when the signal level is ramped down and up before and after the noise burst, respectively—an effect that has been attributed to envelope discontinuities at the tone-noise interface [Bregman, A. S., and Dannenbring, G. L. (1977). Can. J. Psychiatry **31**, 151–159]. The hypothesis of the present study was that the perceptual restoration of speech would also be reduced with similar envelope discontinuities that may occur in real life due to the release time constants of hearing-aid compression. In an effort to make the conditions more relevant to hearing aids, speech was amplitude-compressed and normal-hearing listeners of varying ages were recruited. Envelope amplitude ramps were placed at the onsets/offsets of speech segments of interrupted sentences and the restoration effect was measured in two ways: objectively as the improvement in intelligibility when noise was added in the gaps and subjectively through the perceived continuity measured by subjects' own reporting. Both measures showed a reduction as the ramp duration increased—a trend observed for subjects of all ages and for all ramp configurations. These findings can be attributed to envelope discontinuities, with an additional contribution from reduced speech information due to ramping and temporal masking from loud noise bursts.
 [DOI: 10.1121/1.3125329]

PACS number(s): 43.71.An, 43.66.Mk, 43.66.Ts [RSN] Pages: 3995–4005

I. INTRODUCTION

In everyday listening, sounds of interest are commonly masked by more intense sounds in the background. One way the auditory system deals with this difficulty is through the perceptual restoration of the incomplete signal using context information, as well as linguistic knowledge and syntactic and semantic constraints if the signal is speech (Miller and Licklider, 1950; Warren, 1970; Dannenbring, 1976). One consequence of this restoration is the perception of an interrupted signal as continuous once the interruption is filled with a louder sound (continuity illusion; Thurlow, 1957; Thurlow and Elfner, 1959; Warren *et al.*, 1972). For speech, in addition to perceived continuity, an improvement in intelligibility may also be observed through this restoration (phonemic restoration; Cherry and Wiley, 1967; Wiley, 1968; Warren, 1970; Warren and Obusek, 1971; Powers and Wilcox, 1977; Verschuure and Brocaar, 1983; Bashford and Warren, 1987; Bashford *et al.*, 1988, 1996).

One requirement for the intervening loud sound to induce perceptual restoration of the interrupted signal is that there be no perceptual evidence of a change in the signal. It helps when the inducer sound has the appropriate spectral, temporal, or spatial acoustic characteristics, as well as sufficient intensity, to mask the missing signal (had the signal been present during the interruption). It also helps when the signal onsets and offsets around the interruptions are not perceptible (Huggins, 1964; Warren *et al.*, 1972; Bregman and Dannenbring, 1977; Bashford and Warren, 1979; Verschuure and Brocaar, 1983; Bashford and Warren, 1987; Bregman, 1990; Bashford *et al.*, 1992). Under these conditions, the combination of interrupted signal with the inducer sound produces an ambiguous input to the auditory system, which cannot readily infer whether portions of the signal are masked or missing behind the louder sound. One possible explanation is, following the Gestalt principles, the system tends to decide that the signal should be continuous and the audible speech segments should be part of one speech stream (Bregman, 1990; Woods *et al.*, 1996; Assmann and Summerfield, 2004; Husain *et al.*, 2005; Srinivasan and Wang, 2005). This situation seems to facilitate a filling in of the missing speech with the help of top-down processes by using redundancies and context in speech as well as linguistic knowledge and constraints (Warren, 1970; Bashford and Warren, 1979; Samuel, 1981, 1996; Bashford *et al.*, 1996; Assmann and Summerfield, 2004; Sivonen *et al.*, 2006).

The requirement that there be no evidence for a change in the signal, however, may be violated in real-life listening situations. As a result, perceptual restoration may be negatively affected and the potential benefit of improved speech understanding reduced. Such violations may occur, for example, with front-end processing of hearing aids (Edwards, 2004). Let us hypothetically consider an example stimulus of an interrupted tone combined with a loud noise burst, as shown in the top left corner of Fig. 1. A typical feature used in hearing aids is amplitude compression where low-level portions of the input signal are amplified more than high-

[a)]Author to whom correspondence should be addressed. Electronic mail: deniz_baskent@starkey.com

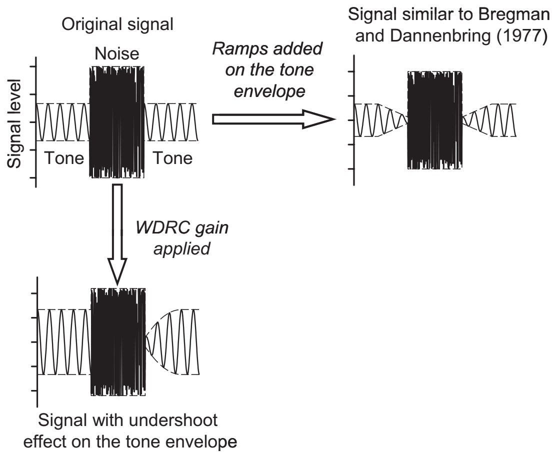

FIG. 1. A signal comprised of a low-intensity interrupted tone combined with a high-intensity noise burst, shown in the upper left corner. The dashed lines indicate the envelope of the combined signal. The upper right corner shows the same combination signal after amplitude ramps were added on the tone envelope preceding and following the noise burst, similar to stimuli used by Bregman and Dannenbring (1977). The lower left corner shows the combination signal after a WDRC gain was applied. Due to the release time constant of compression, a recovery ramp—namely, the undershoot distortion—may be observed on the tone envelope following the noise burst.

level portions. As a result, in a typical compression scenario, tones with a low intensity are amplified with substantial gain. When the signal switches from a low-intensity tone to high-intensity noise, compression is activated and the gain setting changes from substantial to minimal. The reduction in gain due to the activation of compression is the "attack" portion of the compression, which usually happens very fast. By contrast, when the signal switches from a high to low-intensity noise, the gain setting returns from minimal to substantial. This increase in gain is the "release" portion of compression and usually happens slowly, with a pace governed by the release time constant of compression. This slow adjustment in gain may generate an undershoot distortion, that is, an increasing ramp on the tone envelope as shown in the lower left corner of Fig. 1 (Edwards, 2004). Significantly, the distortion between the tone-noise boundary can be harmful for perceptual restoration. Bregman and Dannenbring (1977) produced similar envelope manipulations by adding amplitude ramps on the signal envelope before and after the noise inducer, as shown in the top right corner of Fig. 1. As the duration of the ramps increased, it became difficult for the listeners to perceive the interrupted tone as continuous.

In this study, we hypothesized that the perceptual restoration of speech could also be reduced with similar envelope manipulations—a finding that would have important practical consequences for fitting hearing aids. Based on the previous findings with interrupted tones and as a direct consequence of the violation of the evidence of the continuity rule, we expected three interacting factors. Firstly, envelope discontinuities generated by the ramps were expected to make perceptual restoration of interrupted speech more difficult. Unlike tonal signals, however, speech contains linguistic information that is crucial to its restoration. As a consequence, a second factor potentially affecting the results was a loss in speech information due to the addition of the ramps. A third factor possibly affecting restoration was the potential masking from loud noise bursts onto the speech segments. While the noise level must be high to induce the restoration, this high level then increases the possibility of temporal masking.

Two methods were used to test our hypothesis: an objective measurement of improvement in speech intelligibility due to the addition of noise (Powers and Wilcox, 1977) and a subjective measure in which listeners reported the perceived continuity of interrupted sentences combined with noise. Three different ramp configurations were implemented: one simulating the undershoot distortion, another similar to that used by Bregman and Dannenbring (1977), and one complementing the preceding two and helping to tease out the effects of different factors.

