Perception & Psychophysics
1995, 57 (2), 175-182

Spectral redundancy: Intelligibility of sentences heard through narrow spectral slits

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

The intelligibility of word lists subjected to various types of spectral filtering has been studied extensively. Although words used for communication are usually present in sentences rather than lists, there has been no systematic report of the intelligibility of lexical components of narrowband sentences. In the present study, we found that surprisingly little spectral information is required to identify component words when sentences are heard through narrow spectral slits. Four hundred twenty listeners (21 groups of 20 subjects) were each presented with 100 bandpass filtered CID ("everyday speech") sentences; separate groups received center frequencies of 370, 530, 750, 1100, 1500, 2100, 3000, 4200, and 6000 Hz at 70 dBA SPL. In Experiment 1, intelligibility of single 1/3-octave bands with steep filter slopes (96 dB/octave) averaged more than 95% for sentences centered at 1100, 1500, and 2100 Hz. In Experiment 2, we used the same center frequencies with extremely narrow bands (slopes of 115 dB/octave intersecting at the center frequency, resulting in a nominal bandwidth of 1/20 octave). Despite the severe spectral tilt for all frequencies of this impoverished spectrum, intelligibility remained relatively high for most bands, with the greatest intelligibility (77%) at 1500 Hz. In Experiments 1 and 2, the bands centered at 370 and 6000 Hz provided little useful information when presented individually, but in each experiment they interacted synergistically when combined. The present findings demonstrate the adaptive flexibility of mechanisms used for speech perception and are discussed in the context of the LAME model of opportunistic multilevel processing.

In order to function as an effective communication vehicle, speech must remain intelligible under noisy conditions that mask portions of the signal. Experiments have shown that when phonemes in sentences are completely obliterated by noises, listeners can restore the missing speech sounds and maintain intelligibility (for reviews of phonemic restoration, see Warren, 1984, and Warren, in press). In addition to the complete obliteration of brief segments, extended extraneous sounds may block some spectral regions of speech, but allow others to pass. In the present study, we examined the intelligibility of sentences when only narrow spectral regions could be heard.

An earlier study in this laboratory had indicated that sentences consisting of a very narrow band of frequencies had an unexpectedly high intelligibility (Bashford, Riener, & Warren, 1992). This study dealt with the phonemic restoration of interrupted sentences subjected to narrowband filtering. One of the control conditions employed uninterrupted sentences that had been bandpass filtered with highpass and lowpass slopes (48 dB/octave) set at the same cutoff frequency value of 1500 Hz. Despite the extremely limited frequency range of spectral components and the spectral tilts of the intersecting filter slopes, the intelligibility of key words was better than 95%. Surprisingly, a search of the literature revealed no other direct measurement of the intelligibility of clear sentences that had been bandpass filtered. Since such high intelligibility based upon so little spectral information has implications for models of speech processing, the present study was undertaken to systematically measure listeners' ability to comprehend sentences limited to narrow spectral regions.

There have been many studies of the changes in speech intelligibility produced by systematically changing the cutoff frequencies of highpass and lowpass speech. The frequency importance functions derived from these measurements have been used for the construction of the Articulation Index, or AI (ANSI S3.5, 1969). The initial studies leading to the development of the AI were conducted at Bell Telephone Laboratories to provide a procedure that could be used to design and evaluate communication systems (see French & Steinberg, 1947; Kryter, 1962). The original frequency importance functions were developed for nonsense syllables, in part because of the sensitivity of functions based upon these stimuli to the effects of degrading speech. Since the mid 1980s, there have been measurements of importance functions for sentences and continuous discourse. Because both highpass and lowpass sentential material are completely intelligible for midrange cutoff

This study was supported by a grant from the National Institute of Deafness and Other Communication Disorders (DC 00208) to the first author. The valuable contributions of Christine Van Ryzin and Eric Healy are gratefully acknowledged. Address correspondence to R. M. Warren, University of Wisconsin, Milwaukee, Department of Psychology, Garland Hall, P. O. Box 413, Milwaukee, WI 53201 (e-mail: rmwarren@csd4.csd.uwm.edu).

frequencies, it was necessary to add masking noise to further degrade these signals so that intelligibility would vary with cutoff changes (see Studebaker, Pavlovic, & Sherbecoe, 1987). Pavlovic (1994) summarized some of this recent work and listed 1/3-octave band importance functions for a variety of speech materials, including sentences, short passages, and average speech. In the spectral region representing the highest contribution to intelligibility (approximately 1500–2500 Hz, depending upon the nature of the speech material), individual 1/3-octave bands each had importance values ranging from roughly 0.08 to 0.11 corresponding to only 9% to 11% of the total information content of broadband speech. Yet Bashford et al. (1992) reported that bandpassed sentences with a center frequency of 1500 Hz and a much narrower nominal bandwidth of only 1/8 octave had an intelligibility of over 95%. This high intelligibility of a single narrow band is not inconsistent with the band importance function of AI, since there is considerable redundancy of information across spectral bands, but it does illustrate that AI measurements cannot be used to predict the intelligibility of stand-alone bands of speech.

