G. M. Kuhn

The Phi coefficient as an index of ear differences in dichotic listening

Reprinted from CORTEX,
Vol. IX, 1973, pp. 447-457

LA TIPOGRAFICA VARESE
Via Tonale 49, Varese (Italia)
1973

CORTEX is an international journal devoted to the study of the interrelations of the nervous system and behavior, particularly as these are reflected in the effects of brain lesions on mental functions.

The editor in chief is Prof. Gildo Gastaldi (*Clinica delle Malattie Nervose e Mentali, Via F. Sforza 35, Milano, Italia*). Coeditors are Prof. J. de Ajuriaguerra (Genève), Prof. E. Bay (Düsseldorf), Prof. Arthur L. Benton (Iowa City), Dr. Macdonald Critchley (London), Dr. G. Ettlinger (London), Dr. Norman Geschwind (Boston), Dr. Robert J. Joynt (Rochester, N.Y.), Prof. F. Lhermitte (Paris), Prof. A. R. Luria Moskva), Prof. K. Poeck (Aachen) and Dr. J. Semmes (Bethesda).

Articles may be submitted in English, French, German, Italian or Spanish to the editor in chief or to coeditors. Each article must include a detailed summary both in the original language and in English.

Subscription rate per volume (4 numbers): $ 16. Subscriptions should be sent to *CORTEX c/o La Tipografica Varese, Via Tonale 49 Varese, Italia.*

NOTE

THE PHI COEFFICIENT AS AN INDEX OF EAR DIFFERENCES IN DICHOTIC LISTENING

Gary M. Kuhn[1]

(Haskins Laboratories, New Haven)

Introduction

Studdert-Kennedy and Shankweiler (1970) have applied the index $\frac{R-L}{R+L}$, were

$R =$ the number of correct right-ear responses

$L =$ the number of correct left-ear responses,

to measure ear differences in dichotic listening tests. If the task of the subject in such tests is to identify one stimulus on each dichotic presentation, as was the case above, then this index may yield its maximum value of $\pm$ 1 regardless of the subject's level of performance. But if the task is to identify both stimuli on each dichotic presentation, then the maximum value of the index decreases rapidly, as overall performance rises above 50%.

In order to avoid this ceiling effect in the two-response paradigm, Studdert-Kennedy has suggested that only those trials on which one stimulus is correctly reported should be included in the computation of the ear advantage (see the review in Halwes, 1969: 24). There are, however, a few points about the effect of applying the index in this way that may be worthwhile to keep in mind.

First, the number of one-correct trials may vary considerably across subjects. As a result of such differences in sample size, we cannot necessarily have equal confidence, in the statistical sense, in ear advantages of equal reported magnitude.

Second, the proportion of one-correct trials may vary considerably across subjects. Thus the ear advantage reported over one-correct trials could look very similar for two subjects whose advantages over all trials (measured for statistical significance) were of different magnitude.

Third, it may be that the proportion of one-correct trials varies systematically across levels of performance. But a measure of ear difference that does not take performance into account is one that assumes that overall performance gives no information about ear advantage. Clearly this would be an interesting assumption to test.

[1] Also University of Connecticut, Storrs.

Application to two responses per trial

For these reasons it may be desirable to apply over all trials in a two-response paradigm an index of ear difference whose computed values can be compared directly for statistical significance, independent of the level of performance. Such an index may be derived from the chi².

If we let

L = the number of correct left-ear responses

R = the number of correct right-ear responses

T = the number of dichotic presentations or "trials"

we may establish as the null hypothesis that the ears contribute equally to any observed R + L. Then for a two-response paradigm, we can express the expected outcome of a subject's performance in the following contingency table:

Response category		Ear of presentation: Left	Ear of presentation: Right	
	Correct	$\frac{R+L}{2}$	$\frac{R+L}{2}$	R+L
	Incorrect	$T-\frac{R+L}{2}$	$T-\frac{R+L}{2}$	2T—(R+L)
		T	T	2T

However, a subject's observed performance will be distributed according to the following table:

Response category		Ear of presentation: Left	Ear of presentation: Right	
	Correct	L	R	R+L
	Incorrect	T—L	T—R	2T—(R+L)
		T	T	2T

(1)

To compute the chi² of the difference between the observed and expected frequencies of these two tables we sum the values of the following table:

		Ear of presentation	
		Left	Right
Response category	Correct	$\dfrac{\left[L - \dfrac{R+L}{2}\right]^2}{\dfrac{R+L}{2}}$	$\dfrac{\left[R - \dfrac{R+L}{2}\right]^2}{\dfrac{R+L}{2}}$
	Incorrect	$\dfrac{\left[(T-L) - (T - \dfrac{R+L}{2})\right]^2}{T - \dfrac{R+L}{2}}$	$\dfrac{\left[(T-R) - (T - \dfrac{R+L}{2})\right]^2}{T - \dfrac{R+L}{2}}$

where each cell is of the form

$$\frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

and

$O =$ the observed number of responses for $cell_{ij}$
$E =$ the expected number of responses for $cell_{ij}$.

