Effects of amplitude ramps on phonemic restoration of compressed speech with normal-hearing and hearing-impaired listeners

Deniz Başkent, Cheryl Eiler, and Brent Edwards

Starkey Hearing Research Center, Berkeley, California, USA

Continuity illusion refers to the phenomenon where an interrupted signal is perceived as continuous, once the silent interval is filled with a louder sound. A similar mechanism is believed to help with phonemic restoration of missing speech segments. When speech segments are not audible due to masking of interfering background sounds, listeners may fill in the gaps and have enhanced speech intelligibility, even when those segments are omitted. For example, as a special case of phonemic restoration, it was shown that recognition of gated sentences was better when the silent intervals were filled with loud noise bursts. The present study is a preliminary attempt in exploring hearing aid processing effects on complex listening tasks that are likely to occur in real life, such as restoration of obliterated speech. Specifically, we explored if phonemic restoration might be degraded due to hearing aid compression, which might produce ramps on the speech envelope during the recovery from compression following the loud noise bursts filling the gaps. In Experiment 1, phonemic restoration was measured with normal-hearing (NH) listeners where ramps of varying durations were added on the speech envelope after the noise intervals, to simulate recovery from compression. Phonemic restoration was significantly reduced as the duration of the ramps increased. Experiment 2 shows preliminary results with hearing-impaired (HI) listeners, where phonemic restoration was measured for a number of configurations. The results showed a large variability in phonemic restoration by HI listeners, and audibility and gating period were observed to be important factors affecting the results. The combined results imply that hearing aid compression might have detrimental effects on phonemic restoration; however, more data is needed to determine how applicable these results would be to HI listeners.

INTRODUCTION

When the silent intervals of an interrupted signal are filled with a louder signal, the interrupted signal can be perceived as continuous (continuity illusion). A similar phenomenon is also observed with speech (Warren, 1970). Powers and Wilcox (1977) showed the effect of phonemic restoration using gated sentences. Speech recognition was measured with NH listeners in two conditions: with gated sentences only and with gated sentences where the silent intervals were filled with loud noise bursts. There was an improvement of 10 to 15% in speech recognition scores as a result of adding noise bursts, even though no additional speech information was provided. This observation implies that phonemic restoration can enhance speech perception in noisy lis-

Auditory signal processing in hearing-impaired listeners. 1st International Symposium on Auditory and Audiological Research (ISAAR 2007). T. Dau, J. M. Buchholz, J. M. Harte, T. U. Christiansen (Eds.).
ISBN: 87-990013-1-4. Print: Centertryk A/S.

tening environments where speech segments are not audible due to masking from loud background sounds.

Bregman and Dannenbring (1977) observed that continuity illusion could be weaker if the tone intensity was altered by using falling or rising ramps around the noise burst filling the interruption. The stimulus used in the study is shown in Fig. 1 (a).

Let us hypothetically consider that the tone with the noise filling the interruption period is presented through a hearing aid with compressive gain. In a typical scenario, the tone at the moderate level would be moderately amplified. However, the loud noise would probably be compressed, as shown in Fig. 1(b). At the offset of the noise burst the system would recover from the compression, which would produce an increasing ramp on the tone due to the release time constant of the compressor. The reduction in the tone level, the duration of the amplitude ramp, and the levels of the tone and the noise would be determined by the compressor parameters such as the knee points, compression ratio, and release time constant. The similarity of the stimuli shown in Fig. 1 (a) and (b) suggests that hearing aid compression can potentially reduce continuity illusion, as well as phonemic restoration, as the continuity illusion and phonemic restoration seem to operate on similar principles.

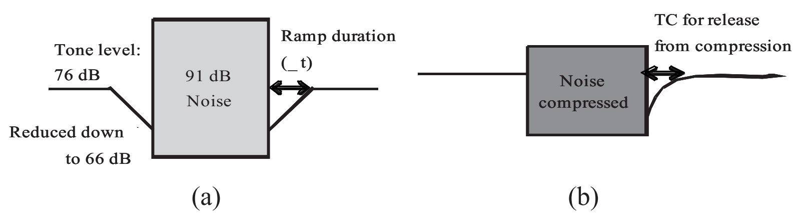

Fig. 1: Comparison of the stimulus used by Bregman and Dannenbring (1977), shown in (a), with the hypothetical representation of the same stimulus after it was compressed, shown in (b).