In the present study, 100 sentences were presented to each of 420 subjects. These stimuli were bandpass filtered at nine different center frequencies, ranging from 370 to 6000 Hz. In Experiment 1, the bands were 1/3 octave wide and had filter slopes of 96 dB/octave. In Experiment 2, the bands were extremely narrow, with intersecting highpass and lowpass slopes of 115 dB/octave; this resulted in a nominal bandwidth of 1/20 octave (as defined by 3-dB down-points). All the sentences were presented at a slow-peak amplitude level of 70 dBA, and effects of prior experimental conditions were avoided by presenting the individual listeners with only one filtering condition.

In addition to measuring intelligibility on the basis of single passbands, we also determined the intelligibility of these sentences when two widely separated bands (center frequencies of 370 and 6000 Hz), which had low intelligibility when heard separately, were presented simultaneously.

EXPERIMENT 1

Method

Subjects. The 220 subjects (11 groups of 20 subjects) were native English speakers with no known hearing problems, recruited from introductory psychology courses at the University of Wisconsin-Milwaukee. They were given money or course credit for participation.

Stimuli. The experimental stimuli consisted of 100 CID (Central Institute of the Deaf) sentences, originally designed to represent "everyday American speech" (Silverman & Hirsh, 1955). They are arranged in 10 sets of 10 sentences each, with each set containing 50 key words. The key words vary in syllabic composition and position within each sentence. An additional 10 sentences were used which were presented before the experimental stimuli. These practice sentences were taken from the high-predictability subset of the SPIN (Speech Perception in Noise) test of Kalikow, Stevens, and Elliot (1977; e.g., "Throw out all this useless junk.").

The stimuli were produced in an audiometric chamber (IAC Series 400A) by a male speaker with a General American dialect and a voicing frequency of approximately 100 Hz. The output of an AKG microphone (Model C414B-ULS) was amplified (AKG N46E1 preamplifier), passed through a Ramsa 8-channel mixer (WR-133), and digitized at a sampling rate of 22 kHz with 16-bit resolution, using a Digidesign sound accelerator card running on a Macintosh IIfx computer.

To prepare the bandpassed experimental sentences, the signal was converted from digital to analog form. It was then passed successively through two bandpass filters, each having slopes of 48 dB/octave (Rockland Model 852) resulting in an attenuation of 96 dB/octave, and recorded at 15 ips on an Otari 8-track recorder (Model MX5050). Nine recorded versions of the stimuli were produced with corresponding center frequencies of 370, 530, 750, 1100, 1500, 2100, 3000, 4200, and 6000 Hz. The amplitude envelopes of the individual broadband sentences were rescaled for each filtered version prior to the analog recording so that the slow peak of each filtered sentence (measured after transduction through Sennheiser HD 250 headphones) was within ±2 dB of the nominal 70-dB presentation level, as measured with a Brüel & Kjaer Model 2204 precision sound-level meter operating in its slow-peak mode. The spectra of all the stimuli were examined (after transduction) using a Hewlett-Packard 3561A Dynamic Signal Analyzer to confirm that the specifications of the bandpass filtering were met and to ensure that the signals contained no spectral artifacts. Across conditions, the slow amplitude peaks of the sentences for all the narrow bands ranged from 32 to 38 dB above the noise floor.

The recorded stimuli were presented diotically through headphones. All single bands were presented at 70 dBA SPL. The levels of the 370- and 6000-Hz bands recorded simultaneously for the test of spectral interaction were each reduced to approximately 67 dB to achieve a combined level approximating that of the individual bands when presented separately.

Procedure. Separate groups of 20 listeners were assigned to each of the nine filtering conditions. Two additional groups of 20 subjects heard the 370-Hz and the 6000-Hz bands presented together. One of these groups heard the two bands diotically, and the other group heard the bands presented dichotically; 10 listeners heard the low band at the right ear and the high band at the left ear, and for the remaining 10 listeners the ear of presentation was reversed. The order of the sentence sets was pseudorandomized and blocked within conditions so that each set of 10 sentences occurred twice in each serial position. The listeners were tested individually while seated in the audiometric chamber with the experimenter. Testing took about 30 min for each listener, and each listener was exposed to only one filtering condition.

