Notes and Comment

In defense of the right and left audiograms: A reply to Coren (1989) and Coren and Hakstian (1990)

PIERRE L. DIVENYI and KARA M. HAAPT
*Speech and Hearing Research Facility, Veterans Affairs
Medical Center, Martinez, California*

Coren (1989) noticed that the various air-conduction pure-tone thresholds obtained from one ear of an individual in quiet tend to display a high correlation. In addition, Coren and Hakstian (1990) reported a consistently high correlation between pure-tone thresholds in the two ears of the same individual. On the basis of these observations, these authors proposed that pure-tone thresholds measured at audiometric frequencies in one ear, and those measured in both ears of an individual, be collapsed into a single average threshold value. The present paper disputes the usefulness and the appropriateness of such a proposal.

In two recent articles, Coren (1989) and Coren and Hakstian (1990) noted that pure-tone thresholds in one ear of a given individual (measured at 6 audiometric frequencies covering the five-octave range between 250 and 8,000 Hz), as well as pure-tone threshold estimates in the two ears of the same individual, exhibit a high degree of correlation. On the basis of this observation, these authors propose to replace, in group data, (1) the pure-tone thresholds obtained at the 6 frequencies for each ear of an individual with a single number representing their average, and (2) the pure-tone thresholds obtained in the two ears of each individual with a single number representing either the average threshold of the two ears or the average threshold in the better ear. In other words, these authors would like to substitute a single number for either one or both of the left and right audiograms every time the object of study is a group of individuals, rather than a single individual. Although it is undeniable that pure-tone audiometric thresholds obtained at different frequencies in one or both ears of an individual often show high positive zero-order correlations,¹ there are several valid reasons for not collapsing the (usually 12) threshold values into one number. The following paragraphs represent an attempt to bring

up several points in support of the audiogram, and to argue that Coren's, and Coren and Hakstian's, proposal is inappropriate, of limited use, and misleading.

On Discounting Ears Headfirst: A Case Against "Counting Heads, Not Ears"

Audiological research, as a matter of principle, is aimed at issues related to the hearing status, rather than that of the ears, of individuals. The measure that serves as the starting point (and, quite often, the only point) to determine an individual's hearing status is the audiogram, that is, estimation of pure-tone audibility thresholds in quiet at various frequencies (usually six). The overall hearing level and the shape of the right and left audiograms, as well as the difference between the two audiograms, provide information on the basis of which the audiologist can establish the hearing status of the individual. The audiograms and the assessment of hearing for each ear of the same individual can, and often do, differ. It is the audiological significance of differences between left and right thresholds at each frequency that constitutes the most powerful argument against Coren and Hakstian's (1990) suggestion—namely, to consider, in group data, only one threshold estimate per frequency (i.e., left-right average or one of the ears) instead of two. There are, however, additional reasons for arguing in favor of treating the two audiograms as separate data, even if thresholds measured at all frequencies in the two ears of each individual in a group were found to be correlated, and even if such a correlation had been ignored by a great number of early and contemporary auditory studies.²

First, Coren and Hakstian's (1990) proposition is inappropriate because, statistically speaking, pooling data from correlated groups (or ears) yields zero-order statistics in which the correlation term does not appear. Thus, although the authors are correct in stating that the variance of the difference between the means of two variables is inversely related to the magnitude of the correlation between the two variables, these statistics will not enter into consideration as long as left-ear and right-ear thresholds are not compared. Since none of the studies cited by Coren as using incorrect statistical inference actually compares right-ear and left-ear data, any correlation between the thresholds in the two ears is irrelevant.

Second, Coren and Hakstian's (1990) proposition is uninformative. If audiograms in the two ears of an individual are, indeed, correlated, one should ask how, in a group of individuals with vastly different hearing, such a correlation is likely to arise. For one, it is known that thresholds in individuals with normal hearing oscillate

The authors gratefully acknowledge comments by James Jerger and two anonymous reviewers on an earlier version of the manuscript. This research was supported by the Department of Veterans Affairs and by a grant from the National Institute on Aging. Correspondence should be addressed to P. L. Divenyi, Speech and Hearing Research, V. A. Medical Center, Martinez, CA 94553. (e-mail: internet.marva4!pdivenyi@ucdavis.edu).

about the audiometric zero, that is, groups of subjects without hearing loss should have uncorrelated right-left audiograms. Furthermore, since correlation of pairs of data is dominated by only a few high-valued data pairs, adding a small number of subjects having a moderate-to-high bilateral hearing loss to a group of individuals with normal (or nearly normal) hearing will produce a correlation much higher than that of the group without those few subjects.