II. METHODS

The first type of ramp configuration simulated the undershoot distortion that can occur with hearing-aid compression (Fig. 1, lower left panel). Ideally, the effect of this configuration should be measured with hearing-impaired listeners who are the real users of these devices. However, the effect of hearing impairment on perceptual restoration is still under investigation and as yet, unknown (Başkent *et al.*, 2007). We therefore decided to focus on the effect of simulated undershoot without the interference from hearing impairment—recruiting normal-hearing listeners for the present study. To make both conditions more realistic and the results more relevant to real hearing-aid users, we (1) made an effort to select listeners of varying ages to more closely parallel the elderly user population, and (2) amplitude-compressed speech before applying the experimental conditions. As a result, in addition to the simulated undershoot in the envelope, other potential perceptual changes due to compression were included in the processed signal.

A. Listeners

A total of 26 normal-hearing listeners participated in the study—all native speakers of American English whose ages varied from 18 to 79 years (with an average age of 37). All subjects' hearing thresholds were better than 20 dB Hearing Level (HL) at audiometric frequencies of 250–4000 Hz. Twenty- two listeners participated in both objective and subjective tests, three only in the objective test, and one in the subjective test alone. Subject participation was determined by their availability for testing.

B. Stimuli

Sentences from the Hearing in Noise Test (HINT) database (Nilsson *et al.*, 1994) and Harvard database (IEEE, 1969) were used for training and data collection sessions, respectively. Although the IEEE sentences have fewer contextual cues (Rabinowitz *et al.*, 1992), this database has the advantage of having a large number of stimuli, which enabled multiple measurements without re-using the sentences with each subject. All sentences were spoken by a single male speaker. The filler noise was a speech-shaped steady noise produced from the long-term speech spectrum of the IEEE sentences.

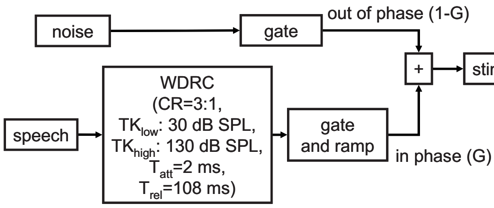

FIG. 2. Summary of the signal processing.

Figure 2 shows a summary of the signal processing; the following sections provide further details.

1. Amplitude compression of speech

The first step of the signal processing involved the compression of sentences with wide-dynamic range compression (WDRC), with a compression ratio of 3:1. The low and high knee points (the levels between which the compression was applied) were set to 30 and 130 dB Sound Pressure Levels (SPLs), respectively. This range was wider than what is typically used in hearing aids, but was chosen so that all speech amplitude components were inside the compression region. The *RC* time constants for the attack and release times were 1 and 30 ms, respectively, measured to be equivalent to the ANSI time constants of 1.7 and 108.4 ms, respectively (ANSI, 2003).

2. Gating

In the next step, the amplitude-compressed sentences and the noise were interrupted periodically by using a gating function of 50% duty cycle and a period of 450 ms, which corresponded to an interruption rate of 2.2 Hz. There were several reasons for selecting this rate. (1) Perceptual restoration of speech is robust with this rate (Warren *et al.*, 1972; Houtgast, 1974; Powers and Wilcox, 1977). (2) Restoration was best when the interruption duration was less than the average word length (Bashford *et al.*, 1988). 2.2 Hz produces interruptions of 225 ms in duration, smaller than the average word durations measured at 378 and 383 ms for HINT and IEEE sentences, respectively. (3) Performance during the pilot study was in the mid-range of the psychometric function for most listeners, minimizing ceiling and floor effects. (4) Speech segments (225 ms in duration) were long enough to implement a range of ramp durations, including the values used by Bregman and Dannenbring (1977).

The gating function started with the *on phase* for speech and *off phase* for noise (Fig. 2) while the rise/fall time of 5 ms was implemented with a cosine ramp. Speech and noise segments overlapped at the transition, as shown in Fig. 3(a), to prevent a reduction in level that could be detected by listeners. Figure 3(b) shows a sample stimulus after interrupted speech was combined with interrupted noise, but before the amplitude ramps were added. Stimuli similar to these (with no ramps) were used to measure baseline performance.

3. Amplitude ramps

In the final step, the amplitude ramps were implemented with cosine ramps that were applied to the gating function

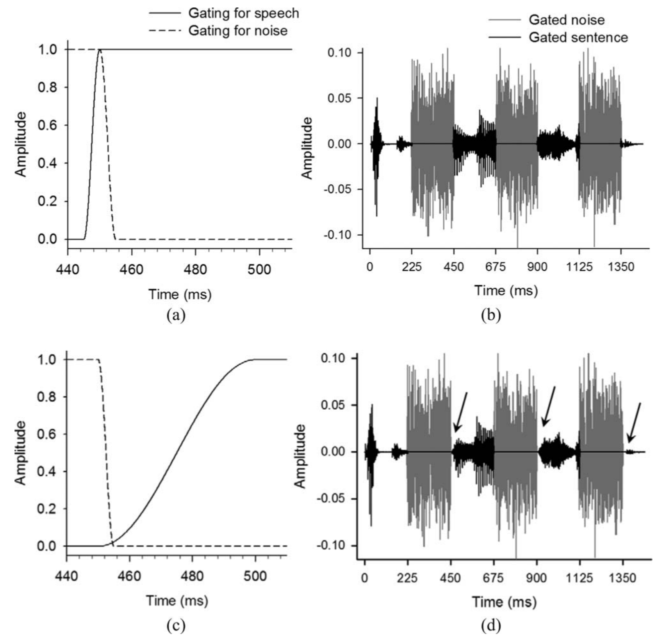

FIG. 3. (a) The envelopes of speech and noise shown at the transition point before an amplitude ramp was added. Note that there is no level discontinuity due to the overlap between the signals. (b) An example stimulus before the amplitude ramps were added. Speech and noise are shown in black and gray, respectively. (c) The same transition from (a), shown after an amplitude ramp of 50 ms added on the speech envelope at the onset of the speech segment. (d) The effect of the amplitude ramp is shown on the stimulus from (b), as indicated by the arrows.

for speech. An example of an envelope ramp at the transition from noise to speech is shown in Fig. 3(c), while the effect of adding this ramp to the overall stimulus is shown in Fig. 3(d).

Ramp durations of 10, 50, and 100 ms simulated time constants comparable to those of syllabic compression systems (Van Tasell, 1993; Souza, 2002) and covered the range of values used by Bregman and Dannenbring (1977).

The ramps were implemented in three different configurations.

(1) *Onset*. Ramps were placed at the onset of the speech segments following the noise bursts—simulating the undershoot (as shown in the top two rows of Fig. 4).
(2) *Both*. Ramps were placed both at the onset and offset of the speech segments—similar to those used by Bregman and Dannenbring (1977) (as shown in the middle rows of Fig. 4).
(3) *Offset*. Ramps were at the offset of the speech segments before the noise bursts (as shown in the bottom rows of Fig. 4).

