Increasing the intelligibility of speech through multiple phonemic restorations

JAMES A. BASHFORD, JR., KERI R. RIENER, and RICHARD M. WARREN
University of Wisconsin-Milwaukee, Milwaukee, Wisconsin

Outside of the laboratory, listening conditions are often less than ideal, and when attending to sounds from a particular source, portions are often obliterated by extraneous noises. However, listeners possess rather elegant reconstructive mechanisms. Restoration can be complete, so that missing segments are indistinguishable from those actually present and the listener is unaware that the signal is fragmented. This phenomenon, called temporal induction (TI), has been studied extensively with nonverbal signals and to a lesser extent with speech. Earlier studies have demonstrated that TI can produce illusory continuity spanning gaps of a few hundred milliseconds when portions of a signal are replaced by a louder sound capable of masking the signal were it actually present. The present study employed various types of speech signals with periodic gaps and measured the effects upon intelligibility produced by filling these gaps with noises. Enhancement of intelligibility through multiple phonemic restoration occurred when the acoustic requirements for TI were met and when sufficient contextual information was available in the remaining speech fragments. It appears that phonemic restoration is a specialized form of TI that uses linguistic skills for the reconstruction of obliterated speech.

In everyday life, portions of speech and other acoustic signals are often obliterated by extraneous transient noises such as coughs, slamming doors, and traffic sounds. However, we possess sophisticated mechanisms that can restore the segments of these signals that have been obliterated by noise. The rule governing restoration is simple, and its consequences manifold. Restoration can occur when an extraneous sound is capable of masking a signal and when there is sufficient contextual information to establish the identity of an obliterated fragment. This restoration, which has been called *temporal induction* (TI), is complete, so that when appropriate conditions are met, listeners cannot distinguish restored sounds from those actually present.

A review of the literature has indicated that there are three basic types of TI (Warren, 1984). The simplest form of restoration has been called homophonic continuity. If three levels of the same sound (e.g., 60-, 70-, and 80-dB levels of a noise each lasting 300 msec) are presented in succession and repeated without pauses, the 60-dB level appears to be on continuously with the other levels heard as intermittent pulsed additions. Apparent continuity of the faintest sound occurs despite the fact that this level is actually present for less than half of the time. Homophonic continuity is a subtractive process. Thus, when

an 80-dB sound is alternated with an 82-dB level of the same sound, the higher, 82-dB level appears to be the softer: it is heard as a weaker pulsed addition to the continuous 80-dB level (80 dB subtracted from 82 dB leaves a residue of 77.7 dB).

The form of TI that has been studied most extensively is heterophonic continuity. When two levels of different sounds (for example, sinusoidal tones of different frequencies) alternate in time, the fainter sound is heard as continuous if the louder sound is a potential masker of the fainter sound. This relationship between masking potential and illusory continuity is well established and has served as the basis for the rule proposed by Houtgast (1972), which states that when a tone is alternated regularly with another sound, the tone is heard as continuous when a change from the tone to the louder sound involves no perceptible change in neural activity at any frequency region (the alternating sounds in his experiments were each 125 msec). Houtgast's rule has provided a procedure for investigating the nature of cochlear mechanics and neural response. By determining the upper amplitude limit for the illusory continuity of the fainter sound (called the *pulsation threshold*), it has been possible to obtain estimates of the width of critical bands, the extent of lateral suppression, and the shape of tuning curves (for example, see Fastl, 1975; Houtgast, 1972, 1973, 1974a, 1974b; Kronberg, Mellert, & Schreiner, 1974; Verschuure, 1974, 1977, 1978; Weber, 1983).