To illustrate this point, we performed a correlation analysis on pure-tone thresholds measured in 200 elderly individuals (Jerger, Jerger, Oliver, & Pirozzolo, 1989).³ By definition, a sizable portion of these individuals had bilateral hearing loss (mainly at the higher frequencies), although the hearing of many of the subjects was relatively normal. The 12 between-frequency and between-ear zero-order correlations computed for this set of data ranged from .142 to .833 across the six frequencies and the two ears. A multivariate analysis of variance, with left and right audibility thresholds as the two dependent variables, revealed significant effects of overall hearing level (i.e., thresholds averaged across six frequencies and two ears) [Wilks's $\Lambda = .378$, $F(2,1187) = 977.629$, $p < .001$] and frequency [$\Lambda = .979$, $F(10,2374) = 2.552$, $p < .005$], and a significant average hearing level $\times$ frequency interaction [$\Lambda = .763$, $F(10,2374) = 34.370$, $p < .001$]. Neither the effect related to subjects nor that of any one of the subject-related interactions was significant. After controlling for those variables and interaction terms that had significant effects (by way of a partial correlation analysis), the correlation diminished: the "raw" overall right-left correlation of .867 decreased to .806 when controlling for average hearing level, to .579 when controlling for average hearing level and frequency, and to .484 when controlling for average hearing level, frequency, and their interaction—a correlation still significantly different from 0, but by no means high.

The finding that overall hearing level is the strongest controlling factor in left-right correlations would predict that similarly high correlations could be found in a group of *randomly paired ears*, as long as this group contained some high-valued pairs. Since the Jerger et al. (1989) database consisted of audiological results of elderly individuals, and since a number of these individuals had significant hearing loss, we predicted that we would find a high correlation between pairs of thresholds measured in left and right ears belonging to different individuals. To test this prediction, we scrambled the ear data of the Jerger et al. database and generated audiogram pairs with the sole constraint that no individual's two ears could form a pair. Left-right correlations of 30 random pairings occupied a range between .284 and .425 (z - r transformed mean: $.374 \pm .042$). This "phantom left-right symmetry" disappeared after controlling for overall level and frequency: after taking out these two factors, the partial correlation came close to 0.

Two conclusions can be drawn from the above correlation analysis. For one, it is inaccurate to consider $r = 0$ to be the null hypothesis for left-right pure-tone thresh-

old correlations uncorrected for overall level and frequency effects.⁴ Furthermore, the fact that the correlation was close to 0 after controlling for level and frequency is tantamount to saying that the relatively high "raw" correlations, to begin with, were due to a certain number of high left-right threshold pairs. The presence of such high threshold pairs, however, is likely whenever the subject sample includes individuals with bilateral hearing loss, and when high-frequency thresholds are expected to be higher than low-frequency thresholds—as they usually are. Thus, the statement that left and right audiograms in a randomly selected group of individuals will be highly correlated is scarcely more informative than saying that a human hand tends to have five fingers.

Finally, Coren and Hakstian (1990) attribute the left-right correlations to environmental and physiological factors having the potential to affect hearing thresholds in both ears of an individual to an equal degree. Our third objection to Coren and Hakstian's proposition is that this is misleading because it implies that the left-right audibility threshold correlation derives solely from comparable *sensitivity* measures in the two ears. Although one should not dispute the probable influence of such factors in groups of individuals with definite hearing loss, it is unclear how these factors could have generated the high left-right correlations reported by the authors for their subsample of young subjects with normal hearing.⁵ Indeed, according to Coren and Hakstian, left-right correlations for this subsample were only slightly lower than those calculated for the complete sample, which included individuals of all ages (up to 92 years). Since, in the absence of any hearing loss, auditory thresholds are expected to be 0 dB hearing level on the average, the only factor that could account for the reported high correlation of left-right thresholds is a possibly idiosyncratic *constant-response criterion* that each subject adopts. Pure-tone audiometry, as used in most clinics, relies on a simplified version of the Method of Limits (which was also the procedure used by Coren, 1989, and Coren & Hakstian, 1990). This is a procedure known for producing results in which the effects of sensory capability and response proclivity are inseparable (Campbell & Moulin, 1968). If, for example, a subject with normal hearing has the habit of responding to anything that vaguely resembles a tone, his thresholds measured by this method will be uniformly low in *both ears*, unless he deliberately chooses a different strategy for each ear, which is unlikely. It was demonstrated long ago (Watson, Franks, & Hood, 1972) that the entire usable range of the psychometric function for audibility in quiet can extend from a mere 3.5 dB (strictest criterion) to as much as 10 dB (laxest criterion), that is, it can cover the same range as the difference between detection thresholds corresponding to d' values of 4.0 and 0.1. Since subjects with normal hearing would be expected to differ from each other only with respect to their response criteria, pooling the audibility thresholds of 10 normal subjects who display a lax response criterion with those of 10 other normal subjects who use a conservative criterion should yield a high left-right correlation. Moreover, since test-retest scores