The sum of the evaluated expressions of these four cells is a chi^2 with one degree of freedom, which can be used as a measure of the observed ear advantage and as a test of the H_0.

Note that the sum of the correct row equals

$$\frac{2\left[R - \dfrac{R+L}{2}\right]^2}{\dfrac{R+L}{2}} = \frac{\dfrac{(R-L)^2}{2}}{\dfrac{R+L}{2}} = \frac{(R-L)^2}{R+L}$$

and that the sum of the incorrect row equals

$$\frac{2\left[(T-L) - (T - \dfrac{R+L}{2})\right]^2}{T - \dfrac{R+L}{2}} = \frac{\dfrac{(R-L)^2}{2}}{T - \dfrac{R+L}{2}} = \frac{(R-L)^2}{2T-(R+L)}.$$

A simplified form for stating the four-cell sum is then

$$chi^2 = \frac{(R-L)^2}{R+L} + \frac{(R-L)^2}{2T-(R+L)}.$$

This sum is a linear transform of the index $\dfrac{R-L}{R+L}$ with the y intercept changing as a function of performance.

If the observed number correct actually is the same for the two ears, then the four-cell sum will take on a value of Ø. At the other extreme, if one ear reports every stimulus correctly and the other ear none, then the generated value will reach its maximum of

$$\frac{T^2}{T} + \frac{T^2}{T} = 2T.$$

We may normalize the scale of possible values of the four-cell sum. A normalizing procedure which reduces the maximum value of the sum to 1.Ø and its distribution under the H_0 to that of a normal variate consists of dividing the computed chi² by N, (N = 2T here), and then taking the square root of the resulting value.

$$chi^2_{norm.} = \sqrt{\frac{chi^2_{comp.}}{N}}$$

It turns out that the values of this normalized four-cell sum are equal in magnitude to those of the "phi coefficient," since

$$chi^2 = N\ phi^2$$

(Walker and Lev, 1953:272). The phi coefficient is a correlation coefficient for two independent, dichotomously measured dimensions. If dimension 1 is either R or L and dimension 2 is either C or I then the 2 × 2 contingency table from which the strenght of their phi could be evaluated would look like the following:

		1		
		L	R	
2	C	b	a	a+b
	I	d	c	c+d
		b+d	a+c	a+b+c+d

where a, b, c, and d are cell frequencies. For the special instance where the column totals are equal,

$$a + c = b + d$$

the computational formula for the phi coefficient reduces to

$$phi = \frac{a - b}{\sqrt{(a + b)(c + d)}}$$

(Walker and Lev, 1953:272).

Given this relationship between phi and the chi² and the fact that the column sums in the observed frequencies contingency table (1) are indeed identical, we may compute the value of the normalized chi² index from

$$phi = \frac{R - L}{\sqrt{(R + L)[2T - (R + L)]}}$$

One reason for favoring computation of the value of this index through the phi formula is that the direction of the ear advantage, i.e., its sign, is retained. Computed in this way, the index can be thought of as yielding a value of correlation between correct performance and "right earedness": a negative value indicates a left-ear advantage.

By way of example, suppose that in a dichotic listening test of 100 trials, where the task was to report both stimuli on each trial, a subject gave the following performance:

$$L = 72$$
$$R = 88.$$

His ear advantage would be

$$\text{phi} = \frac{R - L}{\sqrt{(R + L)\,[2T - (R + L)]}} = \frac{88 - 72}{\sqrt{(160)(40)}} = .20$$

Critical values

A table of the smallest significant or "critical" values of the index may be constructed by letting

chi^2 = the value of $\text{chi}^2_{1\ \text{df}}$ with the desired level of significance
N = 2T, i.e., the total number of responses

and solving the equation

$$\text{phi} = \sqrt{\frac{\text{chi}^2}{N}}$$

The following table has been calculated in this fashion and is included here for illustrative purposes. From the table we see that the probability of obtaining an ear advantage as large as the one observed for the hypothetical subject above is < .01.

