Spectral integration of speech bands in normal-hearing and hearing-impaired listeners

Joseph W. Hall III,^{a)} Emily Buss, and John H. Grose

Department of Otolaryngology/Head and Neck Surgery, University of North Carolina School of Medicine, Chapel Hill, North Carolina 27599

(Received 21 December 2007; revised 12 May 2008; accepted 13 May 2008)

This investigation examined whether listeners with mild–moderate sensorineural hearing impairment have a deficit in the ability to integrate synchronous spectral information in the perception of speech. In stage 1, the bandwidth of filtered speech centered either on 500 or 2500 Hz was varied adaptively to determine the width required for approximately 15%–25% correct recognition. In stage 2, these criterion bandwidths were presented simultaneously and percent correct performance was determined in fixed block trials. Experiment 1 tested normal-hearing listeners in quiet and in masking noise. The main findings were (1) there was no correlation between the criterion bandwidths at 500 and 2500 Hz; (2) listeners achieved a high percent correct in stage 2 (approximately 80%); and (3) performance in quiet and noise was similar. Experiment 2 tested listeners with mild–moderate sensorineural hearing impairment. The main findings were (1) the impaired listeners showed high variability in stage 1, with some listeners requiring narrower and others requiring wider bandwidths than normal, and (2) hearing-impaired listeners achieved percent correct performance in stage 2 that was comparable to normal. The results indicate that listeners with mild–moderate sensorineural hearing loss do not have an essential deficit in the ability to integrate across-frequency speech information. © 2008 Acoustical Society of America.

[DOI: 10.1121/1.2940582]

PACS number(s): 43.66.Dc, 43.66.Sr, 43.71.Ky [BCM]

Pages: 1105–1115

I. INTRODUCTION

The present study investigated speech recognition associated with a single relatively narrow band of speech information or two relatively narrow bands that were widely separated in frequency. In normal-hearing listeners, narrow bands of speech from widely spaced spectral regions can combine to produce a percent correct well above the sum of the percent correct values associated with each band separately (e.g., Grant and Braida, 1991; Warren *et al.*, 1995; Lippmann, 1996; Kasturi *et al.*, 2002). For example, whereas a low- and a high-frequency region may each support less than 20% correct speech identification, both bands presented together may support 70% identification (e.g., Grant and Braida, 1991). Although such effects are not predicted adequately by the articulation theory (e.g., French and Steinberg, 1947), they should not be regarded as surprising. Grant and Braida commented that the substantial improvement in performance achieved when spectrally separated bands are presented together is due to the fact that the bands are likely to contain nonredundant, complementary information. One way to conceptualize this is in terms of the possible set of meaningful utterances that are compatible with the bandlimited information presented to the listener. Each band presented alone might be consistent with a large set of distinct utterances. However, when the bands are presented together the combined information is consistent with a reduced set of candidate utterances.

The ability to perceive speech on the basis of sparse cues that are separated in frequency could have importance for speech understanding in noisy backgrounds. For example, when the signal to noise ratio is very low, a listener may not have access to the entire spectrum of a speech target and good performance may depend upon the ability to integrate speech fragments that are separated in frequency (e.g., Miller and Licklider, 1950; Howard-Jones and Rosen, 1993; Assmann and Summerfield, 2004; Buss *et al.*, 2004; Cooke, 2006; Hall *et al.*, 2008). This may be particularly significant from the perspective of hearing impairment because some evidence indicates that listeners with sensorineural hearing loss may have a diminished ability to integrate synchronous, frequency-distributed information in the perception of speech (Turner *et al.*, 1995; Turner *et al.*, 1999; Healy and Bacon, 2002). Such evidence has arisen from studies using “vocoder” paradigms where speech is divided into a number of frequency bands and the envelope of each band is used to modulate a carrier stimulus in the same frequency region as the original speech, with either a tone or noise band serving as the carrier (e.g., Shannon *et al.*, 1995; Turner *et al.*, 1995; Dorman *et al.*, 1997; Turner *et al.*, 1999). In one such study, it was found that although hearing-impaired listeners could perform as well as normal for consonant perception based upon a single band containing temporal envelope information, the performance of the impaired listeners was worse than normal for two or more bands carrying quasi-independent temporal envelope information (Turner *et al.*, 1999). The normal performance for the single band case was interpreted as being consistent with temporal modulation transfer function studies (Viemeister, 1979) that have gener-

^{a)}Electronic mail: jwh@med.unc.edu

ally indicated a normal ability of hearing-impaired listeners to code the temporal envelope of a stimulus, provided that effects of stimulus audibility are taken into account (e.g., Bacon and Gleitman, 1992; Moore *et al.*, 1992). Turner *et al.* (1999) and Healy and Bacon (2002) have suggested that the speech results related to the combination of information from more than one spectral band containing temporal information may indicate a reduced ability of hearing-impaired listeners to combine spectrotemporal information across frequency.