Before beginning the actual experiment, the listeners in each group were acquainted with the effects of filtering. They were presented with the separate list of 10 practice sentences twice—first broadband, and then filtered in the same manner as that for the experimental sentences that followed. Correct responses to the initial broadband presentation of the sentences were intended to serve as a criterion for participation in the formal experiment. However, none of the listeners were rejected (they all achieved a perfect score with the 10 broadband sentences).

Following the initial practice phase of the experiment, each group of listeners was presented with the 10 sets of 10 CID sentences. The groups differed only in the center frequency of the band presented to them. The listeners were instructed to repeat each sentence as accurately as possible and were encouraged to guess if unsure. The experimenter scored the percentage of key words reported correctly for each sentence.

Results

Figure 1 presents the mean percent intelligibility scores obtained across the 10 blocks of CID sentences

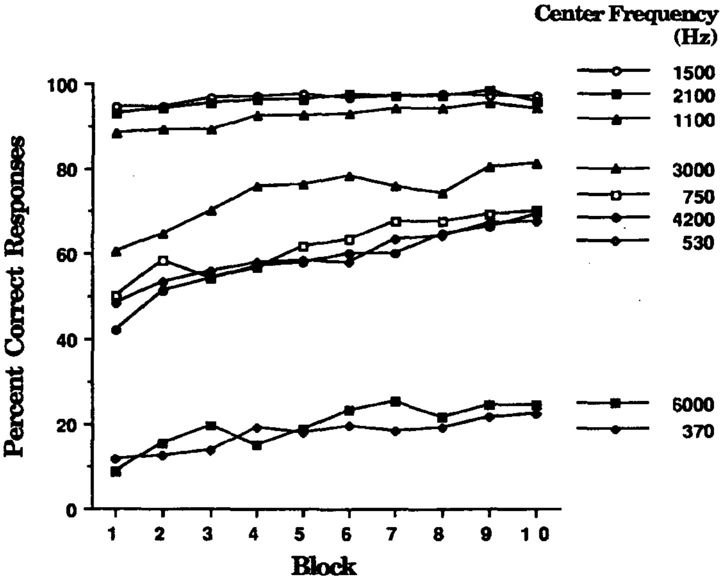

Figure 1. Mean percentage of key words reported correctly across 10 blocks of narrowband CID sentences (1/3 octave; 96-dB/octave slopes) presented at nine different center frequencies.

for each filtering condition. It can be seen that, despite the narrow bandwidth and steep filter slopes, intelligibility was high for all but the lowest (370 Hz) and the highest (6000 Hz) center frequencies.

An analysis of variance (ANOVA) yielded significant main effects of center frequency [$F(8,171) = 285.6, p < .0001$] and block [$F(9,1539) = 28.8, p < .0001$] and a significant interaction of those factors [$F(72,1539) = 1.5, p < .01$]. Subsequent analyses employing Tukey's HSD tests revealed that the speech bands centered at 370 and 6000 Hz were statistically equivalent and produced the lowest intelligibility for each of the blocks ($p < .01$), whereas the midrange speech bands with center frequencies of 1100, 1500, and 2100 Hz were equal and produced the highest intelligibility ($p < .05$), averaging about 95% across blocks. Intermediate intelligibility scores were obtained for the 530-, 750-, and 4200-Hz speech bands, which were statistically equivalent and, at most blocks, produced somewhat lower intelligibility ($p < .05$) than did the 3000-Hz speech band.

Simple effects tests revealed that improvement occurred across blocks in each filtering condition [$F(9,171) \geq 2.14, p \leq .035$]. However, a planned ANOVA performed on the data obtained for the 9th and 10th blocks of sentences yielded a nonsignificant block effect [$F(1,171) = 0.1, p > .80$] as well as a nonsignificant interaction of block and center frequency [$F(8,171) = 0.2, p > .90$]. This suggested that, for all conditions, the listeners' ability to deal with the effects of spectral distortion reached an asymptotic level by the 9th block of sentences.

The higher intelligibility obtained for the bandpassed sentences centered at 1100, 1500, and 2100 Hz is consistent with the long-held view that the middle frequency range is the richest information-bearing region of the speech spectrum (Egan & Wiener, 1946; French & Steinberg, 1947; Kryter, 1962). However, as will be discussed subsequently, previous reports in the literature would not lead us to expect the high intelligibility found for all center frequencies except the lowest (370 Hz) and the highest (6000 Hz).