At the same time that Houtgast (1972) related heterophonic continuity of interrupted tones to the masking potential of the interrupter, Warren, Obusek, and Ackroff (1972) proposed a somewhat broader rule designed to encompass not only the heterophonic continuity covered by

Portions of this work were supported by grants from the National Institutes of Health (DC00208) and the Air Force Office of Scientific Research (88-0320). These experiments were outlined in a paper presented at the 119th meeting of the Acoustical Society of America, with a brief abstract appearing in the *Journal of the Acoustical Society* (Vol. 87, Suppl. 1, 1990, S71). The authors are with the Department of Psychology at the University of Wisconsin-Milwaukee, Milwaukee, WI 53201.

Houtgast's rule but also homophonic continuity as well as the phonemic restoration of speech interrupted by louder noises, which had been described a few years earlier (Warren, 1970). This general rule stated "If there is contextual evidence that a sound may be present at a given time, and if the peripheral units stimulated by a louder sound include those which would be stimulated by a fainter sound, then the fainter sound may be heard as present." This restoration of obliterated sounds was called auditory induction (Warren et al., 1972).

While both homophonic continuity and heterophonic continuity involve restoration of segments identical to prior and subsequent portions of the signal, phonemic restorations involve the perceptual synthesis of obliterated fragments based upon portions of a time-varying signal. This third type of auditory induction has been called *contextual catenation*. Phonemic restoration is not the only type of contextual catenation that has been investigated. Dannenbring (1976) found that when tonal glides were interrupted by a louder broadband noise, the tone could be heard to glide through the broadband noise for intervals of a few hundred milliseconds. A different type of tonal restoration was observed by Sasaki (1980) and by DeWitt and Samuel (1990), who reported that missing notes of melodies played on a piano could be restored when those notes were replaced by broadband noise.

Confirmation of the predicted spectral requirements for verbal induction has come in part from studies employing single interruptions of sentences and words: restoration of single speech fragments has been reported to be more robust when the extraneous sound is acoustically similar to the missing speech (e.g., Layton, 1975; Samuel, 1981a, 1981b). Further, in a study employing multiple interruptions, Bashford and Warren (1987b) have obtained evidence of a detailed correspondence between masking and the illusory continuity of speech, as well as evidence of a large effect of linguistic context upon the durational limit of verbal induction. When listeners in that study were presented with narrowband discourse periodically interrupted by different narrow bands of noise, the duration of interruption required for the detection of gaps in speech increased (relative to a silent-gap condition) by as much as 225 msec when gaps were filled with noise having the same center frequency, but increased by only 50 msec with noise bands two octaves removed in center frequency. Similar variation in the durational limits of continuity was observed as a function of context; that is, addition of broadband noise to gaps within broadband speech produced a 250-msec increase in the threshold for gap detection with normal discourse, but only a 100-msec increase both for isolated words and for the same discourse passage read with reversed word order.

Other studies dealing with the effect of interpolated noise have examined the intelligibility of the interrupted speech and have reported an increase in intelligibility when silent gaps in sentences were filled with noise (Bashford & Warren, 1987a; Cherry & Wiley, 1967; Holloway, 1970; Powers & Wilcox, 1977; Verschuure & Brocaar,

1983). However, Miller and Licklider (1950) found that the intelligibility of interrupted word lists did not benefit from noise interpolation even though the effects of interruption were smoothed and the speech seemed more "natural." The present study addresses this discrepancy along with other questions concerning the noise-induced enhancement of speech intelligibility.

Experiment 1 examined acoustic requirements for the restoration of intelligibility using bandpass-filtered sentences that were interrupted either by silence or by different narrow bands of noise.¹ The experiment was undertaken to determine whether the acoustic requirements for restoration of intelligibility match the requirements reported for the induction of continuity of both verbal and nonverbal signals. Experiment 2 examined the effects of contextual information upon intelligibility of broadband speech that was periodically interrupted either by silence or by broadband noise. The speech signals were presented under uniform conditions and consisted of word lists, sentences having low predictability of key words, and sentences having high predictability of key words.