One question of importance is whether the deficits in across-frequency integration suggested by the above-mentioned studies occur exclusively for the type of speech material they used (only temporal envelope information), or whether it also occurs for more conventionally filtered speech. A recent study by Grant *et al.* (2007) is pertinent to this question. That study examined the intelligibility of speech filtered into relatively narrow spectral bands for both normal-hearing listeners and listeners with sensorineural hearing impairment. They examined several conditions involving either audio only or audio-visual cues. The conditions most relevant to the present study involved either a relatively low-frequency band alone (298–375 Hz) or the low-frequency band plus a high-frequency band (4762–6000 Hz). Both normal-hearing and hearing-impaired listeners achieved approximately 20% correct performance for the low band alone, but whereas the normal-hearing listeners improved to approximately 60% for both low and high bands presented together, the hearing-impaired listeners improved to only about 40% correct with both bands present. Grant *et al.* noted several possible accounts for the poorer performance of the hearing-impaired listeners when both bands were present, including an essential deficit in the ability to integrate across-frequency information, relatively great hearing loss in the region of the high-frequency band, and poor processing of the high-frequency information due to upward spread of masking, a manifestation of poor frequency selectivity that is relatively common in listeners with sensorineural hearing loss (e.g., Tyler *et al.*, 1984; Gagné, 1988). One feature of the Grant *et al.* study that makes it somewhat difficult to compare to the previous vocoder-based studies is that although the Grant *et al.* study examined the impact of adding a high-frequency speech band to a low-frequency band, the study did not specifically determine the intelligibility associated with the high band presented alone.

In the present study, speech recognition was assessed for low and high bands presented alone, and for the bands presented together. Furthermore, stimulus features intended to minimize effects related to upward spread of masking were employed. The rationale was to use an approach that allowed a test of whether listeners with sensorineural hearing impairment have an essential deficit in the ability to integrate across-frequency speech information apart from factors that could be related to a peripherally based reduction in frequency selectivity. Two experiments were performed. The first experiment tested listeners with normal hearing, examining the intelligibility of speech presented in quiet and speech presented in a masking noise background that was intended to simulate a 40–50 dB hearing loss. The main purpose of this study was to obtain information about the ro-

bustness of across-frequency speech integration under unmasked and masked conditions and to provide baselines against which to compare data from hearing-impaired listeners tested in the second experiment.

II. EXPERIMENT 1: NORMAL-HEARING LISTENERS

A. Methods

1. Listeners

Two sets of listeners with normal hearing participated, one tested in quiet (four males and seven females) and the other tested in background noise (three males and five females). The mean age of the normal-hearing listeners was 35.2 years with a standard deviation of 12.3 years. All listeners were screened to have audiometric thresholds of 20 dB HL or better at octave frequencies from 250 to 8000 Hz.

2. Rationale and stimuli

The speech material consisted of Bamford–Kowal–Bench sentences (Bench *et al.*, 1979), with each sentence containing from three to five key words. There were 21 lists of 16 sentences each. This corpus of sentences allowed testing to be completed without replicating any sentence for all of the listeners tested in quiet and all but two of the listeners tested in noise. For these two listeners, parts of list 1 and list 2 were repeated. In some conditions, speech was filtered into a single band centered at one of two frequencies, and in other conditions bands were available at both center frequencies simultaneously. The two bands were arithmetically centered on 500 and 2500 Hz, and filtering was performed via convolution with a 12 Hz resolution.

An important part of the rationale underlying the methods was related to upward spread of masking. This rationale was not of direct interest with regard to the normal-hearing listeners but is important from the standpoint of hearing impairment. Many hearing-impaired listeners are prone to greater than normal upward spread of masking (e.g., Tyler *et al.*, 1984; Gagné, 1988), and we wished to minimize the possibility that such masking could underlie differences between the results of the normal-hearing and hearing-impaired listeners. This objective was met with two types of stimulus manipulation. One manipulation was a level boost of the high-frequency speech band relative to the low-frequency band. Dubno *et al.* (2006) used a similar method in a previous study of speech perception in hearing-impaired listeners. The speech level was 85 dB SPL prior to bandpass filtering for the low band and 97 dB SPL prior to bandpass filtering for the high band. The other stimulus manipulation involved a condition where the low and high bands were presented to opposite ears (details in the following), preventing peripheral masking of the high band by the low band.

In the conditions meant to simulate hearing loss, a pink noise was presented at a level of approximately 37 dB/Hz SPL at 1 kHz. This noise resulted in masked thresholds of approximately 50–55 dB SPL at octave frequencies from 500 to 4000 Hz.

FIG. 1. Criterion normalized bandwidth for the high band vs criterion normalized bandwidth of the low band, with the correlation (r) shown in the box. Data are for normal-hearing listeners. The left-hand panels show data for ear 1 stimulation, and the right-hand panels show data for stimulation where the low band was presented to ear 1 and the high band was presented to ear 2. The upper and lower panels show data for listeners tested in quiet and noise, respectively.

3. Procedure

The listener sat in a double-walled sound booth and was instructed to repeat as many words as possible after each sentence was presented and to guess for words that were not intelligible. No feedback was provided. The experimenter was positioned in front of a visual display that showed the current sentence and monitored the listener's response via a talk-back loop. The experimenter recorded errors following each listener response. Stimuli were presented over Sennheiser HD 265 earphones.

A critical consideration was the specific bandwidth to which the speech was filtered at each center frequency, as the goal was to have the speech intelligibility fall within a relatively narrow range of poor performance (i.e., 15%–25% correct) in the single-band conditions. Whereas most previous approaches have used fixed bandwidths, the present approach instead allowed bandwidth to vary across listeners (Noordhoek *et al.*, 1999), partly because of the potential for intersubject variation in speech performance for a fixed bandwidth among hearing-impaired listeners. Another important reason for the adaptive approach in the first stage of testing was that it was desirable to home in on an appropriate speech bandwidth rapidly. Because of travel and other con-

siderations, many of the hearing-impaired listeners were available for a relatively limited testing time and it was therefore critical to use efficient testing strategies.