In the present study, we explored if such amplitude ramps that may occur due to hearing aid compression on the speech envelope would affect phonemic restoration. In Experiment 1, phonemic restoration was measured with NH listeners using a method similar to the one used by Powers and Wilcox (1977). The potential effects of compression recovery were measured by inserting rising ramps on speech segments following the noise bursts. In Experiment 2, the baseline phonemic restoration was explored with HI subjects, as a preliminary step to investigating effects of hearing aid processing.

EXPERIMENT 1: EFFECTS OF AMPLITUDE RAMPS ON PHONEMIC RESTORATION WITH NORMAL-HEARING LISTENERS

Subjects

Twenty four NH listeners, ages varying from 18 to 79 years with an average of 37 years, participated in Experiment 1. All subjects were native speakers of American English and the audiometric thresholds were measured to be 20 dB HL or better for frequencies ranging from 250 to 4000 Hz.

Stimuli

A short training with HINT sentences (Nilsson *et al.*, 1994) was provided prior to data collection. IEEE sentences (IEEE, 1969) were used during data collection. A speech-shaped steady noise (SSN) produced from the long-term speech spectrum (LTSS) of the IEEE sentences was used to fill the silent intervals. The stimuli were presented with the TDT System III over Sennheiser HD 580 headphones binaurally in a sound-proof booth. The speech presentation level was 65 dB SPL. The level of the noise varied in the training, and was fixed at 75 dB SPL during data collection.

In an attempt to make speech stimuli more realistic, sentences were compressed with fast WDRC with a compression ratio of 3:1 prior to gating. The speech and silent gaps with equal durations of 225 msec, similar to durations reported by Powers and Wilcox (1977), were produced with gating. The same gating function with the opposite phase was used to produce the gated SSN that was used to fill in the silent segments. Signal processing is summarized in Fig. 2.

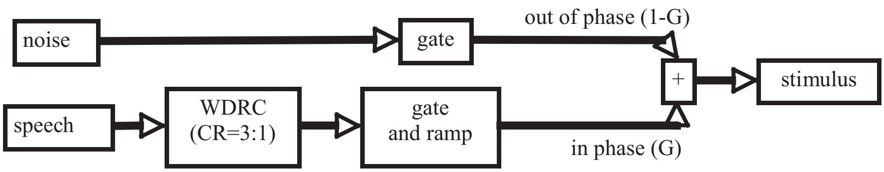

Fig. 2: Summary of processing of the stimuli.

Possible effects of release time constant of compression were explored by producing amplitude ramps on the speech envelope following the noise bursts. The ramps were implemented using a raised cosine function with durations of 10, 50, and 100 msec.

Methods

Ten sentences were used for each condition and each condition was tested twice. After the presentation of each sentence, the subject was asked to repeat what they heard. The percent correct scores for each condition were calculated by counting the number of the words identified correctly. No repetition was allowed and no feedback was provided except in the training. The order of the conditions and the order of the sentences were randomized. Each subject heard each sentence only once.

In a follow-up test, perceived continuity of the sentences was also measured subjectively. In this test, only gated sentences combined with the gated noise, but with varying ramp durations, were used. The subjects responded "continuous" or "broken" after each sentence presentation, and the number of the sentences reported to be heard as continuous was counted to calculate the percentage. Each condition was repeated twice, with ten sentences each. By the time the subjective test was run, the subjects had already heard the sentences from the objective test run previously.

Speech recognition results

Fig. 3 (a) shows the average percent correct scores from the NH listeners with the gated sentences, shown with open circles, and with the gated sentences combined with the gated noise, shown with filled circles, as a function of the ramp duration. The error bars show one standard deviation. As the ramp duration increased both performance lines dropped significantly. This is expected as when the ramps were added on the speech envelope some speech information was lost. Fig. 3 (b) shows the phonemic restoration in percent correct scores, denoted by filled triangles and calculated by taking the difference between the scores with and without the noise for each ramp condition from Fig. 3(a). The ramp duration of 0 msec shows the baseline phonemic restoration of 18%, similar to scores reported by Powers and Wilcox (1977). As the ramp duration was increased, there was a significant reduction in phonemic restoration, as revealed by a one-factor Repeated Measures ANOVA with the main effect of ramp duration (p=0.01).