GENERAL METHODS

All speech stimuli were produced by the same male speaker and initially recorded (bandpass filtered from 100 Hz to 10 kHz) on an Ampex 440-C eight-track recorder as described in detail by Bashford and Warren (1987a). For use in the present experiments, the stimuli were rerecorded on an Otari MTR-10 two-track recorder at 15 ips, with the levels of individual sentences (or words) adjusted on input (Grason-Stadler Model 1292 passive attenuator) so that, upon playback, the slow-peak amplitudes of individual items in all stimulus sets varied by less than 2 dB SPL, as measured using a Brüel and Kjaer Model 2204 precision sound-level meter. Speech and interrupting pink noise (when present) were passed to the listeners' headphones by separate electronic switches (Coulbourn Model S84-04) that were triggered alternately with 10-msec rise/fall to produce a 50% duty cycle. The pink noise (equal power per octave) was bandpassed from 100 to 8,000 Hz. On/off times for speech were 175 msec in Experiment 1 and 200 msec in Experiment 2 (full-cycle times 350 and 400 msec, respectively). Listeners in both experiments (a total of 458) were tested individually in an audiometric room with stimuli delivered through diotically wired TDH-49P headphones mounted in MX 41/AR cushions. Speech signals were always presented at a slow-peak level of 70 dBA, and interrupting noise was delivered at 80 dBA.

EXPERIMENT 1

Method

Subjects. The 210 listeners in this experiment were recruited from the introductory psychology course at the University of Wisconsin-Milwaukee. They were given either course credit or \$2 for their participation.

Stimuli. The speech stimuli were bandpass-filtered versions of the CID (Central Institute of the Deaf) sentences used previously

by Bashford and Warren (1987a). These sentences, which were originally designed to serve as samples of "everyday American speech" (Silverman & Hirsh, 1955), are arranged in 10 lists of 10 sentences, with each list containing a total of 50 phonetically balanced key words that vary in syllabic complexity and in positioning within sentences. Using the standard scoring procedure, intelligibility was measured in terms of the percentage of key words that were identified correctly. In the present study, the recorded sentences were filtered on-line using a Rockland Model 852 filter with high-pass and low-pass slopes of 48 dB/octave intersecting at 1500 Hz. For separate groups of listeners, these filtered sentences were interrupted either by silence or by one of five $\frac{1}{2}$ -octave bands of pink noise, derived by passing the output of a Brüel and Kjaer Model 1405 noise generator through a Brüel and Kjaer 2121 frequency analyzer. The noise-band center frequencies were 375, 750, 1500, 3000, and 6000 Hz, as verified using a Hewlett Packard Model 3561A Spectrum Analyzer.

Procedure. Separate groups of 30 listeners each were randomly assigned to the six interruption conditions. Within conditions, the order in which a listener received the 10 lists of sentences was blocked and pseudorandomized so that, across listeners, each 10-sentence list occurred three times in each serial position. Each listener heard each sentence only once. As described above, the listeners were tested individually, with the interrupted sentences (175 msec on-time and 175 msec off-time) delivered diotically through headphones at a slow-peak level of 70 dBA SPL and with interrupting noise (when present) delivered at 80 dBA SPL. Listeners were instructed to repeat each sentence as best they could during a 3-sec interstimulus interval, and they were encouraged to guess when unsure. The experimenter checked whether the key words were correctly identified.

To obtain a baseline measure of the effects of filtering alone upon intelligibility, an additional 30 listeners were asked to repeat uninterrupted versions of the sentences presented at the same ampli-

tude and under the same (1500-Hz) narrowband filtering condition employed in the main experiment.

Results

For the 30 listeners presented with uninterrupted versions of the filtered CID sentences (the baseline condition), average repetition accuracy for the key words was better than 95%.² Figure 1 presents percent repetition accuracy (along with standard errors) for the five interpolated noise-band conditions collapsed across blocks and plotted as a function of noise-band center frequency (overall accuracy for listeners in the silent-gap condition is also shown).