Figure 2 presents the intelligibility scores and the standard errors for the 10th block of sentences when presented as single bands centered at 370 and 6000 Hz, and when these widely separated bands were presented together under both diotic and dichotic listening conditions. An ANOVA performed on all of the data for the

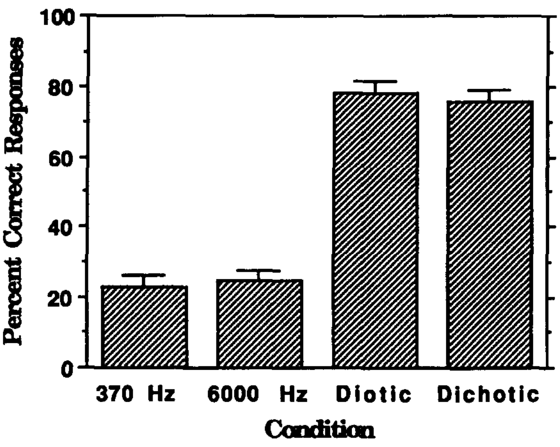

Figure 2. Mean percentage of key words reported correctly for the 10th block of narrowband sentences (1/3-octave bands with center frequencies of 370 and 6000 Hz) presented either individually or in combination under diotic and dichotic listening conditions. Standard error bars are included.

two dual-band conditions yielded a significant block effect [$F(9,342) = 5.81, p < .0001$], a nonsignificant effect of presentation mode [$F(1,38) = 1.63, p > .20$], and a nonsignificant interaction of presentation mode and blocks [$F(9,342) = 0.47, p > .80$], indicating that intelligibilities for diotic and dichotic modes of presentation were equivalent. As can be seen in Figure 2, the increase in intelligibility for dual-band presentation appears to be more than additive. The intelligibility of the 370-Hz band was 23% and that of the 6000-Hz band was 24% when they were heard alone, and when the bands were combined, intelligibility was 78% under diotic presentation and 76% under dichotic presentation. Separate planned comparisons (two-tailed, one-sample t tests) showed that performance in the dual-band conditions was significantly greater [$t(19) > 8.63, p < .0001$] than the sum of the single-band conditions.

EXPERIMENT 2

In Experiment 2, the stimuli and conditions were the same as those in Experiment 1, except for the use of much narrower bands and slightly steeper slopes. This experiment was designed to determine the extent to which the speech processing system can utilize information restricted to extremely limited spectral ranges. In addition, dual bands were used to test for spectral interaction, as in Experiment 1. Since diotic and dichotic presentation resulted in equivalent intelligibility scores in the first experiment, only diotic presentation was used in this experiment.

Method

Subjects. Two hundred additional listeners (10 groups of 20 subjects) were recruited from introductory psychology courses at the University of Wisconsin-Milwaukee. The qualifications and compensation procedures used for Experiment 1 were employed in this experiment as well.

Stimuli. The stimuli were derived from the digital broadband recordings used in Experiment 1. The 9 center frequencies were those used in Experiment 1 (370, 530, 750, 1100, 1500, 2100, 3000, 4200, and 6000 Hz were used). Following digital-to-analog conversion, the signal was subjected to 2/3-octave prefiltering (Rockland Model 852, with slopes of 48 dB/octave) at each center frequency. (This prefiltering eliminated spectral components outside the desired passband, which could overload the second filter, and resulted in an increased signal-to-noise ratio for the experimental stimulus.) Following prefiltering, the sentences were filtered with intersecting 115-dB/octave slopes (Wavetek/Rockland Model 751A Brickwall filter) produced by setting the high and low cutoffs of the bandpass filter at the same frequency. This final filtering produced stimuli with a nominal bandwidth (3-dB downpoints) of 1/20 octave. The amplitude envelopes of each of the 110 broadband sentences (10 practice and 100 experimental sentences) were rescaled so that, when recorded and played back on an Otari two-track (Model MTR-10) recorder, they would be at the level employed in Experiment 1. The slow amplitude peaks of the sentences for all the narrow bands ranged from 32 to 37.5 dB above the noise floor.

Procedure. The procedure was the same as that used with the 1/3-octave bands in Experiment 1, except that for the dual-band condition (simultaneous presentation of the 370- and 6000-Hz bands), only diotic presentation was employed.

Results

Figure 3 presents the mean percent intelligibility scores obtained across the 10 blocks of CID sentences in each filtering condition of Experiment 2. It can be seen that, despite the extremely narrow bandwidth (intersecting highpass and lowpass filter slopes of 115 dB/octave), intelligibility was relatively high for the middle range of center frequencies.