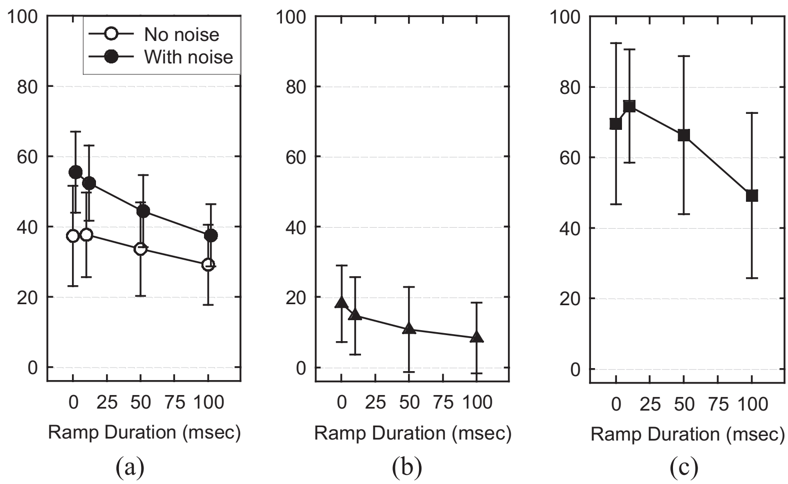

Fig. 3: (a) Percent correct scores with sentences with silent intervals and with noise intervals, averaged across all NH subjects. (b) Phonemic restoration shown in percent correct scores and calculated by taking the difference between the no noise and with noise scores from (a). (c) Percent of sentences perceived as continuous, averaged across all subjects.

Subjective results

Fig. 3 (c) shows the average scores for the percent of the sentences perceived as continuous, denoted by filled squares. The error bars show one standard deviation. Similar to the objective measure of the phonemic restoration, shown in Fig. 3 (b), the perceived continuity decreased significantly as the ramp duration increased (p<0.01).

EXPERIMENT 2: PHONEMIC RESTORATION WITH HEARING-IMPAIRED LISTENERS

Subjects

Six listeners with mild to moderate hearing loss and between the ages of 58 and 86 years, with an average of 69 years, participated in Experiment 2. The audiometric thresholds of individual subjects are shown in Fig. 4 and the age information for the subjects is provided in Table 1. All subjects were native speakers of American English.

Stimuli

A short training with HINT sentences was provided prior to data collection. To make the task easier for this subject group, HINT sentences were also used during the data collection. HINT noise provided with the HINT sentences was used to fill the silent intervals. The stimuli were presented with the TDT System III over Sennheiser HD 580 headphones binaurally in a sound-proof booth. Unlike Experiment 1, sentences were not compressed.

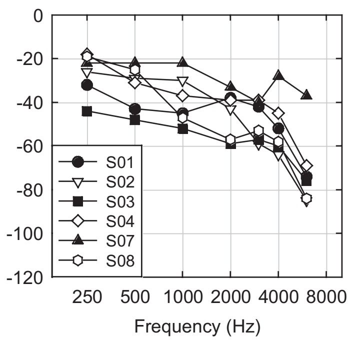

Fig. 4: Audiometric thresholds of hearing-impaired listeners.

Subject	Age (Years)	VC (dB)	Number of the runs for each condition		
			222/2 22	333/3 33	444/2 22
S01	64	-12	on ce	once	---
S02	86	-3	once	once	---
S03	78	0	once	once	---
S04	62	-11	once	twice	twice
S07	58	-2	twice	twice	once
S08	64	-10	twice	twice	once

Table 1: Summary of subject ages, VC settings, and number of the runs for each experimental configuration.

The input speech presentation level was 65 dB SPL, and the input level of the noise was 65, 70, and 75 dB SPL. The stimuli were linearly amplified using the NAL-R prescription. The subjects were also permitted to manually change the volume control (VC) for maximum comfort prior to data collection. The VC settings for individual subjects are shown in Table 1. Three gating configurations were used: 1) Fast interruption rate with 50% duty cycle with the gating on and off periods of 225/225 msec, 2) slow interruption rate with 50% duty cycle with gating on and off periods of 333/333 msec, 3) higher duty cycle with gating on and off periods of 444/222 msec. The gating for the noise was the same except for the last configuration, where the on and off gate durations for the noise were 222/444 msec. The subjects were tested for a different number of runs for different configurations, as summarized in Table 1.