The adaptive procedure was carried out separately for a band centered on 500 Hz, and for another band centered on 2500 Hz. At each center frequency the bandwidth was changed adaptively (by a factor of 1.21), with the bandwidth increasing following two sentences in a row where no key word was reported correctly, and with bandwidth decreasing following a sentence in which any key word was correctly reported. The run was stopped following eight reversals in bandwidth adjustment, and the threshold bandwidth was taken as the geometric mean of the bandwidths at the last six reversals. This threshold value will be referred to as the criterion speech bandwidth. Testing conducted prior to this study on ten normal-hearing listeners showed that the criterion speech bandwidth estimated from this stepping rule was associated with approximately 15%–25% correct (mean of 23.2% correct and standard deviation of 4.3% for the low band and mean of 20.3% correct and standard deviation of 4.7% for the high band) when listeners were retested with the bandwidth fixed at this criterion value.¹ There were three conditions evaluated in this stage of testing: (1) the low band

FIG. 2. Speech recognition percent correct vs total normalized bandwidth, with the correlation (r) shown in the box. Data are for normal-hearing listeners. The left-hand panels show results for monaural stimulation, and the right-hand panels show results for dichotic stimulation. The upper and lower panels show data for listeners tested in quiet and noise, respectively.

presented to ear 1; (2) the high band presented to ear 1; and (3) the high band presented to ear 2. For each of the normal-hearing listeners, ear 1 was randomly selected to be either the right or the left ear.

In the second stage of testing, fixed block trials were used to determine the percent correct speech identification obtained when the low and high bands were presented together. Each estimate of percent correct was determined using a single BKB list consisting of 16 sentences (one list), and three estimates were obtained in each condition. In this stage of testing, the bandwidth at each frequency was held constant at the criterion speech bandwidths that had been obtained in stage 1. There were two conditions in this stage of testing, each involving the simultaneous presentation of the low band and the high band: (1) both the low and high bands presented to ear 1 and (2) the low band presented to ear 1 and the high band presented to ear 2. The final estimate of percent correct in each condition was computed by applying an arcsine transformation (Studebaker, 1985) to the three replicate data points, averaging these values, and then converting the result back to percent correct.

B. Results and discussion

Figure 1 shows the relationships between the bandwidths estimated for the normal-hearing listeners at 500 and 2500 Hz for speech presented in both quiet and in noise. Data are expressed as the criterion speech bandwidth divided by the center frequency, a value referred to as the criterion normalized bandwidth. Panels on the left-hand side are associated with conditions where each band was presented to ear 1, and panels on the right-hand side are associated with conditions where the low band was presented to ear 1 and the high band was presented to ear 2. One finding apparent in Fig. 1 is that there was no obvious relation between the cri-

terion normalized bandwidths at each center frequency: that is, listeners requiring a relatively wide criterion normalized bandwidth at 500 Hz did not necessarily require a wide bandwidth at 2500 Hz. The criterion normalized bandwidth ranged from 0.27 to 0.57 at 500 Hz, and from 0.22 to 0.48 at 2500 Hz. The criterion normalized bandwidth was relatively similar for the listeners tested in quiet and those tested in noise. This finding is consistent with an analysis that indicated that the signal-to-noise ratio was approximately 14 dB or greater in the presence of the pink noise masker, suggesting that speech audibility was not a limiting factor. A repeated measures analysis of variance was performed to examine this, with a within subjects factor of speech band condition (low band presented to ear 1, high band presented to ear 1, and high band presented to ear 2) and a between subjects factor of masker (quiet versus masking noise). In this analysis, the dependent variable was the criterion normalized bandwidth at each frequency. The analysis indicated a significant effect of speech band condition ($F_{2,34}=9.6$; $p=0.001$), but no effect of masking noise ($F_{2,34}=0.4$; $p=0.53$) and no interaction ($F_{2,34}=1.3$; $p=0.28$). Post hoc testing revealed that the significant effect of speech band was due to the criterion normalized bandwidth being slightly wider for the low band than for the high band.

Figure 2 plots the percent correct performance for the monaural and dichotic conditions where the low and high bands were presented simultaneously against a measure of the total speech bandwidth available in these conditions. The total speech bandwidth was defined as the sum of the criterion normalized bandwidths at 500 and 2500 Hz. For the monaural condition, the criterion normalized bandwidth for 2500 Hz was associated with ear 1, and for the dichotic condition, the criterion normalized bandwidth for 2500 Hz was associated with ear 2. It was of interest to plot the monaural

TABLE I. Correlation between percent correct obtained with the low- and high-frequency bands presented together either monaurally or dichotically and the criterion normalized bandwidth of the low band (low), the criterion normalized bandwidth of the high band (high), or the sum of these normalized bandwidths (both). Correlations that are significant at the 0.05 and 0.01 levels of probability are noted by an asterisk or double asterisk, respectively.

	Normal listeners								
	Quiet			Masking noise			Hearing-impaired listeners		
	Low	High	Both	Low	High	Both	Low	High	Both
Monaural	0.65*	0.35	0.66*	0.87**	0.16	0.81*	0.09	-0.34	0.53
Dichotic	0.45	0.51	0.68*	0.59	0.20	0.85**	0.23	-0.49	0.78*

and dichotic percent correct performance against these bandwidth metrics because of the possibility that listeners having relatively wide criterion bandwidths might perform relatively well in the case where the bands were presented simultaneously. This did indeed appear to be the case, with a trend for higher percent correct performance in cases where the total speech bandwidth was relatively wide (see Fig. 2). One interpretation of this effect is that some listeners are relatively poor in extracting information from a single band (and therefore require a large bandwidth for criterion performance), but that such listeners are better able to use the information when two bands are presented together. The columns of Table I dealing with normal-hearing listeners (left and middle) show associated correlations for the low and high bands alone and for these bands together. As can be seen, for the present listeners, there was some indication that the bandwidth of the low-frequency band may have contributed relatively more to the performance obtained when both bands were present, particularly for monaural presentation.