The sentence intelligibility scores were subjected to an ANOVA, which yielded significant main effects of filtering condition [$F(8,171) = 74.1, p < .0001$] and

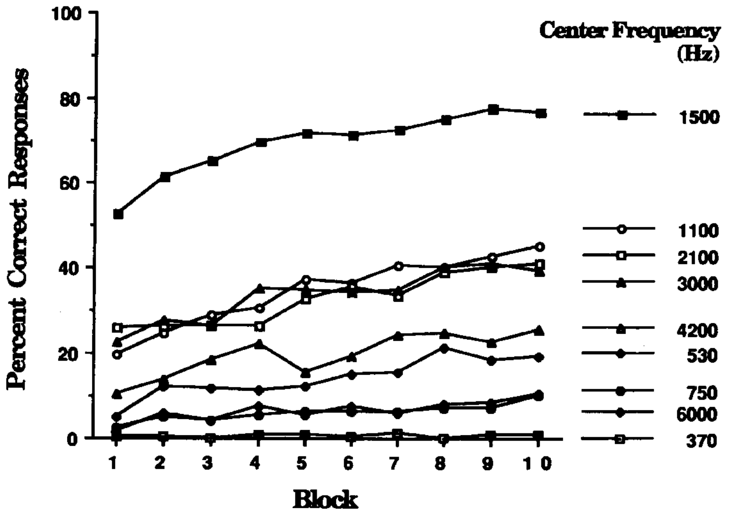

Figure 3. Mean percentage of key words reported correctly across 10 blocks of narrowband CID sentences (intersecting 115-dB/octave slopes producing 1/20-octave bands) at nine different center frequencies.

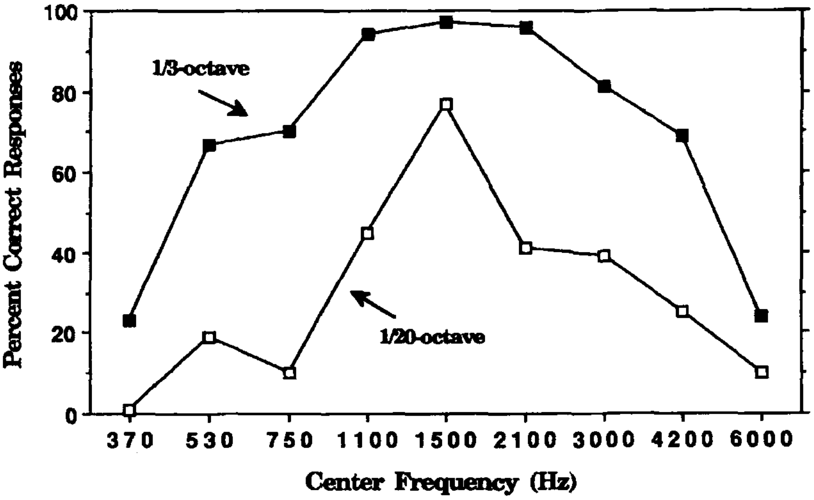

Figure 4. Mean percentage of key words reported correctly for the 10th block of CID sentences presented as 1/3-octave bands (Experiment 1) and 1/20-octave bands (Experiment 2) at nine center frequencies.

block [$F(9,1539) = 35.0, p < .0001$] and a significant filtering condition × block interaction [$F(9,1539) = 2.0, p < .0001$]. Subsequent Tukey HSD testing of the filtering condition effect revealed that sentence intelligibility at each block was highest ($p < .01$) for the 1500-Hz center frequency condition, which yielded an average intelligibility of 77% by the final block of sentences. Simple effects tests revealed that intelligibility improved across blocks [$F(9,171) \geq 2.07, p \leq .04$] for all speech bands except the 370-Hz band [$F(9,171) = 1.05, p > .30$]. However, a planned ANOVA examining performance across the 9th and 10th blocks of sentences yielded a nonsignificant effect of block [$F(1,171) = 0.9, p > .30$] and a nonsignificant interaction between block and filtering conditions [$F(8,171) = 0.3, p > .90$]. This indicated that, at each center frequency, performance reached an asymptotic level by the end of an experimental session—as was found for the 1/3-octave bands in Experiment 1. For the purpose of direct comparison, Figure 4 presents the mean intelligibility scores for the final block of sentences in each filtering condition of Experiments 1 and 2.