4. Presentation of the stimuli

The stimuli were presented binaurally using the TDT System III with Sennheiser HD 580 headphones in a soundproof booth. The system was calibrated using a B&K $\frac{1}{2}$ in. microphone mounted in an artificial ear coupler to measure the frequency response of white noise at the output of the headphones. The absolute level of the noise was determined using a reference tone of 96 dB SPL. This noise level was later used to determine the maximum dB SPL for the stimuli

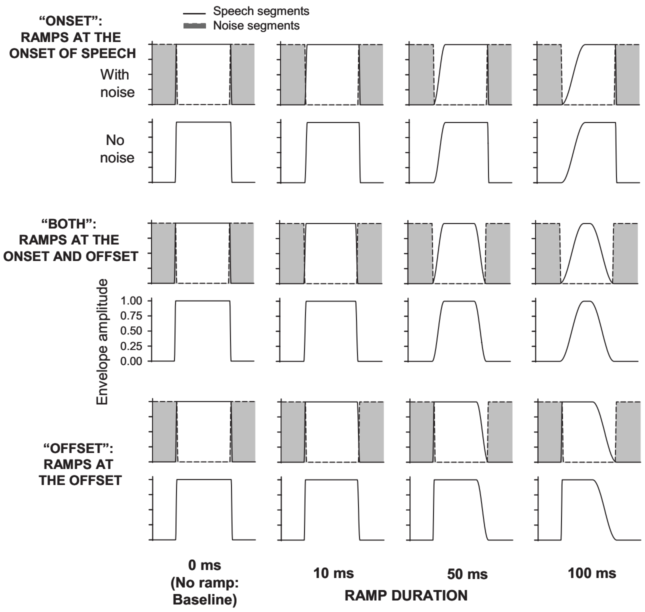

FIG. 4. A summary of the amplitude ramp conditions. The top rows show the envelopes with amplitude ramps added at the onset of speech segments, while the middle rows show the envelopes where the ramps were added both at the onset and offset of the speech segments. The bottom rows show the envelopes where the ramps were added at the offset only. In each row, the left panel shows the baseline condition with no ramp (0 ms) with the ramp duration increasing as one moves toward the right.

at the headphone output. A TDT attenuator was used to adjust the presentation levels with reference to these calculated maximum levels.

Speech was presented at 65 dB SPL. A 1-kHz tone 0.5 s in duration and presented at 60 dB SPL was used to cue the listeners to the beginning of each stimulus. The level of the noise varied during the training whereas a fixed level of 75 dB SPL (again determined by the pilot study) was used during the data collection.

The rms levels of speech stimuli at all stages of signal processing were equalized to 65 dB SPL, the presentation level of the original unprocessed sentence. The motivation for this equalization was to maintain similar energy and loudness levels for speech across all conditions. However, due to this equalization, speech peak levels were most likely higher during the longer-duration ramps.

C. Experimental procedure

The procedure consisted of three sequential stages: Training (to familiarize subjects with the procedure), an objective and subjective test. For listeners who participated in only one of the tests, testing followed the training. The entire procedure was completed in one to three sessions, with a total duration of 2–6 h.

1. Training

Training was similar to the objective test, except that in the training, (1) feedback was provided, (2) sentences were simpler and had more contextual cues, and (3) conditions within the training changed from easy to difficult by a gradual shortening of the speech duty cycle and a speeding up of the interruption rate (both different from those used during actual data collection). A list of ten HINT sentences was used for each training condition. Identical training was given to each subject by keeping the order of both the conditions and the sentences the same. Table I summarizes the conditions used for training and percent correct scores measured within each condition.

2. Objective measure of perceptual restoration

In the objective test, recognition of interruption sentences was measured with and without noise; the increase in intelligibility by the addition of noise was the metric for the perceptual restoration benefit. Subjects were instructed to listen to the processed sentences and verbally repeat as many words as possible. When uncertain, they were encouraged to guess in order to increase the influence of the top-down mechanism. The experimenter judged the accuracy of the repeated words (excluding the articles "the," "a," and "an") and recorded the correct words using the MATLAB GUI. The

TABLE I. Summary of the training conditions and the results obtained with these conditions. The scores denoted by "*" indicated a significant ($p < 0.05$) improvement in intelligibility.

Compression condition for speech segments	Duration of speech segments (ms)	Duration of noise segments (ms)	Speech level (dB SPL)	Noise level (dB SPL)	Raw percent correct scores (average score ± one standard deviation) (%)	Percent correct scores for benefit from perceptual restoration—measured as the difference between no-noise and noise-added conditions (average score ± one standard deviation) (%)
Uncompressed	500	200	65	No noise	94.1 ± 4.9	1.0 ± 5.8
				65	95.1 ± 3.7	
Compressed	300	300	65	No noise	74.6 ± 11.6	$8.4^{*} \pm 10.0$
				70	83.0 ± 7.9	
Compressed	400	200	65	No noise	90.5 ± 5.3	$3.3^{*} \pm 4.6$
				75	93.9 ± 3.3	
Compressed	200	200	65	No noise	62.2 ± 17.4	$28.6^{*} \pm 19.1$
				75	90.8 ± 4.7	

MATLAB program automatically calculated the percent correct scores from the number of the correct words in relation to the total number of words in the sentence, and stored the results in log files for each subject. The experimenter did not know which conditions were being tested, except for an occasional leakage of sound at very high noise levels through the headphones. Each stimulus was presented only once with no repetition and no feedback was provided.

Measurements were repeated twice. Twenty conditions were tested in each round: 2 baseline conditions with no ramp (0 ms) $\times$ 2 noise levels (no noise and 75 dB SPL) and 18 ramp conditions with 3 ramp configurations (onset, both, offset) $\times$3 ramp durations (10, 50, 100 ms) $\times$2 noise levels (no noise and 75 dB SPL). The order of the conditions was randomized for both each round and each listener. A list of ten IEEE sentences was used in each condition, making a total of 400 sentences (20 conditions $\times$ 2 repetitions) used for the entire objective test with each subject.

3. Subjective measure of perceived continuity

In the subjective test, the perceived continuity was measured with interrupted sentences that were combined with the noise bursts. The listeners were instructed to decide if they heard the sentence as interrupted or continuous, and then themselves entered their response using the MATLAB GUI. The MATLAB program automatically calculated the percentage of the sentences heard as continuous and stored the scores in log files.

To help listeners understand the instructions, the test started with ten sentences, half of which were more likely to be perceived as continuous since the noise level was high (75 dB SPL) and half of which were more likely to be perceived as interrupted since the noise level was low (65 dB SPL). After the ten initial sentences, the orders of both the conditions and the sentence lists were randomized.

There were two rounds of measurements. Ten conditions were tested in each round: one baseline condition with no ramp (0 ms) and nine ramp conditions with 3 ramp configurations (onset, both, offset) $\times$3 ramp durations (10, 50, 100 ms). A list of ten IEEE sentences was used in each condition, making a total of 200 sentences (10 conditions $\times$ 2 repetitions) used for the entire subjective test with each subject.