The listeners' percent correct repetition scores were subjected to an analysis of variance (ANOVA) that yielded significant main effects for interrupter [$F(5,174) = 7.72$, $p < .0001$] and blocks [$F(9,1566) = 55.32$, $p < .0001$] and a nonsignificant interaction [$F(45,1566) = 0.609$, $p > .98$]. Subsequent Tukey HSD analysis of the block effect indicated that repetition accuracy increased monotonically up to the ninth block of sentences ($p < .05$ or better), showing that listeners gradually improved in their ability to deal with the combined interference of filtering and interruption. The independent effect of interrupter type was found to be due to the higher repetition accuracy ($p < .05$ or better by Tukey test) obtained for sentences interrupted by the noise band having the same center frequency. The proportion of correct responses with silent gaps was 0.32, compared with 0.41 with the matching noise band, an improvement of 9%. Intelligibility did not

Figure 1. Percent repetition accuracy (and standard error confidence intervals) for narrowband (1500 Hz center frequency) CID sentences ("everyday speech") interrupted by silence or by noise bands of various center frequencies.

differ across the nonmatching noise-band conditions and was statistically equivalent to that obtained in the silent-gap condition.

Discussion

The results obtained in Experiment 1 provide additional support for the hypothesis that phonemic restoration is a specialized linguistic form of auditory induction (Warren et al., 1972; Warren, 1984). As discussed earlier, induction of continuity requires that an interpolated sound be a potential masker of the signal it replaces. The enhancement of intelligibility occurring when narrowband speech and the narrowband noise had matching center frequencies (see Figure 1) demonstrates that this spectral requirement for the illusion of noninterruption also applies to the restoration of intelligibility.

While the improvement in intelligibility of the narrowband speech produced by a frequency-matched narrowband noise follows from temporal induction theory, there are psychoacoustic factors that would be expected to reduce intelligibility. The segments of noise preceding and following each speech fragment can produce forward and backward masking, respectively, with this degradation of the signal being greatest when the noise and signal have overlapping frequencies (see Elliot, 1971). Nevertheless, intelligibility is highest with matching frequencies. It follows that there is an optimal intensity for improving intelligibility by the addition of noise to regularly spaced gaps in speech: the noise should be at a level that would be high enough to be capable of masking the phonemes being replaced (including the high-intensity vowels) but below the level at which nonsimultaneous masking would obliterate information carried by the speech fragments that have not been deleted (including the low-intensity consonants).

EXPERIMENT 2

In an early study dealing with the intelligibility of periodically interrupted speech, Miller and Licklider (1950) observed illusory continuity of isolated monosyllables through brief, noise-filled gaps but found no accompanying increase in repetition accuracy. In contrast, studies employing sentences or prose as stimuli have (as in Experiment 1 above) demonstrated substantial increases in intelligibility due to interpolated noise (Bashford & Warren, 1987a; Cherry & Wiley, 1967; Holloway, 1970; Powers & Wilcox, 1977; Verschuure & Brocaar, 1983). Although this suggests that restoration of intelligibility has certain contextual requirements, another interpretation has been offered by Verschuure and Brocaar (1983). They presented listeners with periodically interrupted versions of syntactically normal sentences that were either semantically normal or anomalous (e.g., "The church bell drank a sailor") and found increased intelligibility with noise interpolation for both sentence types. Considering their semantically anomalous sentences to be equivalent to word lists, Verschuure and Brocaar suggested that the earlier

failure of Miller and Licklider (1950) to obtain an objective effect upon intelligibility was due to the use of trained listeners who presumably had acquired sufficient skill in listening to interrupted speech to obviate the need for interpolated noise. However, semantically anomalous sentences may not be functionally equivalent to word lists. Word lists are articulated as separate utterances, with pauses between the successive items to permit listeners to respond. These pauses are absent in sentences, so that coarticulation links adjacent words and provides reciprocal allophonic cues across word boundaries. There are also additional prosodic and syntactic cues to aid comprehension of sentences. According to the LAME model (lateral access from multilevel engrams), mechanisms employed for comprehending speech are variable and opportunistic, changing with both the particular stimulus and the task requirements, and with interactions occurring simultaneously within and across different levels of linguistic organization (Warren, 1981).