As in Experiment 1, the 370-Hz band and the 6000-Hz band were also presented simultaneously for a test of spectral interaction. Figure 5 presents the mean intelligibility scores and the standard errors for the 10th block of sentences for the high and low bands presented separately and presented together. A planned one-sample t test revealed that the intelligibility of the dual bands (27.8%) was significantly greater [$t(19) = 5.81, p < .0001$] than the summed intelligibility scores for the 370-Hz band (0.9%) and the 6000-Hz band (10%). Thus, although the 370-Hz band had negligible intelligibility when heard alone, it produced a dramatic increase in intelligibility when added to the 6000-Hz band.

DISCUSSION

In this study, we examined the intelligibility of sentences heard through narrow spectral slits. We found that these bandpassed sentences maintained a remarkably high intelligibility even though features characterizing the component phonemes were severely distorted or absent.

The stimuli employed in Experiments 1 and 2 were produced by the bandpass filtering of 100 CID sentences designed to represent "everyday American speech" (Silverman & Hirsh, 1955). Their center frequencies ranged from 370 to 6000 Hz, and they were presented to sepa-

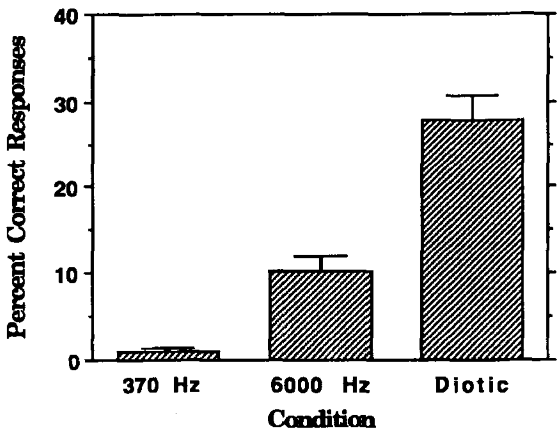

Figure 5. Mean percentage of key words reported correctly for the 10th block of narrowband sentences (115-dB/octave intersecting slopes producing 1/20-octave bands, centered at 370 and 6000 Hz) presented individually or under combined diotic listening. Standard error bars are included.

rate groups of 20 subjects for each of 21 filtering conditions. Steep filter slopes were used throughout (approximately 100 dB/octave), and high intelligibility was found for 1/3-octave band sentences over a wide range of center frequencies (Experiment 1). Sentence intelligibility was still quite high for extremely narrow bands with intersecting lowpass and highpass slopes (Experiment 2), which resulted in nominal bandwidths of 1/20 octave.[1] In both experiments, bands at the limits of the frequency range examined in this study (370 and 6000 Hz) had very low intelligibilities, but when the bands were combined, a synergistic integration of information occurred. The increase in intelligibility in Experiment 2 was quite dramatic: Although the intelligibility of the 370-Hz band was less than 1%, and that of the 6000-Hz band was 10% when they were heard alone, intelligibility rose to 28% when the bands were presented simultaneously. Although relatively modest overadditivity of intelligibility scores for bands of speech presented simultaneously have been reported previously (see Braida et al., 1979, for a summary of the literature), these previous studies did not show a synergistic effect approaching the magnitude found in the present study, presumably because, unlike in the present study, widely separated narrow bands were not used.

As vocal utterances increase in complexity and more closely approximate normal discourse, identification becomes more accurate and more resistant to interference and distortion. Thus, while identification of broadband isolated vowels presented loudly and clearly is quite poor (Bond, 1976; Morton & Carpenter, 1962), isolated monosyllables can be identified with almost perfect accuracy when they are played loudly and clearly. Other studies have shown that when monosyllabic words are acoustically degraded, their intelligibility increases with word familiarity (Epstein, Giolas, & Owens, 1968; Owens, 1961). Epstein et al. concluded that as long as the spectrum up to approximately 2000 Hz remained intact, performance for familiar words was similar to that for continuous discourse, while that for unfamiliar words was decreased. However, when the lowpass cutoff frequency dropped below 2000 Hz, intelligibility for even familiar words was severely affected, while sentence intelligibility remained high. Speaks and Jerger (1965) manipulated the conditional probabilities between words in sentences and demonstrated that intelligibility increased with increasing approximation to normal English sentences under two conditions of lowpass filtering (350 and 500 Hz). These studies demonstrate that higher level information has a great influence upon the ability to recognize (or identify) lower level components of speech. Although intelligibility measures based upon word lists are useful for designing and evaluating communication systems by virtue of their sensitivity to interference, they provide a highly inaccurate measure of the ability to identify acoustically degraded words when they are used, as they normally are, for communication purposes as components of sentences.