III. RESULTS

Figure 5 shows the average objective and subjective scores as a function of ramp duration. The leftmost score in each panel shows the baseline performance with no ramps (indicated by the ramp duration of 0 ms). The results in the top row are for the "onset" ramp configuration that simulated undershoot (also shown in the top rows of Fig. 4). The results in the middle row are for the "both" configuration, similar to the setting used by Bregman and Dannenbring (1977) (also shown in the middle rows of Fig. 4). The results in the bottom row are for the third "offset" ramp configuration (also shown in the bottom rows of Fig. 4). The two left columns show the objective measures of speech intelligibility. In the leftmost column, the open circles and squares

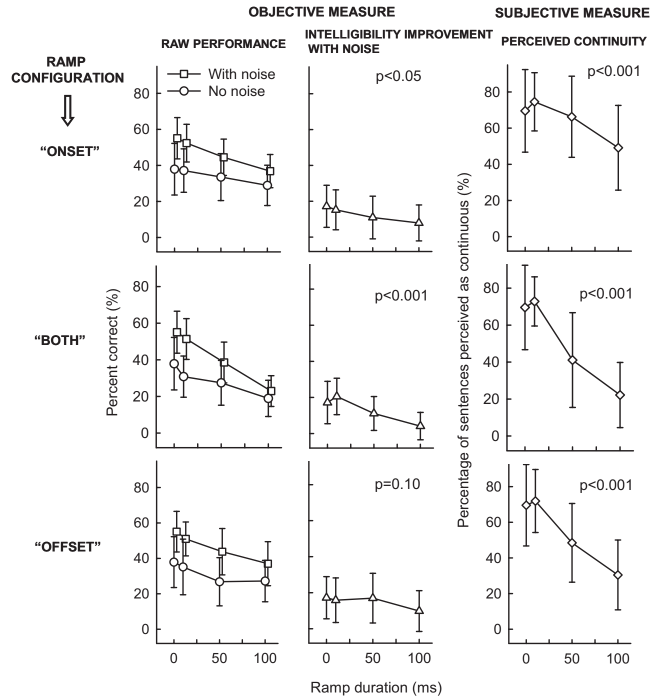

FIG. 5. Average scores from the objective and subjective tests shown as a function of the ramp duration. The panels from top to bottom show the results with the onset, both, and offset ramp configurations, respectively (also shown in Fig. 4). The two left columns show the objective scores. In the leftmost column, the open circles show the performance with interrupted sentences and the open squares show the performance after the noise was added to the interrupted sentences. The difference between these two scores, plotted with open triangles in the middle column, shows the intelligibility benefit. The percentage scores from the subjective measure of perceived continuity are shown in the right column with open diamonds. The significance level of the effect of the ramp duration on the objective and subjective measures is indicated by the p numbers in each panel.

show the raw percent correct scores with the interrupted sentences without and with the noise bursts, respectively. The difference between the two raw scores was attributed to the benefit from perceptual restoration, explicitly shown in the middle column. The right column shows the percentage scores from the subjective measure of perceived continuity.

A. Baseline performance with no ramps

In the baseline condition there were no amplitude ramps on the speech envelope and no discontinuity in the level of the combined signals [Figs. 3(a) and 3(b), and the leftmost column of Fig. 4]. The leftmost scores in each panel of Fig. 5 show the baseline performance. The average restoration benefit was 17.19%, with a standard deviation of 11.71%. Three listeners had negative restoration scores. Despite the difference of using compressed speech in the present study, both mean and variance values were similar to those observed by Powers and Wilcox (1977). The percentage of sentences perceived as continuous was on average 68.10%, with a standard deviation of 22.94%.

B. Effects of the amplitude ramps

Figure 5 shows that in general all scores dropped from the baseline levels as the duration of the ramps increased. First, let us examine the raw percent correct scores in the left

TABLE II. Scores for intelligibility benefit and perceived continuity compared for baseline and 100-ms duration ramp conditions. Superscript letters indicate a significant reduction in the scores due to the ramp, as determined by a posthoc Tukey test.

Ramp configuration	Baseline objective score of restoration benefit (percent correct)	Objective benefit score with the 100-ms ramps (percent correct)	Baseline subjective score of perceived continuity (percentage)	Subjective score with the 100-ms ramp (percentage)
Onset		7.92[a]		49.13[b]
Both	17.19	4.14[b]	69.57	22.17[b]
Offset		9.81		30.43[b]

[a]$p < 0.05$.
[b]$p < 0.001$.

panels of Fig. 5. Scores both with and without noise dropped as a function of the ramp duration. However, the performance with noise was better than without for almost all ramp durations. A two-factor repeated measures (RM) Analysis of Variance (ANOVA) with the factors of ramp duration and noise was conducted with the scores; both factors had significant main effects on performance for all three ramp configurations [ramp duration, $F(3,72) \geq 37.17$, $p < 0.001$ and noise, $F(1,24) \geq 103.80$, $p < 0.001$]. Now let us examine the difference in the raw scores between no-noise and noise conditions, an effect that was attributed to the perceptual restoration (shown more explicitly in the middle columns of Fig. 5). The drop in raw scores with the noise (open squares) was more pronounced than the drop in raw scores with no noise (open circles) as the ramp duration increased; hence the restoration benefit decreased. This reduction was significant for onset and both ramp configurations (middle panels of the first and second rows, respectively, in Fig. 5). The significance level was determined by the interaction of the factors of ramp duration and noise [$F(3,72) \geq 3.79$, $p < 0.05$ and $p < 0.001$, for onset and both configurations, respectively]. Next, let us examine the subjective scores presented in the right column of Fig. 5. These data show that perceived continuity diminished as the ramp duration increased. The main effect of the ramp duration was significant for all three ramp configurations [one-factor RM ANOVA; $F(3,66) \geq 14.83$, $p < 0.001$].

For both objective and subjective measures, the strongest effects among ramp configurations were observed with the both configuration and among ramp durations with the longest ramp duration of 100 ms. As a sample comparison, the objective and subjective scores at 100 ms ramp duration are presented in Table II.

C. Analysis of data for age factor

To explore the effects of age on the results, we reanalyzed the data from Fig. 5 re-plotting the individual baseline scores with no ramps as a function of listener age (see Fig. 6). Regression lines (shown by dashed or solid lines) were superimposed on the data to indicate the trend. The correlation coefficients (r_s) and corresponding p values were calculated with the Spearman rank order correlation since we did not know if the correlations were linear. These data show that recognition of interrupted speech (with or without the noise) decreased as the listener's age increased (left panel in Fig. 6). However, perceptual restoration ability seemed to be independent of age; the changes in the intelligibility benefit (see middle panel in Fig. 6) and the perceived continuity (right panel in Fig. 6) as a function of age were minimal and non-significant.

Figure 7 shows another analysis for age, this time for the amplitude ramp conditions. We divided the listeners into two subgroups (using the age of 40 as the threshold) and averaged the scores separately for each. This division was based on the study by Bergman *et al.* (1976) who reported a clear difference in the intelligibility of interrupted speech between listeners younger and older than 40. With this separation, the objective test had 15 younger and 10 older listeners; and the subjective test had 14 younger and 9 older listeners. Similar to Fig. 5, Fig. 7 shows the results with the ramp configurations of onset, both, and offset, in the panels from top to bottom. The panels from left to right show the raw scores with interrupted speech with or without noise, the objective scores of restoration benefit, and subjective scores of perceived continuity, respectively. The only difference from Fig. 5 is that open symbols with solid lines show the results averaged across younger listeners while the gray symbols with dashed lines show the results averaged across listeners older than 40. Similar to the observations from Fig. 6 with the interrupted speech, the raw scores by the older group were

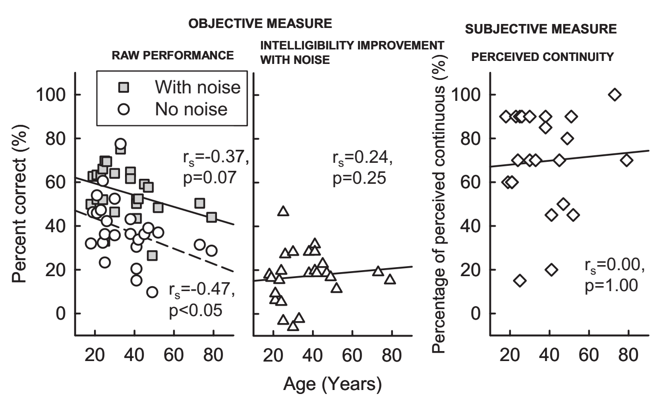

FIG. 6. Individual baseline scores (no ramps) re-plotted from Fig. 5 as a function of age. The dashed and the solid lines superimposed on individual data show the linear regression lines. The correlation coefficients (r_s) and p values, calculated using the Spearman rank order correlation, are indicated in each panel.