Table 1
Four Typical Pure-Tone Audiograms and Their Correlations

Audiogram	Pure-Tone Thresholds (dB)						Correlations		
	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz	A	B	C
A	10	5	0	5	10	5			
B	35	30	25	35	35	30	.868		
C	10	5	10	30	50	70	.212	.203	
D	15	20	15	25	55	30	.500	.434	.677

are known to be correlated within the same individual—as Coren and Hakstian correctly point out—it is most likely that a cautious subject will have relatively *high* thresholds, and a notorious risk-taker *low* thresholds, at 3-, 6-, or 12-month intervals. This would be expected even if their *auditory sensitivities* (as measured independently by means of some criterion-free measure, such as two-alternative forced-choice detection thresholds or evoked otoacoustic emissions) were proven to be comparable, since individuals are known to hold their criteria fairly constant (Campbell, 1967).

Thus, beyond pointing out that left-right auditory-threshold correlations exist, raising the issue of these correlations without probing into their possible causes is misleading.

On the Polyvalence of “Equipollent Components of the Audiogram”

Pure-tone audiograms consist of thresholds measured at frequencies ranging between 250 and 8,000 Hz, that is, a five-octave span. Since A- or C-weighted thresholds obtained at adjacent octaves are not expected to display any great variation, threshold values evaluated across a group of individuals tend to be correlated across frequencies, the way Coren (1989) reports and our Table 1 shows. Arguing that correlated quantities convey redundant information, Coren proposes to replace the six threshold values with a more convenient single average. Although the nonindependence of the component thresholds of the audiogram has been recognized by audiology, and serves as the basis for various traditionally used averaging schemes, adopting the overall average as the only measure has never been considered because of the important information that such a measure would ignore—information required for the diagnosis of a disorder and/or an indication of the need for further testing. Therefore, the appropriateness of giving undue emphasis to the overall average of pure-tone thresholds is questionable.

One task of audiologists is the differentiation of audiograms, both those of different individuals and those obtained in the two ears of the same individual. The *contour* of the thresholds measured at six points of the five-octave frequency range has a high diagnostic value. Table 1 illustrates four frequently encountered audiograms, each considered to represent diagnostically different types: one normal, one indicating a flat hearing loss across frequencies, one indicating a sloping hearing loss beyond 1 kHz, and one showing a notch at 4 kHz. The same table also shows correlations between pairs of the four audiograms; they fall in the range extending from .212 (A-C)

to .868 (A-B). Curiously, although not surprisingly for audiologists, all pairs are positively, and some highly, correlated. Therefore, averaging the thresholds within each of the four audiograms would lead to overlooking profound differences between them, that is, to discarding crucial information. As an alternative to expressing auditory sensitivity in each ear as a single average number, audiologists resort to other methods. The most frequently adopted practice is to represent thresholds as two pure-tone averages: PTA1 (average of the thresholds obtained at 250, 500, and 1,000 Hz) and PTA2 (average of the thresholds obtained at 1, 2, and 4 kHz). Although this representation fails to identify certain particularities of the audiogram (say, a notch), and although the two quantities it creates are still nonindependent, it still detects a sloping hearing loss, which is the most frequent form of sensorineural pathology. However, PTA1 and PTA2 represent only one method designed to reduce the audiogram to two numbers; other methods have also been adopted (e.g., average hearing level and slope, used mainly in Europe; e.g., Dreschler & Plomp, 1980). Thus, two quantities are considered adequate to characterize a typical audiometric profile. To obtain a more refined description of audiograms, however, multivariate analysis techniques are necessary (Festen & Plomp, 1983; Verschuure, van den Wijngaart, Brocaar, & Nagels, 1985). Overall hearing level, as proposed by Coren (1989), is insufficient to characterize audiometric thresholds, except for individuals with a flat hearing loss (i.e., a rather rare class of cases) and individuals with normal hearing (i.e., the most trivial cases).

Conclusions

Coren's (1989) and Coren and Hakstian's (1990) observation, namely, that a correlation between the thresholds measured at different audiometric frequencies in the same ear and a correlation between thresholds measured in the two ears of an individual are common, is indisputable. However, we do not see this as a problem. Despite the lack of independence of the separate threshold measurements, the frequency contour of the audiogram as well as the difference between an individual's right and left audiograms convey important diagnostic information to the audiologist. Furthermore, pure-tone audiometry represents only *one piece of information* in the assessment of the hearing status of a given individual. Pure-tone thresholds are, most frequently, the starting point in the assessment, and the weight assigned to them will depend on the outcome of other audiological tests. Without the *a priori* expectation that the two ears of an individual

should yield comparable (i.e., correlated) results, and the knowledge that thresholds in a normal ear are essentially zero, diagnosis of a disorder would be much more difficult. Collapsing audiograms across frequencies and across ears would suppress important information and is, therefore, unjustified.