	Probability under the H_0 that phi ≥ PHI		
2T	.05	.02	.01
48	.282	.335	.371
80	.219	.260	.288
96	.200	.237	.262
100	.195	.232	.257
120	.178	.212	.235
160	.154	.183	.203
192	.141	.167	.185
200	.138	.164	.182
240	.126	.150	.166
320	.109	.130	.144
384	.100	.118	.131
400	.097	.116	.128
480	.089	.106	.117
640	.077	.091	.101
960	.063	.075	.083
1000	.061	.073	.081

A comparison of indices

For the sake of comparison, the ear advantages of five hypothetical subjects have been computed using three indices:

1. $\frac{R - L}{R + L}$ over all trials
2. $\frac{R - L}{R + L}$ over one-correct trials
3. phi over all trials

If the data for the 100 trials of a two-response paradigm are

Subject	Over all trials		Over one-correct trials	
	L	R	L	R
1	87	93	2	8
2	77	83	2	8
3	67	73	2	8
4	57	63	2	8
5	47	53	2	8

then the values of the ear advantages would be

Subject	$\frac{R - L}{R + L}$ all	$\frac{R - L}{R + L}$ one-correct	phi	P (phi ≥ PHI)
1	.033	.600	.100	not significant at $\alpha = .05$
2	.037	.600	.070	"
3	.042	.600	.065	"
4	.050	.600	.061	"
5	.060	.600	.060	"

50% performance

If a subject's performance level over a given set of trials is 50%, then his errors equal his number correct

$$2T - (R + L) = R + L$$

and his total number of responses equals twice his performance

$$N = 2T = 2(R + L).$$

His four-cell sum is then

$$\text{chi}^2 = \frac{(R - L)^2}{R + L} + \frac{(R - L)^2}{R + L} = \frac{2(R - L)^2}{R + L}.$$

And since

$$\text{phi} = \sqrt{\frac{\text{chi}^2}{N}}$$

we have

$$\text{phi} = \sqrt{\frac{\frac{2(R-L)^2}{R+L}}{2(R+L)}} = \sqrt{\frac{2(R-L)^2}{2(R+L)^2}} = \frac{R-L}{R+L}$$

Since the values of the two indices are identical for the case of 50% performance, it might appear that values of $\frac{R-L}{R+L}$ as computed over one-correct trials could be compared directly for statistical significance. This is not so, of course, if the size of the one-correct sample varies from subject to subject.

One response per trial

Using the same computational formula,

$$\text{phi} = \frac{R-L}{\sqrt{(R+L)[2T-(R+L)]}}$$

the phi index could also be applied to the data of a one-response directed recall listening test.

In this application, only a correct response from the requested ear is counted as correct under either condition of recall. Also, the quantity T is set equal to the number of trials under either condition of recall.

The null hypothesis here is that the ears contribute equally to the requested R + L. Phi is computed once over both conditions.

It does not seem to be appropriate to apply the phi to the data of a one-response, free-recall paradigm, since the performance of either ear may conceivably exceed half the total number of responses, or looked at another way, since an incorrect response cannot be assigned to either ear.

Summary

The phi correlation coefficient is proposed as an index of ear differences in dichotic listening tests. It is proposed specifically for the two-response paradigm, where, as an index of ear difference over all trials, it would be statistically appropriate for correlation with overall performance.

Using the same computational formula, the phi index may also be applied to the results of a one-response, directed-recall listening test.

The interest of the phi index lies in the fact that for a constant size of response set and number of dichotic trials, its values may be directly compared for statistical significance.

Acknowledgments. I would like to acknowledge the assistance of B. McK. Johnson, L. Katz, C. Darwin, and D. Goldstein, who influenced my first thoughts on these matters, and to extend special thanks to M. Studdert-Kennedy and D. Shankweiler for criticizing and discussing an earlier draft of this paper.

References

HALWES, T. G. (1969) *Effects of dichotic fusion on the perception of speech*, Supplement to Haskins Laboratories Status Report on Speech Research.

STUDDERT-KENNEDY, M., and SHANKWEILER, D. (1970) *Hemispheric specialization for speech perception*, "J. Acoust. Soc. Amer.," *48*, 579-594.

WALKER, H. M., and LEV, J. (1953) *Statistical Inference*, Holt, Rinehart and Winston, New York.

G. M. Kuhn, Haskins Laboratories, 270 Crown Street, New Haven, Conn. 06510 U.S.A.