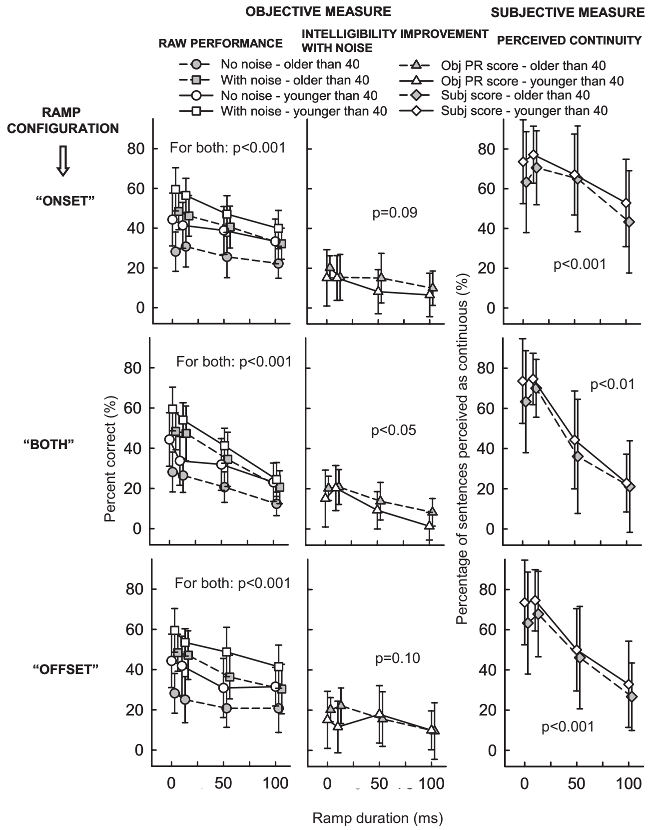

FIG. 7. Data from Fig. 5, averaged for listeners younger and older than 40 years. The p values in each panel indicate a statistically significant difference between the scores of the younger and older listener groups.

substantially lower than the younger group, with or without noise for all ramp configurations (left column of Fig. 7). A two-factor mixed-design ANOVA with the factors of ramp duration and subject group showed that this difference was significant [$F(1,92) \geq 11.82$, $p < 0.001$]. Despite this difference in the raw scores, once the objective scores for the restoration benefit were calculated (as shown in the middle column of Fig. 7) there was no significant difference in performance by the younger and older subject groups for the onset and offset configurations. For the both configuration, the performance by the older group was slightly but significantly better than that by the younger group [$F(1,92)=4.62$, $p < 0.05$], although a posthoc Tukey test did not identify a significant difference for any specific ramp duration. With the subjective measures (shown in the right column of Fig. 7) scores of the older listeners were significantly lower than those of the younger listeners for all three configurations [$F(1,84) \geq 7.24$, $p < 0.01$]. Overall, the scores of the older listeners were lower than those of the younger listeners for the perception of interrupted speech with or without noise and perceived continuity. On the other hand, the restoration benefit was comparable or slightly better for the older listeners. Despite these differences, there was no significant interaction between the main factors of ramp duration and subject group, implying that a similar trend was observed in the data as a function of the ramp duration between the two subject groups.

IV. CONCLUSIONS AND DISCUSSION

A. Effects of amplitude ramps on the perceptual restoration of speech

Baseline scores of the objective and subjective measures of perceptual restoration showed that adding noise to the interruptions increased intelligibility and perceived continuity of interrupted compressed speech, similar to results from earlier studies with interrupted uncompressed speech (Cherry and Wiley, 1967; Wiley, 1968; Warren, 1970; Powers and Wilcox, 1977; Verschuure and Brocaar, 1983; Bashford and Warren, 1987; Bashford *et al.*, 1996). When the amplitude ramps were added on the onsets/offsets of speech segments, both objective and subjective scores decreased significantly from the baseline level as the ramp duration increased.

There are three factors that have potentially contributed to the results observed with the ramps.

(1) *Reduced speech information*. The right panels of Fig. 3 show that one of the direct consequences of adding ramps on the speech envelope was a reduction in speech information—possibly also reducing the linguistic content and context of speech. Both objective and subjective measures of perceptual restoration could have been affected by the loss of speech information, but possibly to different degrees. The increase in intelligibility with intervening noise depends on the filling in by the top-down mechanisms that use speech information and context (Verschuure and Brocaar, 1983; Bashford and Warren, 1987; Bashford *et al.*, 1992, 1996), but these may not be as important for perceived continuity. Continuity illusion works with all signals regardless of linguistic factors, as long as the requirements for perceptual restoration listed in the Introduction are satisfied (Warren *et al.*, 1972; Bregman and Dannenbring, 1977; Bashford and Warren, 1979; Verschuure and Brocaar, 1983; Warren, 1984; Bashford and Warren, 1987; Bregman, 1990; Bashford *et al.*, 1992). Earlier studies indicated that linguistic factors may affect objective and subjective measures of perceptual restoration differently. Intervening noise causes interrupted monosyllables and words (speech with no context) to be perceived as continuous, but usually with no improvement in intelligibility (Miller and Licklider, 1950; Hopkinson, 1967; Kreul, 1971; Samuel, 1981). Improvement is observed with speech that has rich linguistic content and context, such as with sentences or running speech (Schubert and Parker, 1955; Powers and Wilcox, 1977; Verschuure and Brocaar, 1983). Bashford and Warren (1979) showed that it was not necessary to understand speech to perceive it as continuous. The illusion worked even with speech played backwards, although the perceived continuity was more robust with speech played normally rather than backwards. From these observations, loss of speech information would be expected to affect the objective measure more than the subjective. In our study, however, both measures decreased significantly with increasing ramp duration, indicating that there may be factors other than pure information loss that further affected the results. This idea is supported by the results presented in Fig. 5: The average

scores with the interrupted speech with and without the noise (left panels of Fig. 5, open squares and open circles, respectively) showed that both performances dropped as the ramp duration increased; the reduction in scores with the noise, however, was faster. Since theoretically, the information loss is the same between the no-noise and with-noise conditions, the effect of speech information loss should have been similar between the two curves. As it was not, there must have been an additional factor that further reduced the scores with the noise conditions.