The present experiment was designed to permit a direct comparison of restoration effects for naive listeners presented with isolated words or syntactically and prosodically normal sentences that provided varying amounts of semantic context. In addition, the results obtained permitted a comparison of restoration of continuity of speech (reported previously) with restoration of intelligibility. As we shall see, continuity and intelligibility effects produced by interpolated noise are experimentally separable.

Method

Subjects. Separate groups of introductory psychology students (a total of 248 subjects) who did not participate in Experiment 1 served in the eight listening conditions of Experiment 2 (four different types of broadband speech were periodically interrupted either by silence or by broadband pink noise). As in Experiment 1, the listeners were given either course credit or \$2 for their participation. For purposes of counterbalancing to be discussed, groups of 30 listeners were assigned to conditions involving CID sentences or isolated words, and groups of 32 listeners each were assigned to conditions involving the SPIN test sentences described below.

Stimuli. In addition to the 10 lists of CID ("everyday speech") sentences employed in Experiment 1, there were three conditions involving monosyllabic key words: (1) Ten 50-word lists of monosyllables (500 key words) taken from the same phonetically balanced (PB) lists used by Miller and Licklider (1950) and produced in isolation at a rate of one word each 3 sec, (2) 8 25-item lists with the 200 key monosyllables appearing as final words within nonpredictive carrier sentences (e.g., "Ruth hopes she called about the junk"), and (3) 8 matching lists with the same 200 key monosyllables appearing as the final word in highly predictive carrier sentences ("Throw out all this useless junk"). The latter two conditions were derived from the SPIN (speech perception in noise) test of Kalikow, Stevens, and Elliot (1977).

Procedure. As in Experiment 1, the order of presentation for stimulus lists was blocked and counterbalanced, with the 10 lists of CID sentences and the 10 lists of isolated monosyllables occurring three times in each of eight blocks across listeners and with the 8 lists of low-predictable and high-predictable SPIN sentences occurring four times in each block across listeners. The stimuli were interrupted periodically (200-msec on/off time, 400-msec full cycle), with speech presented at a slow-peak level of 70 dBA SPL. The broadband pink noise was bandpassed from 100 to 8,000 Hz and when introduced into the gaps in the speech was delivered at 80 dBA

SPL. In keeping with standard procedures, the listeners assigned to the CID-sentence and isolated-monosyllable conditions were instructed to repeat everything they heard, whereas the listeners assigned to the low- and high-predictable SPIN test conditions were required to repeat only the last (key) word of each sentence. As in Experiment 1, the listeners responded during a 3-sec interstimulus interval and they were encouraged to guess when unsure. Scoring was based upon the percentage of key words identified correctly.

Results

Figure 2 presents average percent correct repetition scores (and standard error confidence intervals) for the eight experimental conditions (four types of speech $\times$ two types of interrupter). Percent accuracy scores were subjected to a three-way ANOVA that yielded significant main effects of interrupter [$F(1,240) = 159.76, p < .0001$], context [$F(3,240) = 465.84, p < .0001$], and blocks [$F(7,1680) = 43.63, p < .0001$] and also yielded significant two-way interactions of interrupter $\times$ context [$F(3,240) = 22.08, p < .0001$], interrupter $\times$ blocks [$F(7,1680) = 5.09, p < .0001$], and context $\times$ blocks [$F(21,1680) = 2.78, p < .0001$]. The three-way interaction was not significant [$F(21,1680) = 0.74, p > .79$].