Another finding apparent in Fig. 2 is that the performance was relatively good regardless of whether stimulation was monaural or dichotic, or whether listeners were tested in quiet or in noise. The level of performance obtained across these conditions (approximately 64%–94% correct) was clearly greater than that obtained by additive combination of information present in the single-band conditions (which were associated with approximately 15%–25% correct). This result is consistent with previous demonstrations of speech band combination effects for frequency-separated bands (e.g., Grant and Braida, 1991; Warren *et al.*, 1995; Lippmann, 1996; Kasturi *et al.*, 2002). A repeated measures analysis of variance was performed with a within-subjects factor of mode of presentation (monaural versus dichotic presentation) and a between-subjects factor of masker (quiet versus masking noise). In this analysis, the dependent variable was the arcsine transformation of percent correct word identification. The analysis indicated no significant effect of presentation mode ($F_{1,17}=0.6$; $p=0.45$), no effect of masking noise ($F_{1,17}=1.9$; $p=0.18$), and no interaction ($F_{1,17}=2.25$; $p=0.15$). The fact that performance was relatively good in conditions where a background masking noise was present indicates that the speech band combination effect is relatively robust in listeners with normal hearing. This result suggests that for hearing-impaired listeners having thresholds similar to those simulated by the masking noise used here, little effect of audibility may be expected.

III. EXPERIMENT 2: HEARING-IMPAIRED LISTENERS

A. Methods

1. Listeners

There were nine hearing-impaired listeners, six females and three males. The listeners had an average age of 43.3 years with a standard deviation of 11.3 years. All listeners had mild-to-moderate sensorineural hearing losses as determined via air- and bone-conduction audiometry. Audiometric data and speech recognition scores (percent correct) for monosyllabic words presented in quiet for these listeners are shown in Table II. The assignment of ear 1 and ear 2 was random except in two cases (listeners 7 and 8). During audiometric testing, the responses of listener 7 to both speech and tones were relatively unreliable for right-ear presentation. This listener was therefore tested using the left ear only. Listener 8 had normal hearing in the right ear and so was tested using the left ear only. For this listener, a 40 dB HL speech-shaped noise was presented to the right ear during filtered speech testing in order to mask speech that may have crossed over from the left earphone.

TABLE II. Air-conduction audiograms (dB HL) and speech recognition scores (% correct) for monosyllabic words. Ear 1 (the ear tested with both the low and high bands presented the same ear) is identified in bold. "NT" indicates not tested.

		250	500	1000	2000	4000	8000	Recognition
HI1	L	20	45	50	50	30	35	88
	R	20	40	50	45	35	35	92
HI2	L	30	40	45	45	50	50	84
	R	30	35	45	50	45	50	84
HI3	L	55	70	75	50	65	75	84
	R	45	55	55	50	60	70	76
HI4	L	25	25	50	50	45	40	88
	R	25	25	45	45	45	35	88
HI5	L	35	35	25	30	40	60	100
	R	25	30	30	30	45	65	100
HI6	L	35	40	40	40	50	60	92
	R	30	35	40	45	45	60	92
HI7	L	45	50	50	45	50	65	80
	R	60	70	65	50	50	95	NT
HI8	L	50	50	50	50	50	80	76
	R	15	5	15	10	10	15	100
HI9	L	50	50	45	40	40	40	84
	R	60	60	65	60	65	75	52

FIG. 3. Criterion normalized bandwidth for the high band vs criterion normalized bandwidth for the low band, with the correlation (r) shown in the box. Data for hearing-impaired listeners are depicted by listener number. The left-hand panels show data for ear 1 stimulation, and the right-hand panels show data for stimulation where the low band was presented to ear 1 and the high band was presented to ear 2. The upper panels allow comparison of the hearing-impaired listeners to the normal-hearing listeners who were tested in quiet, and the lower panels allow comparison of the hearing-impaired listeners to the normal-hearing listeners who were tested in noise (normal data are replotted from Fig. 1). Hearing-impaired listeners having criterion bandwidths outside the range of normal-hearing listeners are identified by an asterisk: Asterisks above and below the symbol signify high-band criterion bandwidths above and below the limits for normal-hearing listeners, respectively; asterisks to the right and to the left of the symbol signify low-band criterion bandwidths above and below the limits for normal-hearing listeners, respectively.

2. Methods and procedure

The methods and procedure were identical to those used in experiment 1, except that all speech was presented in quiet. Audibility of speech was judged at the outset as unlikely to limit performance of hearing-impaired listeners. As indicated in Table II, pure-tone thresholds at frequencies near the center of each speech band were 55 dB HL or better for all hearing-impaired listeners, similar to those of normal-hearing listeners tested in the presence of pink noise, as noted earlier. At threshold bandwidths estimated for noise-masked normal-hearing listeners, the signal-to-noise ratios at threshold were approximately 14 dB or greater, suggesting that speech was likely to be audible for the hearing-impaired listeners under these conditions.