There have been only a few previous reports of the intelligibility of bandpassed sentences, and in all but one, the sentences were degraded further by the addition of noise. Kryter (1960), as part of a larger study dealing with the intelligibility of monosyllabic word lists under a variety of conditions, described the intelligibility scores of 6 listeners hearing 10 bandpassed sentences. However, only two filtering conditions employing single passbands were used—one band was 500–2000 Hz and the other was 1000–2500 Hz. Kryter presented noise along with the sentences, and despite the broad bandwidths of the speech, this additional interference apparently reduced intelligibility well below that found for sentences in the present study. Grant and Braida (1991), in a study of interaction of visual and auditory cues in speech reading (lipreading), used three filtering conditions employing single 1/3-octave bandpassed sentences. They too introduced noise to degrade the intelligibility of the speech bands and reported intelligibility values for auditory information alone that were well below those found for 1/3-octave bands in the same frequency range in the present Experiment 1. The only previous report of the intelligibility of bandpassed sentences played without the addition of noise appears to be that of Bashford et al. (1992). It was noted that one of their control conditions involving sentence lists with highpass and lowpass slopes (48 dB/octave) intersecting at 1500 Hz produced an intelligibility score that was over 95%. Since a literature search revealed no previous reports of such high intelligibility based upon such a narrow band of frequencies, and since this finding had relevance for models and theories of speech perception, the present investigation was undertaken to extend the study of narrowband sentences to other frequencies and conditions.

There is ample evidence that accuracy in identifying phonetic, syllabic, and lexical components becomes much more robust as speech more closely approximates normal discourse. The remarkable ability of sentences to maintain intelligibility despite distortion and interference is consistent with the LAME (lateral access from multilevel engrams) model of speech perception (Warren, 1981). The LAME model is a model of caution that considers that linguistic processing by a listener having mastery of a language is flexible and opportunistic and can involve reciprocal interaction of top-down and bottom-up information, as well as the pooling of information occurring at comparable levels of linguistic complexity, in keeping with both the nature of the signal and the task demands. Thus, the processing strategies employed by listeners vary, and they are matched not only to the extent and level of the linguistic information, but also to the available spectral components. The presence of formants and other spectral features necessary for the identification of words when presented as members of word lists may not be required when the same lexical items are used in discourse. It follows that rules and models applicable to lexical access when words are presented in lists or other special modes may not be applicable when the same words are used for communication in sentences and discourse.

The high intelligibility of spectrally limited sentences in the present study reflects a skill that can be gainfully employed in everyday listening when extraneous sounds mask some spectral regions but not others. The existence of such a skill is also suggested by a phenomenon in which listeners themselves restrict the spectral information used for verbal organization of broadband speech sounds. When a series of several steady-state vowels with durations corresponding to those occurring in speech are repeated loudly and clearly, the individual phonemes lose their identity, and illusory syllables and words are heard (Warren, Bashford, & Gardner, 1990). Typically, a perceptual splitting of the broadband vowels takes place, and listeners report two simultaneous forms that differ in both timbre and phonemic content. A point of special relevance for the present study is that Chalikia and Warren (1991, 1994) found that the splitting of the broadband vowels occurred along spectral fault lines, with the division occurring at about 1500 Hz; frequencies below this value were used for one verbal organization and those above for the other. This fissioning took place despite the harmonically related frequencies within each vowel and despite the phase relations across the spectrum, indicating that there was a single pulsed source for that vowel. The verbal forms heard for each of the two spectral regions followed the phonotactic rules of English; further, they almost always consisted of either English words or of syllables found in English words. Thus, when it is difficult to match a sequence of steady-state vowels to an utterance that could be produced by a speaker, a spontaneous division into two spectral regions takes place, and each region is processed separately and used for accessing a preexisting English syllabary or lexicon. Of course, with normal utterances, the information present in different frequency regions corresponds to a single verbal signal, and the end product of processing is a fused, not a split percept. However, the ability to carry out simultaneous linguistic organization of separate spectral regions can provide a backup and verification procedure when speech comprehension is difficult for intrinsic reasons (e.g., unusual pronunciation and inappropriate or unfamiliar usage of words).[2]

There has been no previous systematic study of the ability to comprehend sentences degraded solely by bandpass filtering. The present study has shown that sentences restricted to narrow spectral slits maintain a remarkably high intelligibility over an extended range of center frequencies, and that information contained in widely separated bands can be integrated to produce an increase in intelligibility that is much greater than simple additivity. These observations, together with other evidence, indicate that listeners possess processing mechanisms and strategies employing limited spectral regions that can enhance comprehension under difficult listening conditions.