This part of the study is an initial step for exploring if the hearing-impaired listeners would also be able to benefit from phonemic restoration. Once this is established, a follow-up study with the ramps, similar to Experiment 1, will be conducted with hearing-impaired listeners.

Results

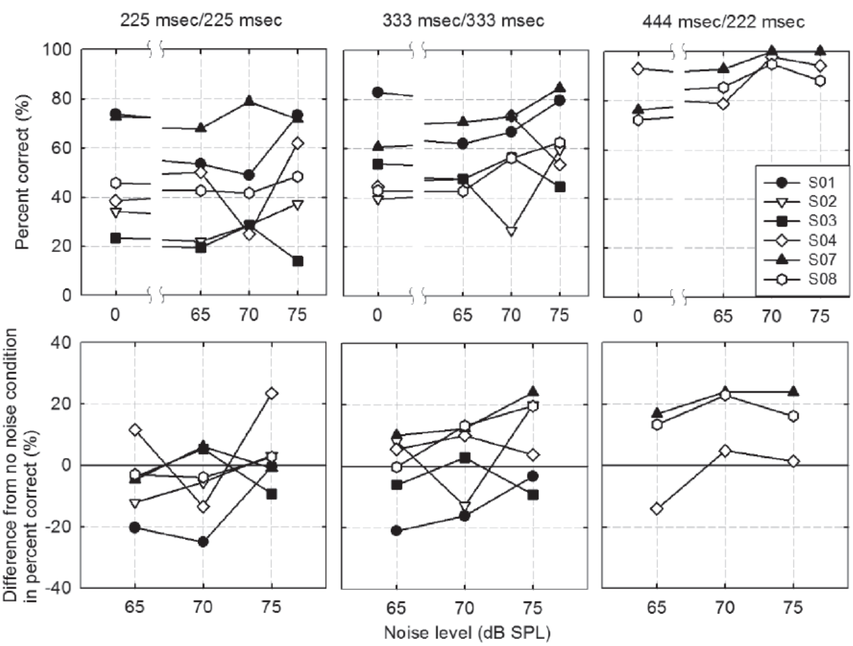

Fig. 5: Percent correct scores, shown for individual HI listeners and as a function of the noise level. The configuration is shown for gating on and off durations for each column above the upper panels. The lower panels show phonemic restoration, calculated by taking the difference between the performance when the noise was added to silent intervals and the performance with the no-noise condition.

Fig. 5 shows the percent correct scores of individual listeners. The columns from left to right show the results for the gating configurations of 225/225, 333/333, and 444/222 msec, respectively. The top panels show the scores for the no noise condition and for the noise levels of 65, 70, and 75 dB SPL, and the lower panels show the phonemic restoration, i.e., how speech recognition changed when noise was added to silent intervals. These scores were calculated by taking the difference in the scores between the noise conditions and the no noise condition.

There was a substantial variability in scores. More phonemic restoration was observed with the slower interruption rate, shown in the middle column of Fig. 5, and with larger duty cycle of speech, shown in the right column of Fig. 5, compared to the fast interruption rate and 50% duty cycle configuration, shown in the left column of Fig. 5. Subject S03 did not show any phonemic restoration effect with either interruption rate, and there was a negative effect of adding noise on speech recognition with subject S01.

Audibility factor

Table 1 shows that subject S01 turned down the volume by 12 dB, an amount that could have affected audibility significantly. A 1/3 octave-band filter analysis with subject's audiometric thresholds and speech levels after the amplification implied that some low and high-frequency components of the stimuli might not have been audible (see Fig. 6). To explore the possibility that audibility might have affected the results, three subjects were retested with the gating on/off condition of 333/333 msec using two additional gain prescriptions: half-gain rule and WDRC. The subjects were also able to set the VC for these new prescriptions. The panels in the left column of Fig. 7 replicated the scores from the left panels of Fig. 5 for these subjects. The middle and right columns of Fig. 7 show the scores with the same subjects with the half-gain rule and using WDRC, respectively.