B. Results and discussion

Figure 3 plots the criterion normalized bandwidth for the 500 Hz center frequency against the criterion normalized bandwidth for the 2500 Hz center frequency. The data for the

normal-hearing listeners are replotted to aid comparison to the data of the hearing-impaired listeners, for whom data are identified by listener number. The upper panels are used to compare with normal-hearing listeners tested in quiet and the lower panels are used to compare with normal-hearing listeners tested in noise. As with the normal-hearing listeners, there was no apparent relation between the criterion normalized bandwidths at the two center frequencies. There were relatively large individual differences in the bandwidth necessary for criterion performance in the hearing-impaired listeners, with criterion normalized bandwidth ranging from approximately 0.28 to 1.06 Hz at 500 Hz, and from approximately 0.14 to 0.54 Hz at 2500 Hz. The criterion speech bandwidths obtained for the hearing-impaired listeners were broadly similar to those obtained by the normal-hearing listeners. Statistical analyses were performed to compare performance of the two groups. Because some of the hearing-impaired listeners were not tested in ear 2, repeated measures analyses of variance were performed to compare the normal-hearing and hearing-impaired listeners for the

two ear 1 conditions (low band and high band) and separate t-tests were performed to compare the normal-hearing and hearing-impaired listeners for the single ear 2 condition (high band). In all analyses, the dependent measure was the criterion normalized bandwidth.

Comparisons of the hearing-impaired listeners to the normal-hearing listeners tested in quiet will be considered first. The repeated measures analysis of variance had a within-subjects factor of condition (low band presented to ear 1 and high band presented to ear 1), and a between subjects factor of hearing loss (present or absent). The analysis indicated no significant effect of condition ($F_{1,18}=1.7$; $p=0.21$) or of hearing impairment ($F_{1,18}=0.15$; $p=0.70$) and no interaction ($F_{1,18}=0.55$; $p=0.47$). The t-test comparing the normal-hearing and hearing-impaired listeners for the high band presented to ear 2 also indicated no significant difference ($t_{16}=0.73$; $p=0.48$). The repeated measures analysis comparing the masked normal-hearing listeners to the hearing-impaired listeners indicated no significant effect of condition ($F_{1,15}=2.49$; $p=0.13$) or of hearing impairment ($F_{1,15}=0.16$; $p=0.69$) and no interaction ($F_{1,15}=0.04$; $p=0.84$). The t-test comparing the masked normal-hearing and the hearing-impaired listeners for the high band presented to ear 2 also indicated no significant difference ($t_{13}=0.23$; $p=0.82$).

At first glance, the result that the relative criterion bandwidth did not differ between normal and hearing-impaired listeners might appear to be in conflict with the results obtained by Noordhoek *et al.* (2000), where hearing-impaired listeners needed a wider bandwidth than normal to obtain 50% correct. There are at least three factors that should be considered in this regard:

- (1) The listeners in the Noordhoek *et al.* study were tested using bandlimited speech presented in a complementary band-stop masking noise. It is possible that relatively poor frequency selectivity, common in sensorineural hearing loss (e.g., Tyler *et al.*, 1984; Stelmachowicz *et al.*, 1985; Leek and Summers, 1996), resulted in a greater masking effect for hearing-impaired than normal-hearing listeners, perhaps accounting for the need for a wider speech bandwidth in the hearing-impaired listeners of that study. The present finding that normal-hearing and hearing-impaired listeners required broadly similar speech bandwidth for criterion performance is consistent with the previous results of Grant *et al.* (2007), which indicated that, for the same narrow speech bandwidth, normal-hearing and hearing-impaired listeners obtained approximately the same, relatively low percent correct.
- (2) The Noordhoek *et al.* study tracked 50% intelligibility and therefore required a larger bandwidth than the present study where a lower intelligibility was tracked. It is possible that the wider bandwidth tracked in the Noordhoek *et al.* study resulted in effects of hearing impairment related either to within-band masking effects or to a smaller than normal number of effectively independent frequency channels at the speech bandwidth associated with normal performance.
- (3) Because the variability was relatively great among the

hearing-impaired listeners in the present study, it is important to consider whether some of the results of the hearing-impaired listeners fell outside the normal range even though there was no overall group difference. This was addressed by assessing whether individual hearing-impaired listeners had criterion normalized bandwidths that were more than 2 s.d. above or below the normal mean. In Fig. 3, listeners having bandwidths above the normal limit are identified by an asterisk above the listener number for the high band and to the right of the listener number for the low band; listeners having bandwidths below the normal limit are identified by an asterisk below the listener number for the high band. No hearing-impaired listener had a bandwidth below the normal limit for the low band. With respect to the group of normal-hearing listeners tested in quiet, this analysis indicated the following. For the low band, none of the hearing-impaired listeners had criterion bandwidths narrower than the normal limit, and listeners 8 and 9 had bandwidths wider than the normal limit; for the high band presented to ear 1, listeners 3, 4, 7, and 8 had criterion bandwidths narrower than the normal limit, and listeners 1, 2, and 9 had criterion bandwidths wider than the normal limit. For the high band presented to ear 2, listeners 3 and 4 had criterion bandwidths narrower than the normal limit, and none of the listeners had criterion bandwidths wider than the normal limit. With respect to the group of normal-hearing listeners tested in masking noise, this analysis indicated the following. For the low band, none of the hearing-impaired listeners had criterion bandwidths narrower than the normal limit, and listener 8 had a bandwidth wider than the normal limit; for the high band presented to ear 1, listeners 3, 4, and 7 had criterion bandwidths narrower than the normal limit, and listeners 1, 2, and 9 had criterion bandwidths wider than the normal limit. For the high band presented to ear 2, none of the hearing-impaired listeners had criterion bandwidths wider or narrower than the normal limit. Overall, the results indicate that the criterion speech bandwidth was variable in the hearing-impaired group, with some listeners requiring narrower than normal values (for the high band) and others requiring wider than normal values (for the low and high bands).