REFERENCES

ANSI-S3.5, 1969 (R1986). *American national standard methods for the calculation of the articulation index*. New York: American National Standards Institute.

Bashford, J. A., Jr., Riener, K. R., & Warren, R. M. (1992). Increasing the intelligibility of speech through multiple phonemic restorations. *Perception & Psychophysics*, **51**, 211-217.

Bashford, J. A., Jr., & Warren, R. M. (1987). Effects of spectral alternation on the intelligibility of words and sentences. *Perception & Psychophysics*, **42**, 431-438.

Bond, Z. S. (1976). Identification of vowels excerpted from neutral and nasal contexts. *Journal of the Acoustical Society of America*, **59**, 1229-1232.

Braida, L. D., Durlach, N. I., Lippman, R. P., Hicks, B. L., Rabinowitz, W. M., & Reed, C. M. (1979). *Hearing aids: A review of past research on linear amplification, amplitude compression, and frequency lowering* (ASHA Monographs No. 19). Rockville, MD: American Speech-Language-Hearing Association.

Chalikia, M. H., & Warren, R. M. (1991). Phonemic transformations: Mapping the illusory organization of steady-state vowel sequences. *Language & Speech*, **34**, 109-143.

Chalikia, M. H., & Warren, R. M. (1994). Spectral fissioning in phonemic transformations. *Perception & Psychophysics*, **55**, 218-226.

Egan, J. P., & Wiener, F. M. (1946). On the intelligibility of bands of speech in noise. *Journal of the Acoustical Society of America*, **18**, 435-441.

Epstein, A., Giolas, T. G., & Owens, E. (1968). Familiarity and intelligibility of monosyllabic word lists. *Journal of Speech & Hearing Research*, **11**, 435-438.

French, N. R., & Steinberg, J. C. (1947). Factors governing the intelligibility of speech sounds. *Journal of the Acoustical Society of America*, **19**, 90-119.

Grant, R. W., & Braida, L. D. (1991). Evaluating the articulation index for auditory-visual input. *Journal of the Acoustical Society of America*, **89**, 2952-2960.

Kalikow, 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.

Kryter, K. D. (1960). Speech bandwidth compression through spectrum selection. *Journal of the Acoustical Society of America*, **32**, 547-556.

Kryter, K. D. (1962). Validation of the articulation index. *Journal of the Acoustical Society of America*, **34**, 1698-1702.

Morton, J., & Carpenter, A. (1962). Judgment of the vowel colour of natural and artificial sounds. *Language & Speech*, **5**, 190-204.

Owens, E. (1961). Intelligibility of words varying in familiarity. *Journal of Speech & Hearing Research*, **4**, 113-129.

Pavlovic, C. V. (1994). Band importance functions for audiological applications. *Ear & Hearing*, **15**, 100-104.

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.

Speaks, C., & Jerger, J. (1965). Method for measurement of speech identification. *Journal of Speech & Hearing Research*, **8**, 185-194.

Studebaker, G. A., Pavlovic, C. V., & Sherbecoe, R. L. (1987). A frequency importance function for continuous discourse. *Journal of the Acoustical Society of America*, **81**, 1130-1138.

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. (1984). Perceptual restoration of obliterated sounds. *Psychological Bulletin*, **96**, 371-383.

Warren, R.M. (in press). Auditory illusions and perceptual processing of speech. In N. Lass (Ed.), *Principles of experimental phonetics*. Saint Louis: Mosby.

Warren, R.M., Bashford, J.A., Jr., & Gardner, D.A. (1990). Tweaking the lexicon: Organization of vowel sequences into words. *Perception & Psychophysics*, **47**, 423-432.

NOTES

1. Usable information may be available some distance from the center frequency (perhaps as far as the 30-dB down-points, as suggested by Kryter, 1960), despite the attenuation and distortion of the spectral profile by the steep filter slopes.

2. The CID "everyday American speech" sentences used in the present study have an intelligibility of over 98% when either lowpassed at 1100 Hz or highpassed at 1700 Hz using steep filter slopes of 96 dB /octave (Bashford & Warren, 1987). This spectral redundancy of broadband speech not only provides listeners with the possibility of using dual processing mechanisms for speech recognition, but suggests that artificial speech recognition systems might benefit from separate processing of high- and low-frequency regions followed by comparison and melding.

(Manuscript received March 7, 1994;
revision accepted for publication August 22, 1994.)