(2) *Level discontinuities due to envelope changes*. Our main hypothesis was that the envelope discontinuities due to the ramps would affect the perceptual restoration of speech negatively. Bregman and Dannenbring (1977) previously showed that envelope manipulations similar to ours reduced the perceived continuity of an interrupted tone combined with a noise burst. In their experiment, because of the nature of the stimulus, a tone, there was no linguistic content and, consequently, no effect due to loss of information. Therefore, the reduced continuity illusion must have been mainly due to the level changes in the tone envelope at the temporal edges with the noise. As the basic principles for perceptual restoration of verbal and nonverbal stimuli seem to be similar (Warren *et al.*, 1972; Bregman and Dannenbring, 1977; Bashford and Warren, 1979; Verschuure and Brocaar, 1983; Warren, 1984; Bashford and Warren, 1987; Bregman, 1990; Bashford *et al.*, 1992), the detrimental effects of the ramps observed on perceived continuity of interrupted tones should also apply to perceived continuity of interrupted speech. This reasoning would explain the ramp effects observed on the subjective scores of perceptual restoration. It is not clear, however, how this factor could affect the restoration benefit observed with the objective measure. For example, it is possible that perceived continuity is required in order to benefit from restoration. As explained in the Introduction, the common understanding about improvement in speech intelligibility as a result of noise is that with its addition, the auditory system cannot tell with certainty whether the speech is being interrupted or is continuous and simply masked by the noise. This ambiguity seems to be useful: The system assumes that the speech is continuous but masked—activating top-down mechanisms to fill in missing speech and improve intelligibility (Repp, 1992; Woods *et al.*, 1996; Srinivasan and Wang, 2005). Bashford *et al.* (1992) and Bregman (1990) suggested a two-stage model of perceptual restoration of speech with intervening noise: The primitive first stage works for all signals (verbal or nonverbal) and simply decides whether the signal continues behind the masker from the available cues (the "whether" question). The knowledge-driven second stage then finds a plausible answer to what the missing parts might be (the "what" question). In a later study, Shinn-Cunningham and Wang (2008) reported anecdotal observations, which seemingly contradict the two-stage model by observing an improvement in intelligibility under conditions where speech would not have been perceived as continuous. Yet, they also added that were the perceived continuity induced, improvement might have been more robust. While identification of the specific mechanisms of perceptual restoration is beyond the scope of the present study, generally, perceived continuity and improved intelligibility appear to be closely related for the perceptual restoration of speech (Verschuure and Brocaar, 1983). Therefore, a reduction in perceived continuity could result in the reduction in the objective benefit from restoration.

(3) *Masking*. Loss of speech information and envelope discontinuities seem to be sufficient to explain the general pattern of the results of the present study. However, a third factor should also be considered: masking due to the intense loud noise bursts. To mask interruptions and speech onsets/offsets (in order to induce robust perceptual restoration), the intervening noise has to be a broadband signal of relatively high intensity. Using a broadband and/or intense noise for this purpose, however, has the adverse effect of increased temporal masking (forward or backward) of speech segments (Dirks and Bower, 1970; Bashford *et al.*, 1992, 1996). The various effects of temporal masking are noteworthy for this study.

(a) *For the baseline condition with no ramps*. Masking from noise bursts can theoretically reduce the audibility of speech segments and decrease the linguistic content, which in turn can affect perceptual restoration negatively. The leftmost scores in the left panels of Fig. 5 show the scores with the baseline conditions. Due to the potential masking, there could have been less speech information available to listeners with the noise conditions than the no-noise conditions. Dirks and Bower (1970), however, showed that temporal masking of speech from loud noise bursts was ignorable at slow interruption rates of noise. While at high rates (such as 100 Hz) they observed substantial masking, at the slow rate of 1 Hz (close to the 2.2 Hz rate of the present study) the masking effect was minimal. These results raise the possibility that the effect from masking may not have been substantial in the baseline conditions.

(b) *For conditions with the ramps*. In this instance there could be an opposite (and positive) effect—the partial inaudibility of the ramps compensating for their detrimental effects, especially for shorter ramp durations. An example of this positive effect was observed in the present study with the 10-ms ramp. Scores with this ramp were almost identical to those of the no-ramp baseline, implying that this ramp was not noticed, possibly in response to the temporal masking from noise. The negative effects of the amplitude ramps occurred with the longer-duration ramps of 50 and 100 ms, which in not being (entirely) masked, produced a perceptual effect.

As the effects of forward masking are generally more pronounced than those of backward masking (Gaskell and Henning, 1999) the present study used different ramp configurations to tease out possible differences. The results, however, showed no clear difference in data with onset and offset configurations—the configurations that would have reflected effects of forward and backward masking separately—raising the possibility that the difference between the two masking mechanisms may have been too small to be reflected in the scores.

Overall, the data indicate that the short duration ramp of 10 ms was not perceptible and did not have any noticeable effect on perceptual restoration. The longer-duration ramps of 50 and 100 ms were not masked, at least not entirely, as they produced a reduction in the scores.

One interesting observation about the overall findings is that both objective and subjective measures decreased monotonically as the ramp duration increased, instead of either measure completely falling apart. Even at the longest ramp duration of 100 ms, there were still positive scores for both measures, albeit much smaller as compared to the baseline. The conventional understanding of continuity illusion sees it as an all-or-none phenomenon. For example, Houtgast (1972) used the threshold between perceived continuity and discontinuity of a signal as a measure for masking experiments. In a later study, however, Warren *et al.* (1994) showed that perceived continuity with nonverbal signals was not all-or-none, but instead worked on a continuum. The results of the present study extend this finding to speech signals. Thus both the intelligibility benefit and perceived continuity of speech do not seem to work as all-or-none mechanisms, but instead, according to a graded effect.

B. Implications for hearing-aid processing: Effects of simulated undershoot

One of the main motivations for the study was to explore the possible effects of undershoot distortion simulated by the onset ramp configuration [shown in Fig. 3(c) and in the lower left corner of Fig. 1]. With this configuration, there was a significant reduction in both the intelligibility benefit and perceived continuity as the ramp duration, i.e., the simulated release time constant, increased. The results imply that in listening environments with fluctuating background noise, compression release times may affect perceptual restoration of speech and hence its intelligibility negatively, depending on the settings of the hearing aid. Note that the configuration used in the present study was an attempt to simulate an extreme case of what may happen with hearing-aid processing in order to observe the effects fully; the results may differ in real-life applications depending on the compression settings used. For example, in the present experiments, the envelope amplitude of speech was reduced to zero immediately after the noise burst at the onset of the amplitude ramp, simulating the greatest change that can theoretically occur in the speech envelope. For a change in that order to happen, the gain applied during the noise and during the speech should differ by 65 dB SPL or more. It is likely that the change in the speech envelope will be less with a more realistic compression system under real-life conditions, and as a result, the disruptive effects of these smaller ramps may be less pronounced than reported here. Additionally, multiband compression where the input stimulus is compressed only at the frequency regions where the listener has reduced dynamic range is commonly used in real-life settings (e.g., Moore *et al.*, 1985; Kuk, 1999; Hansen, 2002). In the present study, the ramps were applied to the overall envelope of broadband speech, as if a single-band compression system was used. If the ramps were applied only to high-frequency bands of speech to simulate multiband compression, the results could again be less dramatic.

C. Applicability of the results to hearing-aid users: Potential effects of age and hearing impairment

As the aforementioned effects of simulated undershoot have been observed with normal-hearing listeners, it is unclear how these results would apply to actual users of hearing aids. Two main factors, advanced age and hearing impairment, could change the results significantly for hearing-aid users. To explore the effects of age, the present study analyzed the data in two different ways: first for the baseline conditions (shown in Fig. 6) and then for the ramp conditions (shown in Fig. 7). The results indicated that the intelligibility of interrupted speech, with or without noise, was negatively correlated, i.e., decreased as the listener's age increased—an expected effect that has previously been observed and attributed to the reduced temporal processing by elderly listeners (Bergman *et al.*, 1976; Gordon-Salant and Fitzgibbons, 1993; Stuart and Phillips, 1996). What was not expected, however, was the lack of change in the restoration benefit scores as a function of age (as seen in the middle panels of Figs. 6 and 7). Previous studies have shown that elderly listeners have difficulty accessing speech information in the dips of a fluctuating background noise (Stuart and Phillips, 1996), again presumably due to reduced temporal processing. We therefore had assumed that the limited access to the speech information in the dips would also make perceptual restoration more difficult for older listeners. Despite the difficulties understanding speech with interruptions, with or without noise, older listeners seemed to make use of the added noise to improve intelligibility as well as younger listeners. A similar observation was made by Madix *et al.* (2005) who, using a different paradigm to measure perceptual restoration, found no difference in performance between younger (19–28 years old) and older (41–62 years old) listeners. A possible explanation for the lack of a decrement in the restoration benefit by elderly listeners might be a compensation by better use of linguistic information (Barrett and Wright, 1979; Wingfield and Tun, 2001). Older listeners may benefit more from context information (Pichora-Fuller *et al.*, 1995) and/or use this information more efficiently (Wingfield *et al.*, 1991) than younger listeners.