Figure 4 plots the percent correct performance for the monaural and dichotic conditions where the low and high bands were presented simultaneously against the measure of the total normalized bandwidth available in these conditions. Figure 4 again replots the normal data in order to aid comparison. The right-most portion of Table I shows associated correlations for the low and high bands alone and for these bands together. As with the normal-hearing listeners, there was a trend for the hearing-impaired listeners with the largest total normalized bandwidth to have higher percent correct scores, although this was statistically significant only for the dichotic case.

The most notable finding was that the performance of the hearing-impaired listeners was generally quite good when both the low and high bands were present and was

FIG. 4. Speech recognition percent correct vs total normalized bandwidth, with the correlation (r) shown in the box. Data for hearing-impaired listeners are depicted by listener number. The left-hand panels show data for monaural stimulation, and the right-hand panels show data for dichotic stimulation. The upper panels allow comparison of the hearing-impaired listeners to the normal-hearing listeners who were tested in quiet, and the lower panels allow comparison of the hearing-impaired listeners to the normal-hearing listeners who were tested in noise (normal data are replotted from Fig. 2). Hearing-impaired listeners having percent correct outside the range of normal-hearing listeners are identified by an asterisk: Asterisks above and below the symbol signify percent correct above and below the limits for normal-hearing listeners, respectively; asterisks to the right and to the left of the symbol signify total normalized bandwidths above and below the limits for normal-hearing listeners, respectively.

comparable to that for the normal-hearing listeners. t -tests were performed on the arcsine-transformed percent correct data to evaluate possible differences between the normal and hearing-impaired listeners. This testing indicated that the hearing-impaired listeners did not differ significantly from the normal-hearing listeners tested in quiet ($t_{18}=0.60$; $p=0.56$) or in noise ($t_{15}=1.49$; $p=0.16$). This finding also held for dichotic presentation both in quiet ($t_{16}=0.59$; $p=0.57$) and in noise ($t_{13}=0.91$; $p=0.38$). Because the performance of hearing-impaired listeners is often marked by high variability, it is also important to evaluate possible outliers within the impaired group. This was again assessed by determining whether hearing-impaired listeners fell more than 2 s.d. above or below the normal mean for either monaural or dichotic stimulation. Again, this was evaluated with respect to the arcsine-transformed percent correct data. The results were that none of the data of the hearing-impaired listeners fell outside the normal limit.

As noted in Sec. II, pilot data on ten normal-hearing listeners showed that the criterion speech bandwidth estimated from the adaptive testing was associated with approximately 15%–25% correct when listeners were retested in fixed blocks at this criterion bandwidth. Because no fixed-block testing was performed with hearing-impaired listeners for a single band at this criterion bandwidth, an additional analysis was done to evaluate the assumption that the initial, adaptive stage of testing converged on about the same percent correct for the normal-hearing and hearing-impaired listeners. The analysis was based upon the data from the adaptive tracks of both the normal-hearing and hearing-impaired listeners for either the low band or high band presented to ear

1. In the analysis, the bandwidths visited in each adaptive track were binned into equal log steps and psychometric functions were estimated with a linear fit (proportion correct plotted against bandwidth). Because fits based upon individual, raw data were relatively poor, data within each group were combined and normalized to the mean criterion bandwidth for each group. For example, if the mean criterion bandwidth (computed on log transform data) was 300 Hz for a listener but was 200 Hz for the group, the bandwidths for that individual were multiplied by a factor of 2/3. This procedure clustered the functions of all listeners around the centroid of the group without affecting the individual function slopes. The results of this procedure are shown in Fig. 5, with the size of the symbol reflecting the number of points contributing to the associated estimate. The line fitted to these data was used to estimate the proportion of correct responses associated with the mean criterion speech bandwidths obtained in the adaptive tracks; these values of proportion correct were approximately 0.14–0.17 for both the normal-hearing and hearing-impaired listeners. This analysis therefore supports the assumption that the adaptive threshold testing converged upon approximately the same level of speech recognition performance for the two groups of listeners.

IV. GENERAL DISCUSSION

A. Integration of spectrally separated speech information

The central question evaluated in this study was whether listeners with mild–moderate sensorineural hearing impair-

FIG. 5. Proportion correct plotted as a function of speech bandwidth (see the text for details). Functions are fitted to data from the adaptive track stage of testing for hearing-impaired listeners and normal-hearing listeners in quiet and in noise.

ment have an essential deficit in the ability to integrate information from simultaneous, frequency-separated narrow bands of filtered speech. In normal-hearing listeners, narrow bands of speech from widely spaced spectral regions can combine to produce a percent correct well above the sum of the percent correct values associated with each band separately. The results of the present study indicated that the hearing-impaired listeners showed a similar ability to combine speech information from frequency-separated bands. This occurred both for monaural and for dichotic presentation.

Previous vocoder-based speech results have suggested that sensorineural hearing loss may be associated with relatively poor across-frequency integration of speech information, and one interpretation considered by Grant *et al.* (2007) in their filtered speech study was also based upon poor across-frequency integration by hearing-impaired listeners. The results of the present study are not consistent with an interpretation that sensorineural hearing loss is associated with an essential deficit in the ability to integrate across-frequency speech information. However, the present results are not necessarily in conflict with those of the previous vocoder studies or with the results of Grant *et al.* (2007). For example, it is possible that the different pattern of results in the present study and the previous vocoder studies is related to differences in stimuli. Such stimulus differences could include those related to processing (vocoding versus filtering) and to level (high-frequency speech energy in the present study was boosted in level in order to avoid effects related to upward spread of masking). The issue of the level of high-frequency speech energy is also relevant to comparisons between the present study and the study of Grant *et al.* Although Grant *et al.* noted that a deficit in the ability to integrate across-frequency speech information could have been the basis for their filtered speech results, they also noted that other factors could have been at work, including increased upward spread of masking in the impaired listeners. Because the present stimuli had features designed to mini-

mize effects of upward spread of masking, the present results should not be interpreted as being in conflict with those of Grant *et al.* (2007).