The interrupter $\times$ context interaction was found through simple effects tests to be due in part to the absence of a noise-induced restoration of intelligibility for monosyllables presented in isolation [$F(1,240) = 0.0, p > .9$]. Noise interpolation in the remaining conditions did produce a significant increase in repetition accuracy [$F(1,240) \geq 43.4, p < .001$] for low-predictable SPIN sentences (11.4%), CID sentences (14.3%), and high-predictable SPIN sentences (19.1%). Subsequent three-way ANOVAs for each pairwise combination of sentential context conditions were conducted, yielding a significant interrupter $\times$ context effect [$F(1,120) = 8.04, p < .01$] only for

low- versus high-predictable SPIN sentences. Thus, restoration of intelligibility was somewhat greater for monosyllables in predictive carrier sentences than for the same words in neutral carriers.

Simple effects analysis of the interrupter $\times$ blocks interaction indicated that the interrupter effect was significant at each block [$F(1,1226) \geq 21.2, p < .0001$], whereas the effect of blocks was significant only with interruption by silence [$F(7,1680) = 74.5, p < .0001$]. Thus restoration of intelligibility occurred throughout the eight blocks, but whereas performance in the interpolated-noise conditions remained constant over blocks, performance in the silent-gap conditions gradually improved. This finding contrasts with the results of Experiment 1, in which repetition accuracy for CID sentences improved across blocks both with and without interpolated noise. It seems likely that this was due to the fact that listeners in Experiment 1 were required to adjust both to interruption and to a highly restricted stimulus bandwidth.

Discussion

The gradual improvement found for repetition accuracy with successive blocks of sentences in the silent-gap conditions indicates that, as suggested by Verschuure and Brocaar (1983), listeners given sufficient exposure to this novel form of interruption might eventually be able to make optimal use of information provided by remnants of the speech signal without the need for interpolated noise. However, Verschuure and Brocaar also suggested that the lack of improvement in intelligibility found in previous studies when noise was added to silent gaps in word lists (Dirks & Bower, 1970; Miller & Licklider, 1950) was attributable to the use of trained listeners who no longer benefited from noise interpolation. Our present

Figure 2. Percent repetition accuracy (and standard error confidence intervals) for four types of broadband speech interrupted by silence or by broadband pink noise.

finding that the repetition accuracy of untrained listeners did not improve with noise interpolation in the isolated-monosyllable condition conflicts with that hypothesis.

The inability of noise to enhance intelligibility of interrupted word lists is especially interesting in view of reports in the literature that the addition of noise to gaps in isolated words does produce illusory continuity (Bashford & Warren, 1987b; Miller & Licklider, 1950). Apparently, even though the information required to identify the missing phonetic components is lacking with isolated monosyllables, illusory continuity occurs nevertheless, as it does with nonverbal acoustic signals.

The finding that intelligibility increases with low-predictability SPIN sentences indicates that sentential factors other than semantically based predictability play important roles in the restoration of intelligibility. Low-predictable sentences are syntactically legal, and syntactic requirements can eliminate some lexical candidates from consideration; however, there are still great numbers of lexical candidates when semantic restrictions are minimal. The coarticulation linking adjacent words might provide information concerning the phonemes of the key words contiguous with other words. In addition, suprasegmental prosodic cues including such features as intonation and stress might produce a normal-sounding framework that could facilitate restoration. As would be anticipated, the availability of semantic pointers to key words in the high-predictability SPIN test enhances performance both when interruptions remain silent and when they are filled with noise. As can be seen in Figure 2, the slopes of the lines marking the improvement in intelligibility produced by the introduction of noise in gaps is steeper for high-SPIN than low-SPIN sentences, suggesting that the availability of semantic information is improved by the addition of noise. However, this statement must be tempered by the evidence furnished by the low-SPIN sentences that prosodic features play an important role.