The right panel of Fig. 6 showed no effect of age on the baseline subjective scores. However, once analyzed for all conditions, including both the baseline and the ramps (as shown in the right panel of Fig. 7), the scores of the older subjects were significantly lower than those of the younger group. Regardless of this shift in scores, however, the trend in subjective data was almost identical between the subject groups, implying that the negative effects of the ramps were similar for both groups of listeners.

Overall, this analysis established that (1) elderly listeners also benefit from perceptual restoration, and (2) this benefit, as well as perceived continuity, is reduced by level discontinuities in a pattern similar to that present with younger listeners. As a result, the findings of the present study would not be expected to differ with hearing-aid users due specifically to their increased age.

The remaining question involved the impact of hearing impairment on the current findings. Earlier studies have shown that, even after audibility was ensured with proper amplification, hearing-impaired listeners had more difficulty in extracting speech from fluctuating background noise (Festen and Plomp, 1990; Eisenberg *et al.*, 1995). One possible explanation is increased forward masking. Nelson and Pavlov (1989) demonstrated longer time constants for forward masking with moderately impaired listeners and Dubno *et al.* (2003) later found a negative correlation between forward masking thresholds and speech intelligibility in fluctuating background noise. This factor, combined with the loss in speech redundancy due to other potential suprathreshold deficits, might make it more difficult for hearing-impaired listeners to have access to speech segments between loud noise bursts—therefore reducing the benefits of perceptual restoration. On the other hand, stronger masking from noise segments might be an advantage for hearing-aid users since the undershoot ramps might be less detectible and therefore less disruptive. We are aware of only one study that has explored perceptual restoration with hearing-impaired listeners (Başkent *et al.*, 2007). The preliminary results showed that while some hearing-impaired listeners benefited from perceptual restoration, this benefit disappeared as the severity of the hearing loss increased. With large variability in the results and compromised audibility as possible confounding factors, however, more data are needed before drawing conclusive results.

V. SUMMARY

The mean and the variance of the baseline restoration benefit and perceived continuity with compressed interrupted speech were similar to values published for uncompressed interrupted speech in previous studies.

All three ramp configurations reduced both objective and subjective scores as the ramp duration increased. One ramp configuration was similar to the configuration used by Bregman and Dannenbring (1977) who showed that perceived continuity of interrupted tones reduced with similar amplitude ramps. They attributed this effect to the discontinuities in the tone envelope due to the ramps. We hypothesized that similar level discontinuities would reduce perceptual restoration of speech as well. Our data supported this hypothesis, with the potential contribution of two additional factors, speech information loss and temporal masking from loud noise bursts, to the results.

The reduction in the scores as a function of ramp duration was graded and the perceptual restoration effect did not entirely disappear even at the longest ramp durations. This observation supports the idea that perceptual restoration benefit and perceived continuity of speech are not an all-or-none mechanism.

One ramp configuration simulated the undershoot distortion. The reduction in perceptual restoration with this configuration indicated the potentially negative effect of hearing-aid processing. The applicability of these results with normal-hearing listeners to actual users of hearing aids would depend on two main factors, hearing impairment and the advanced age of hearing-aid users. An analysis for age showed that the trends for the perceptual restoration data were similar between younger and older listeners, ruling out age as a potential factor that would affect the applicability of the results. The effect of hearing impairment remains unknown at this point. If future research shows that hearing impairment does not prevent benefiting from perceptual restoration, the findings of the present study could have important implications for hearing-aid users.

ACKNOWLEDGMENTS

The authors would like to thank John Galvin and Qian-Jie Fu from House Ear Institute for providing the recordings for the IEEE sentences, Bill Woods for constructive suggestions, the subjects for their efforts, and Anastasios Sarampalis, Nancy Robinson-Kime, and the anonymous reviewers for their comments on a previous version of the manuscript.

ANSI (**2003**). Specification of hearing aid characteristics, S3.22 (American National Standards Inst., New York).

Assmann, P. F., and Summerfield, Q. (**2004**). "The perception of speech under adverse acoustic conditions," *Speech Processing in the Auditory System*, Springer Handbook of Auditory Research Vol. **18**, edited by S. Greenberg, W. A. Ainsworth, A. N. Popper, and R. R. Fay (Springer-Verlag, New York).

Barrett, T. R., and Wright, M. (**1979**). "Age-related facilitation in recall following semantic processing," J. Gerontol. **36**, 194–199.

Bashford, J. A., and Warren, R. M. (**1979**). "Perceptual synthesis of deleted phonemes," J. Acoust. Soc. Am. **65**, 423–426.

Bashford, J. A., and Warren, R. M. (**1987**). "Multiple phonemic restorations follow the rules for auditory induction," Percept. Psychophys. **42**, 114–121.

Bashford, J. A., Myers, M. D., Brubaker, B. S., and Warren, R. M. (**1988**). "Illusory continuity of interrupted speech: Speech rate determines durational limits," J. Acoust. Soc. Am. **84**, 1635–1638.

Bashford, J. A., Riener, K. R., and Warren, R. M. (**1992**). "Increasing the intelligibility of speech through multiple phonemic restorations," Percept. Psychophys. **51**, 211–217.

Bashford, J. A., Warren, R. M., and Brown, C. A. (**1996**). "Use of speech-modulated noise adds strong 'bottom-up' cues for phonemic restoration," Percept. Psychophys. **58**, 342–350.

Başkent, D., Eiler, C., and Edwards, B. (**2007**). "Effects of amplitude ramps on phonemic restoration of compressed speech with normal-hearing and hearing-impaired listeners," Proceedings of the First International Symposium on Auditory and Audiological Research (ISAAR), Helsingør, Denmark, pp. 585–593.

Bergman, M., Blumenfeld, V. G., Cascardo, D., Dash, B., Levitt, H., and Margulies, M. K. (**1976**). "Age-related decremenet in hearing for speech. Sampling and longitudinal studies," J. Gerontol. **31**, 533–538.

Bregman, A. S. (**1990**). *Auditory Scene Analysis: The Perceptual Organization of Sound* (MIT, Cambridge).

Bregman, A. S., and Dannenbring, G. L. (**1977**). "Auditory continuity and amplitude edges," Can. J. Psychiatry **31**, 151–159.

Cherry, C., and Wiley, R. (**1967**). "Speech communications in very noisy environments," Nature (London) **214**, 1164.

Dannenbring, G. L. (**1976**). "Perceived auditory continuity with alternately rising and falling frequency transitions," Can. J. Psychiatry **30**(2), 99–114.

Dirks, D. D., and Bower, D. (**1970**). "Effect of forward masking and backward masking on speech intelligibility," J. Acoust. Soc. Am. **47**, 1003–1008.

Dubno, J. R., Horwitz, A. R., and Ahlstrom, J. B. (**2003**). "Recovery from prior stimulation: Masking of speech by interrupted noise for younger and older adults with normal hearing," J. Acoust. Soc. Am. **113**, 2084–2094.

Edwards, B. (**2004**). "Hearing aids and hearing impairment," *Speech Pro-*

cessing in the Auditory System, Springer Handbook of Auditory Research Vol. **18**, edited by S. Greenberg, W. A. Ainsworth, A. N. Popper, and R. R. Fay (Springer-Verlag, New York).