Although the present results on the effect of sensorineural hearing impairment on the ability to integrate across-frequency speech information were relatively straightforward, there are nevertheless reasons to interpret them with some caution. One reason is that the listeners of this study had mild-moderate hearing losses. It is possible that speech integration results would be different for listeners with more severe hearing loss. A potential difficulty in interpreting results obtained from listeners with severe hearing loss is that reduced speech perception abilities (even with the complete speech spectrum available) might put a ceiling on the magnitude of speech combination effects for frequency-separated bands. Another reason for caution in interpreting the present results is that the listeners of this study had relatively flat hearing loss configurations. It is possible that sloping hearing losses may be associated with different speech integration abilities.

B. Criterion speech bandwidths at 500 and 2500 Hz

Although the primary purpose of the present study was to examine the effect of hearing impairment on the ability to integrate speech information across frequency, the findings on the criterion speech bandwidth are also of interest. There was considerable variability among the normal-hearing listeners on the criterion bandwidth measure, and even greater variability among the hearing-impaired listeners. Perhaps most striking in this regard is that, for monaural presentation when both the low and high bands were present simultaneously, more hearing-impaired listeners fell outside the range for normal-hearing listeners (two listeners below and three above) than inside this range in terms of total normalized bandwidth (see Fig. 4). This result was not predicted and it is not readily accounted for. One possible interpretation of the wider bandwidth required by some hearing-impaired listeners is based on the fidelity with which speech information is encoded: if encoding of information is somehow impaired, then criterion performance would require more information via greater bandwidth. It is more challenging to account for the finding that some hearing-impaired listeners required a narrower than normal total speech bandwidth for criterion performance (a finding that appears to have been dominated by the bandwidth of the high band). One possibility is that some hearing-impaired listeners may adapt to the abnormal speech patterns available at the outputs of their relatively wide auditory filters. For example, whereas the temporal envelope of a speech stimulus at the output of a relatively wide auditory filter may be abnormal, it may nevertheless carry information that listeners can learn to use to differentiate among speech sounds. One specific possibility is that hearing-impaired listeners may learn to make use of relatively high-rate modulation cues related to the fundamental frequency, which may, in turn, provide information related to voicing and pitch (e.g., Arehart, 1994).

This study also yielded information about the relation between bandwidths required to support a criterion level of

performance at two separated frequency regions. While there was no experimental hypothesis about this relationship, one expectation that seems reasonable is that listeners who require a relatively narrow bandwidth at one region would also require a narrow bandwidth at another region. However, across the listeners tested here (normal-hearing listeners tested in quiet and in noise and hearing-impaired listeners) the correlation between the criterion speech bandwidths associated with the two frequency regions was consistently close to zero. This would imply that performance was not dominated by some general speech processing factor that applies across bandwidth in speech perception. Instead, it may point to the importance of processing factors that are specialized with respect to frequency region and vary independently across listeners.

V. CONCLUSIONS

- (1) Listeners with sensorineural hearing impairment showed an ability to combine information from frequency-separated bands of speech that was similar to that demonstrated by normal-hearing listeners. This occurred both for monaural and dichotic stimulation. These results are consistent with an interpretation that listeners with mild-moderate sensorineural hearing loss do not have an essential deficit in the ability to combine across-frequency speech information.
- (2) Neither normal-hearing nor hearing-impaired listeners showed a significant correlation between the criterion speech bandwidths at the two frequency regions examined here.
- (3) Speech recognition performance when both the low and high bands were presented simultaneously tended to be better for listeners having relatively wide criterion bandwidths. This was true for both normal-hearing and hearing-impaired listeners.
- (4) On average, listeners with sensorineural hearing impairment required criterion speech bandwidths that were similar to normal at center frequencies of 500 and 2500 Hz. However, there was relatively great variation in the criterion speech bandwidths of the hearing-impaired listeners. In some conditions, there were hearing-impaired listeners who had criterion speech bandwidths that were narrower than the normal limit and others who had criterion speech bandwidths that were wider than the normal limit.

ACKNOWLEDGMENTS

This work was supported by NIH NIDCD Grant No. R01 DC00418. Professor Brian Moore provided several very helpful comments. Two anonymous reviewers also offered valuable suggestions that improved the quality of this manuscript.

¹Although the percent correct converged upon in typical tracking procedures involving stimuli presented over independent trials can be calculated in a straightforward manner (Levitt, 1971), such a calculation is more complex in the current procedure, where a number of words are presented within a sentence and the probability of correctly identifying a particular key word may depend in part on the semantic context of the other words

in the sentence. For this reason, the percent correct associated with the adaptively measured criterion bandwidth was determined empirically.