SUMMARY AND CONCLUSIONS

The results obtained in Experiment 1 provide additional support for the theory that considers phonemic restoration to be a specialized, linguistic form of auditory induction (Warren, 1984; Warren et al., 1972). Induction of continuity requires that an interpolated sound be a potential masker of the signal it replaces, and as shown in Experiment 1, the enhanced intelligibility occurring when the narrowband speech and the narrowband noise have matched center frequency indicates that this spectral requirement also applies to the restoration of intelligibility. Experiment 2 measured the effects of context upon the restoration of intelligibility and provides evidence that several factors play an important role. Although word-list intelligibility did not benefit from interpolated noise, restoration did occur when words were embedded within a sentential framework. It appears that several types of cues are employed for restoring intelligibility in frag-

mented sentences. Possible facilitating cues include prosodic and semantic features as well as coarticulation. It is suggested that many sources of information at different linguistic levels are employed in an interactive fashion in keeping with the LAME model of speech perception (Warren, 1981).

It appears that filling gaps in speech with noise invokes a two-stage process of perceptual synthesis. The first, more primitive stage of auditory induction applies to both verbal and nonverbal signals interrupted briefly by a potential masker. This initial form of induction produces an apparent continuity of the signal, rendering interruptions undetectable, but when applied to speech does not involve the synthesis of specific, meaningful components. A second, knowledge-driven stage of induction can apply linguistic rules to connected speech, resulting in the interpolation of appropriate speech segments and restoration of the message. It is suggested that care must be taken in studies dealing with the perception of interrupted speech to distinguish between these two levels of induction.

REFERENCES

- BASHFORD, J. A., JR., & WARREN, R. M. (1979). Perceptual synthesis of deleted phonemes. In J. J. Wolf & D. H. Klatt (Eds.), *Speech communication papers presented at the 97th meeting of the Acoustical Society of America* (pp. 423-426). New York: Acoustical Society of America.
- BASHFORD, J. A., JR., & WARREN, R. M. (1987a). Effects of spectral alternation on the intelligibility of words and sentences. *Perception & Psychophysics*, **42**, 431-438.
- BASHFORD, J. A., JR., & WARREN, R. M. (1987b). Multiple phonemic restorations follow the rules for auditory induction. *Perception & Psychophysics*, **42**, 114-121.
- CHERRY, E., & WILEY, R. (1967). Speech communications in very noisy environments. *Nature*, **214**, 1164.
- DANNENBRING, G. L. (1976). Perceived auditory continuity with alternately rising and falling frequency transitions. *Canadian Journal of Psychology*, **30**, 99-114.
- DEWITT, L. A., & SAMUEL, A. G. (1990). The role of knowledge-based expectations in music perception: Evidence from musical restoration. *Journal of Experimental Psychology: General*, **119**, 123-144.
- DIRKS, D. D., & BOWER, D. (1970). Effect of forward and backward masking on speech intelligibility. *Journal of the Acoustical Society of America*, **47**, 1003-1007.
- ELLIOT, L. L. (1971). Backward and forward masking. *Audiology*, **10**, 65-76.
- FASTL, H. (1975). Pulsation patterns of sinusoids vs. critical band noise. *Perception & Psychophysics*, **18**, 95-97.
- HOLLOWAY, C. M. (1970). Passing the strongly voiced components of noisy speech. *Nature*, **226**, 178-179.
- HOUTGAST, T. (1972). Psychophysical evidence for lateral inhibition in hearing. *Journal of the Acoustical Society of America*, **51**, 1885-1894.
- HOUTGAST, T. (1973). Psychophysical experiments on "tuning curves" and "two-tone inhibition." *Acustica*, **29**, 168-179.
- HOUTGAST, T. (1974a). Masking patterns and lateral inhibition. In E. Zwicker & E. Terhardt (Eds.), *Facts and models in hearing* (pp. 258-265). Berlin: Springer-Verlag.
- HOUTGAST, T. (1974b). The slopes of masking patterns. In E. Zwicker & T. Terhardt (Eds.), *Facts and models in hearing* (pp. 269-272). Berlin: Springer-Verlag.
- KALKOW, D. N., STEVENS, K. N., & ELLIOT, L. L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *Journal of the Acoustical Society of America*, **61**, 1337-1351.