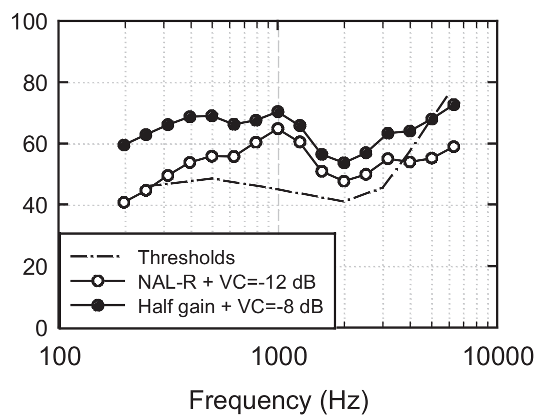

Fig. 6: 1/3 octave-band analysis shown for subject S01. The dashed lines show the audiometric thresholds, the open symbols show the LTSS of the stimuli after NAL-R prescription and VC were applied, and the filled symbols show the LTSS after the half-gain prescription and the VC were applied.

The results with different gain prescriptions showed that audibility might have been a factor that affected the phonemic restoration performance with HI listeners. Performance by the subjects S01 and S03, for example, improved greatly when other prescriptions and a higher VC were used, which both would have increased the audibility for these subjects (Fig. 6).

CONCLUSION

The present study showed that the rising amplitude ramps placed on the speech envelope following the noise bursts may have a negative effect on phonemic restoration. This finding is complementary to the results by Bregman and Dannenbring (1977) who reported that continuity illusion of tones became weaker if similar ramps were placed on the tone envelope. These amplitude ramps can also be interpreted as a crude simulation of the recovery from hearing-aid compression. Therefore, the results may indicate that for a range of release time constants, hearing aid compression might have an adverse effect on phonemic restoration.

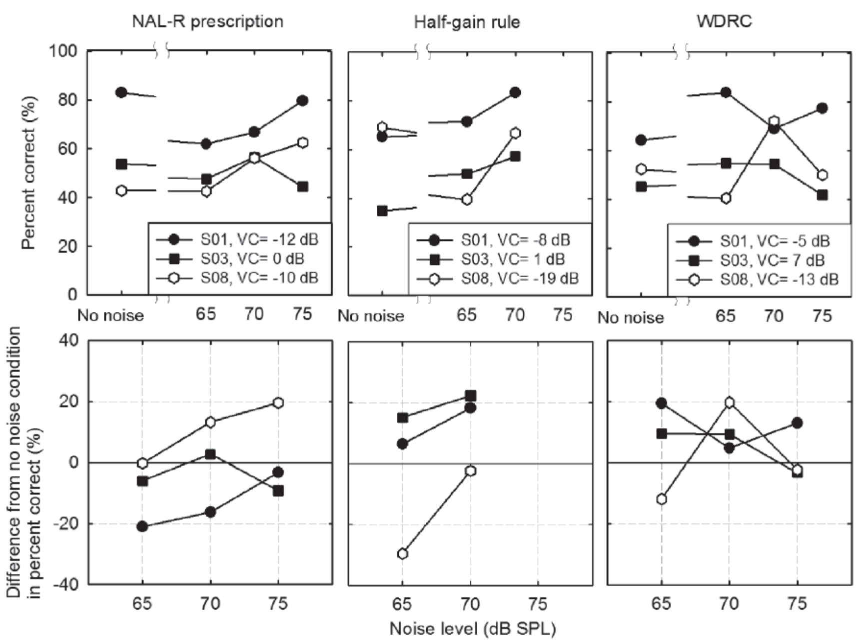

Fig. 7: Similar to Fig. 5, except the scores are shown for different gain rules. The VC settings for each subject and for each prescription are shown in the legends.

The present study also explored phonemic restoration with HI listeners. There was a large variability in the results. Some HI listeners were observed to be able to benefit from phonemic restoration in some experimental configurations, especially with slower gating rates and with gain prescriptions that provided higher presentation levels. Further experiments will be conducted to understand the possible factors such as hearing impairment, advanced age, reduced spectral or temporal sensitivity, or reduced cognitive ability that might have prevented the other listeners from receiving this benefit.

REFERENCES

Bregman, A. S. and Dannenbring, G. L. (**1977**). "Auditory continuity and amplitude edges," Canad. J. Psychol./Rev. Canad. Psychol. **31**, 151-159.

Institute of Electrical and Electronics Engineers (**1969**). IEEE recommended practice for speech quality measurements.

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. Amer. **95**, 1085-1099.

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.

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

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