- Arehart, K. H. (1994). "Effects of harmonic content on complex-tone fundamental-frequency discrimination in hearing-impaired listeners," *J. Acoust. Soc. Am.* **95**, 3574–3585.
- Assmann, P. F., and Summerfield, A. Q. (2004). "The perception of speech under adverse conditions," in *Speech Processing in the Auditory System*, edited by S. Greenberg, W. A. Ainsworth, A. N. Popper, and R. R. Fay (Springer, New York).
- Bacon, S. P., and Gleitman, R. M. (1992). "Modulation detection in subjects with relatively flat hearing losses," *J. Speech Hear. Res.* **35**, 642–653.
- Bench, J., Kowal, A., and Bamford, J. (1979). "The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children," *Br. J. Audiol.* **13**, 108–112.
- Buss, E., Hall, J. W. III, and Grose, J. H. (2004). "Spectral integration of synchronous and asynchronous cues to consonant identification," *J. Acoust. Soc. Am.* **115**, 2278–2285.
- Cooke, M. (2006). "A glimpsing model of speech perception in noise," *J. Acoust. Soc. Am.* **119**, 1562–1573.
- Dorman, M. F., Loizou, P. C., and Rainey, D. (1997). "Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs," *J. Acoust. Soc. Am.* **102**, 2403–2411.
- Dubno, J. R., Horwitz, A. R., and Ahlstrom, J. B. (2006). "Spectral and threshold effects on recognition of speech at higher-than-normal levels," *J. Acoust. Soc. Am.* **120**, 310–320.
- French, N., and Steinberg, J. (1947). "Factors governing the intelligibility of speech sounds," *J. Acoust. Soc. Am.* **19**, 90–119.
- Gagné, J. P. (1988). "Excess masking among listeners with sensorineural hearing loss," *J. Acoust. Soc. Am.* **83**, 2311–2321.
- Grant, K. W., and Braida, L. D. (1991). "Evaluating the articulation index for auditory-visual input," *J. Acoust. Soc. Am.* **89**, 2952–2960.
- Grant, K. W., Tufts, J. B., and Greenberg, S. (2007). "Integration efficiency for speech perception within and across sensory modalities by normal-hearing and hearing-impaired individuals," *J. Acoust. Soc. Am.* **121**, 1164–1176.
- Hall, J. W., Buss, E., and Grose, J. H. (2008). "The effect of hearing impairment on the identification of speech that is modulated synchronously or asynchronously across frequency," *J. Acoust. Soc. Am.* **123**, 955–962.
- Healy, E. W., and Bacon, S. P. (2002). "Across-frequency comparison of temporal speech information by listeners with normal and impaired hearing," *J. Speech Lang. Hear. Res.* **45**, 1262–1275.
- Howard-Jones, P. A., and Rosen, S. (1993). "Uncomodulated glimpsing in 'checkerboard noise'," *J. Acoust. Soc. Am.* **93**, 2915–2922.
- Kasturi, K., Loizou, P. C., Dorman, M., and Spahr, T. (2002). "The intelligibility of speech with 'holes' in the spectrum," *J. Acoust. Soc. Am.* **112**, 1102–1111.
- Leek, M. R., and Summers, V. (1996). "Reduced frequency selectivity and the preservation of spectral contrast in noise," *J. Acoust. Soc. Am.* **100**, 1796–1806.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Lippmann, R. (1996). "Accurate consonant perception without mid-frequency speech energy," *IEEE Trans. Speech Audio Process.* **4**, 66–69.
- Miller, G. A., and Licklider, J. C. R. (1950). "The intelligibility of interrupted speech," *J. Acoust. Soc. Am.* **22**, 167–173.
- Moore, B. C., Shailer, M. J., and Schooneveldt, G. P. (1992). "Temporal modulation transfer functions for band-limited noise in subjects with cochlear hearing loss," *Br. J. Audiol.* **26**, 229–237.
- Noordhoek, I. M., Houtgast, T., and Festen, J. M. (1999). "Measuring the threshold for speech reception by adaptive variation of the signal bandwidth. I. Normal-hearing listeners," *J. Acoust. Soc. Am.* **105**, 2895–2902.
- Noordhoek, I. M., Houtgast, T., and Festen, J. M. (2000). "Measuring the threshold for speech reception by adaptive variation of the signal bandwidth. II. Hearing-impaired listeners," *J. Acoust. Soc. Am.* **107**, 1685–1696.
- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). "Speech recognition with primarily temporal cues," *Science* **270**, 303–304.
- Stelmachowicz, P. G., Jesteadt, W., Gorga, M. P., and Mott, J. (1985). "Speech perception ability and psychophysical tuning curves in hearing-impaired listeners," *J. Acoust. Soc. Am.* **77**, 620–627.

- Studebaker, G. A. (1985). "A 'rationalized' arcsine transform," *J. Speech Hear. Res.* **28**, 455–462.
- Turner, C. W., Chi, S. L., and Flock, S. (1999). "Limiting spectral resolution in speech for listeners with sensorineural hearing loss," *J. Speech Lang. Hear. Res.* **42**, 773–784.
- Turner, C. W., Souza, P. E., and Forget, L. N. (1995). "Use of temporal envelope cues in speech recognition by normal and hearing-impaired listeners," *J. Acoust. Soc. Am.* **97**, 2568–2576.
- Tyler, R. S., Hall, J. W., Glasberg, B. R., Moore, B. C. J., and Patterson, R. D. (1984). "Auditory filter asymmetry in the hearing impaired," *J. Acoust. Soc. Am.* **76**, 1363–1376.
- Viemeister, N. F. (1979). "Temporal modulation transfer functions based upon modulation thresholds," *J. Acoust. Soc. Am.* **66**, 1364–1380.
- Warren, R. M., Riener, K. R., Bashford, J. A., Jr., and Brubaker, B. S. (1995). "Spectral redundancy: Intelligibility of sentences heard through narrow spectral slits," *Percept. Psychophys.* **57**, 175–182.