- KRONBERG, H., MELLERT, V., & SCHREINER, C. (1974). Dichotic and monaural pulsation thresholds. *Proceedings of the 8th International Congress on Acoustics (London)*, 1, 143.
- LAYTON, B. (1975). Differential effects of two nonspeech sounds on phonemic restoration. *Bulletin of the Psychonomic Society*, 6, 487-490.
- MILLER, G. A., & LICKLIDER, J. C. R. (1950). The intelligibility of interrupted speech. *Journal of the Acoustical Society of America*, 22, 167-173.
- POWERS, G. L., & WILCOX, J. C. (1977). Intelligibility of temporally interrupted speech with and without intervening noise. *Journal of the Acoustical Society of America*, 61, 195-199.
- SAMUEL, A. G. (1981a). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General*, 110, 474-494.
- SAMUEL, A. G. (1981b). The role of bottom-up confirmation in the phonemic restoration illusion. *Journal of Experimental Psychology: Human Perception & Performance*, 7, 1124-1131.
- SASAKI, T. (1980). Sound restoration and temporal localization of noise in speech and music sounds. *Tohoku Psychologica Folia*, 39, 79-88.
- SILVERMAN, S. R., & HIRSH, I. J. (1955). Problems related to the use of speech in clinical audiometry. *Annals of Otology, Rhinology, & Laryngology*, 64, 1234-1245.
- VERSCHUURE, J. (1974). Transient phenomena in pulsation threshold measurement. *Audiology*, 13, 90.
- VERSCHUURE, J. (1977). Pulsation threshold patterns and neurophysiological tuning. In E. F. Evans & J. P. Wilson (Eds.), *Psychophysics and physiology of hearing* (pp. 237-247). London: Academic Press.
- VERSCHUURE, J. (1978). *Auditory excitation patterns: The significance of the pulsation threshold method for the measurement of auditory nonlinearity*. Unpublished doctoral dissertation, Erasmus University, Rotterdam, Netherlands.
- VERSCHUURE, J., & BROCAAR, M. P. (1983). Intelligibility of interrupted meaningful and nonsense speech with and without intervening noise. *Perception & Psychophysics*, 33, 232-240.
- WARREN, R. M. (1970). Perceptual restoration of missing speech sounds. *Science*, 167, 392-393.
- WARREN, R. M. (1981). Mode of representation in production and perception, Chairman's comments. In T. Myers, J. Laver, & J. Anderson (Eds.), *The cognitive representation of speech* (pp. 34-37). Amsterdam: North-Holland.
- WARREN, R. M. (1982). *Auditory perception: A new synthesis*. New York: Pergamon.
- WARREN, R. M. (1984). Perceptual restoration of obliterated sounds. *Psychological Bulletin*, 96, 371-383.
- WARREN, R. M., OBUSEK, J. M., & ACKROFF, J. M. (1972). Auditory induction: Perceptual synthesis of absent sounds. *Science*, 176, 1149-1151.
- WEBER, D. L. (1983). Do off-frequency simultaneous maskers suppress the signal? *Journal of the Acoustical Society of America*, 73, 887-893.

NOTES

1. A study employing conditions similar to those of Experiment 1 was reported earlier at a meeting of the Acoustical Society of America (Bashford & Warren, 1979; see also Warren, 1982, pp. 154-155). The findings obtained for interpolated noise bands are in agreement with those obtained in the present study; however, the earlier experiment employed fewer subjects and did not include a silent-gap condition.
2. The bases for high intelligibility of the narrowband speech is beyond the scope of the present study. However, it is an effect of intrinsic interest since it indicates that intelligibility can be achieved using a quite limited frequency range that is also subject to the spectral tilt produced by the intersecting filter slopes (48 dB/octave). We are engaged in further examination of this phenomenon.

(Manuscript received August 31, 1990;
revision accepted for publication October 1, 1991.)