Eisenberg, L. S., Dirks, D. D., and Bell, T. S. (**1995**). "Speech recognition in amplitude-modulated noise of listeners with normal and listeners with impaired hearing," J. Speech Hear. Res. **38**, 222–233.

Festen, J. M., and Plomp, R. (**1990**). "Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing," J. Acoust. Soc. Am. **88**, 1725–1736.

Gaskell, H., and Henning, G. B. (**1999**). "Forward and backward masking with brief impulsive stimuli," Hear. Res. **129**, 92–100.

Gordon-Salant, S., and Fitzgibbons, P. J. (**1993**). "Temporal factors and speech recognition performance in young and elderly listeners," J. Speech Hear. Res. **36**, 1276–1285.

Hansen, M. (**2002**). "Effects of multi-channel compression time constants on subjectively perceived sound quality and speech intelligibility," Ear Hear. **23**, 369–380.

Hopkinson, N. T. (**1967**). "Combined effects of interruption and interaural alternation on speech intelligibility," Lang Speech **10**, 234–243.

Houtgast, T. (**1972**). "Psychophysical evidence for lateral inhibition in hearing," J. Acoust. Soc. Am. **51**, 1885–1894.

Houtgast, T. (**1974**). "Lateral suppression in hearing: A psychophysical study on the ear's capability to preserve and enhance speech contrasts," Ph.D. thesis, Free University of Amsterdam, Amsterdam, The Netherlands.

Huggins, A. W. F. (**1964**). "Distortion of the temporal pattern of speech: Interruption and alternation," J. Acoust. Soc. Am. **36**, 1055–1064.

Husain, F. T., Lozito, T. P., Ulloa, A., and Horwitz, B. (**2005**). "Investigating the neural basis of the auditory continuity illusion," J. Cogn Neurosci. **17**, 1275–1292.

IEEE (**1969**). "IEEE recommended practice for speech quality measurements," IEEE Trans. Audio Electroacoust. **AU-17**, 225–246.

Kreul, E. J. (**1971**). "Speech intelligibility for interaural alternated speech with and without intervening noise for words and nonsense," Lang Speech **14**, 99–107.

Kuk, F. K. (**1999**). "Theoretical and practical considerations in compression hearing aids," Trends Amplif. **1**, 5–39.

Madix, S. G., Thelin, J. W., Plyler, P. N., Hedrick, M., and Malone, J. (**2005**). "The effects of age and context on phonemic restoration in young adult females," American Academy of Audiology Annual Conference, Washington, D.C., poster presentation.

Miller, G. A., and Licklider, J. C. R. (**1950**). "The intelligibility of interrupted speech," J. Acoust. Soc. Am. **22**, 167–173.

Moore, B., Laurence, R., and Wright, D. (**1985**). "Improvements in speech intelligibility in quiet and in noise produced by two-channel compression hearing aids," Br. J. Audiol. **19**, 175–187.

Nelson, D. A., and Pavlov, R. (**1989**). "Auditory time constants for off-frequency forward masking in normal-hearing and hearing-impaired listeners," J. Speech Hear. Res. **32**, 296–306.

Nilsson, M., Soli, S., and Sullivan, J. (**1994**). "Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and noise," J. Acoust. Soc. Am. **95**, 1085–1099.

Pichora-Fuller, K. M., Schneider, B. A., and Daneman, M. (**1995**). "How young and old adults listen to and remember speech in noise," J. Acoust. Soc. Am. **97**, 593–608.

Powers, G. L., and Wilcox, J. C. (**1977**). "Intelligibility of temporally interrupted speech with and without intervening noise," J. Acoust. Soc. Am. **61**, 195–199.

Rabinowitz, W. M., Eddington, D. K., Delhorne, L. A., and Cuneo, P. A. (**1992**). "Relations among different measures of speech reception in subjects using a cochlear implant," J. Acoust. Soc. Am. **92**, 1869–1881.

Repp, B. H. (**1992**). "Perceptual restoration of a "missing" speech sound: Auditory induction or illusion?," Percept. Psychophys. **51**, 14–32.

Samuel, A. G. (**1981**). "Phonemic restoration: Insights from a new methodology," J. Exp. Psychol. Gen. **110**, 474–494.

Samuel, A. G. (**1996**). "Does lexical information influence the perceptual restoration of phonemes?," J. Exp. Psychol. Gen. **125**, 28–51.

Schubert, E. D., and Parker, C. D. (**1955**). "Addition to Cherry's finding on switching speech between the two ears," J. Acoust. Soc. Am. **27**, 792–794.

Shinn-Cunningham, B. G., and Wang, D. (**2008**). "Influences of auditory object formation on phonemic restoration," J. Acoust. Soc. Am. **123**, 295–301.

Sivonen, P., Maess, B., Lattner, S., and Friederici, A. D. (**2006**). "Phonemic restoration in a sentence context: Evidence from early and late ERP effects," Brain Res. **1121**, 177–189.

Srinivasan, S., and Wang, D. (**2005**). "A schema-based model for phonemic restoration," Speech Commun. **45**, 63–87.

Stuart, A., and Phillips, D. P. (**1996**). "Word recognition in continuous and interrupted broadband noise by young normal-hearing, older normal-hearing, and presbyacusic listeners," Ear Hear. **17**, 478–489.

Souza, P. (**2002**). "Effects of compression on speech acoustics, intelligibility, and sound quality," Trends Amplif. **6**, 131–165.

Thurlow, W. R. (**1957**). "An auditory figure-ground effect," Am. J. Psychol. **70**, 653–654.

Thurlow, W. R., and Elfner, L. F. (**1959**). "Continuity effects with alternately sounding tones," J. Acoust. Soc. Am. **31**, 1337–1339.

Van Tasell, D. J. (**1993**). "Hearing loss, speech and hearing aids," J. Speech Hear. Res. **36**, 228–244.

Verschuure, J., and Brocaar, M. P. (**1983**). "Intelligibility of interrupted speech with and without intervening noise," Percept. Psychophys. **33**, 232–240.

Warren, R. M. (**1970**). "Perceptual restoration of missing speech sounds," Science **167**, 392–393.

Warren, R. M. (**1984**). "Perceptual restoration of obliterated sounds," Psychol. Bull. **96**, 371–383.

Warren, R. M., and Obusek, C. J. (**1971**). "Speech perception and phonemic restorations," Percept. Psychophys. **9**, 358–362.

Warren, R. M., Obusek, C. J., and Ackroff, J. M. (**1972**). "Auditory induction: Perceptual synthesis of absent sounds," Science **176**, 1149–1151.

Warren, R. M., Bashford, J. A., Jr., Healy, E. W., and Brubaker, B. S. (**1994**). "Auditory induction: Reciprocal changes in alternating sounds," Percept. Psychophys. **55**, 313–322.

Wiley, R. (**1968**). "Speech communication using the strongly voiced components only," Ph.D. thesis, Imperial College, University of London, London, UK.

Wingfield, A., and Tun, P. A. (**2001**). "Spoken language comprehension in older adults: Interactions between sensory and cognitive change in normal hearing," Semin. Hear. **22**, 287–302.

Wingfield, A., Aberdeen, J. S., and Stine, E. A. (**1991**). "Word onset gating and linguistic context in spoken word recognition by young and elderly adults," J. Gerontol. **46**, 127–129.

Woods, W. S., Hansen, M., Wittkop, T., and Kollmeier, B. (**1996**). "A scene analyzer for speech processing," *Proceedings of the ESCA Workshop on the Auditory Basis of Speech Perception, ABSP-1996* (International Speech Communication Association (ISCA), Keele, U.K.), pp. 232–235.