REFERENCES

- CAMPBELL, R. A. (1967). Estimation of an observed criterion or beta. *Journal of the Acoustical Society of America*, **42**, 267-268.
- CAMPBELL, R. A., & MOULIN, L. L. (1968). Signal detection audiometry: An exploratory study. *Journal of Speech & Hearing Research*, **11**, 402-410.
- CHABA (1988). Report of the Working Group on Speech Understanding and Aging. *Journal of the Acoustical Society of America*, **83**, 859-895.
- COREN, S. (1989). Summarizing pure-tone hearing thresholds: The equipollence of components of the audiogram. *Bulletin of the Psychonomic Society*, **27**, 42-44.
- COREN, S., & HAKSTIAN, A. R. (1990). Methodological implications of interaural correlation: Count heads not ears. *Perception & Psychophysics*, **48**, 291-294.
- DIVENYI, P. L., & OLIVER, S. K. (1991). *Perception of speech in speech noise by elderly and young listeners: Effect of the spatial arrangement of signal and masker*. Manuscript submitted for publication.
- DRESCHLER, W. A., & PLOMP, R. (1980). Relations between psychophysical data and speech perception for hearing-impaired subjects: I. *Journal of the Acoustical Society of America*, **68**, 1608-1615.
- FESTEN, J. M., & PLOMP, R. (1983). Relations between auditory functions in impaired hearing. *Journal of the Acoustical Society of America*, **73**, 652-662.
- JERGER, J. F., JERGER, S., OLIVER, T., & PIROZZOLO, F. (1989). Speech understanding in the elderly. *Ear & Hearing*, **10**, 79-89.
- MARKIDES, A. (1980). The relation between hearing loss for pure tones and hearing loss for speech among hearing-impaired children. *British Journal of Audiology*, **14**, 115-121.
- VERSCHUURE, J., VAN DEN WUNGAART, W. S. I. M., BROCAAR, M. P., & NAGELS, M. M. (1985). Methods of analysis of large numbers of audiograms. *Audiology*, **24**, 2-14.
- WATSON, C. S., FRANKS, J. R., & HOOD, D. C. (1972). Detection of tones in the absence of external masking noise: I. Effects of signal intensity and signal frequency. *Journal of the Acoustical Society of America*, **52**, 633-643.
- WILEY, T. L., OVIATT, D. L., & BLOCK, M. G. (1987). Acoustic-immittance measures in normal ears. *Journal of Speech & Hearing Research*, **30**, 161-170.

NOTES

1. Analysis of audiometric data collected in our laboratory (Divenyi & Oliver, 1991) as well as analysis of an audiological database consisting of 200 elderly subjects (Jerger et al., 1989) confirms the presence of consistent between-frequency and between-ear pure-tone threshold correlations.

2. Despite the indisputable occurrence of pooling threshold data across ears, our cursory survey of the literature seems to indicate that it is highly debatable whether the "vast majority of early studies" and a "significant portion" of current ones are, as Coren and Hakstian (1990) suggest, guilty of the practice of considering as independent pure-tone thresholds measured in the left and right ears of individuals. However, consensus is not the issue. Rather, one should ask oneself whether there are valid reasons for evaluating auditory data across ears, as opposed to individual subjects. As such, the choice of the ear as the unit of analysis was eminently defensible in all three relatively recent studies cited by Coren and Hakstian as guilty of applying statistically incorrect procedures. Markides (1980) was testing the application of a simple formula for expressing hearing loss *across different types* of audiograms. (It is assumed that there were individuals whose left and right audiograms were placed in different categories and, therefore, were kept separate in the analysis.) Verschuure et al. (1985) described a nonstandard statistical procedure (i.e., principal component analysis) for the classification of audiograms with respect to orthogonal features. To demonstrate the general applicability of their method, the authors *deliberately* chose audiograms obtained in ears of members of the same family with a documented history of a presumably genetic, rarely encountered audiometric configuration (i.e., a subject sample known for its *correlated* audiograms). Wiley, Oviatt, and Block (1987) were investigating middle ear functions in individuals with normal hearing. They measured pure-tone thresholds for the sole purpose of excluding ears with hearing outside normal limits. (An indication that they succeeded in excluding some ears is that they ended up with an odd number.) We fail to see in what way the correlation of right and left pure-tone thresholds could have influenced the results, let alone the conclusions, of these three studies.

3. We thank James Jerger and his colleagues for making their database available to us for this analysis.

4. In their correlation analysis, Coren and Hakstian (1990) controlled for age as a factor. This operation did not appreciably decrease the magnitude of their zero-order correlations, possibly because intersubject variability in audiological data of elderly individuals is notoriously high (CHABA Report, 1988).

5. We presume that the majority of Coren's (1989) subjects under 30 years of age were audiometrically normal.

(Manuscript received September 25, 1990;
revision accepted for publication January 16, 1992.)