G21-2-1

PERCEPTUAL CONFUSIONS AMONG LETTERS OF THE ALPHABET

An initial look at call-sign mis-identifications

Gerry Thomas

A common problem encountered when DX'ing the medium-wave frequencies is failing to hear a completely readable ID. Many times, one or more letters of the call-sign are garbled or masked to such an extent that they only reach the status of "sounded like." This degradation of the call-sign can be the result of several factors -- impulse noise (QRN), speech or tonal interference (QRM), long- and short-term amplitude variation (QSB and beating), and so on. However, not the least significant "degrader" of signal quality is the "white noise" that is inherent in the receiver, the DX'er's environment, etc. This "shushing" sound can be an effective masker of any number of different acoustical stimuli and has been in use as such in scientific laboratories for years. As a result, there is a fairly large body of scientific literature dealing with the confusibility of speech sounds which occur in a background of white noise. Unfortunately, (for our purposes), the vast majority of this research has used phonemes (basic speech sounds) as target stimuli and not letters of the alphabet, per se. Those few studies that are referenced as using letters of the alphabet as targets were either (1) conducted under government contract during WW II and are difficult to locate today, (2) used only some of the letters of the alphabet, or (3) failed to report exact probabilities of confusion for all letters. Regarding the research on phoneme perception, it is possible to predict the letter confusibility (since letters are constructed of phonemes) that might occur in a DX'ing situation. However, this alternative alone was not favored for a couple of reasons. First, since most communications receiver's headphones/speakers have a frequency response of between 200 Hz and 2300 Hz, the relevance of studies employing different bands of frequencies is questionable. Secondly, some of the findings reported in the phoneme literature are, at first glance, somewhat counter-intuitive and require an empirical test in a DX'ing situation. For example, it has been reported that the phoneme with the sound "ee" is sometimes confused with the phoneme with an "oo" sound (in the presence of white noise or low pass filtering). Very briefly, a possible explanation for this occurrence has to do with the fact that the "ee" and "oo" have first formants (vocal tract resonances) which are very close in frequency, but second formants which are very different in frequency. Since higher frequencies tend to be masked more effectively than lower frequencies, the relatively greater degradation of the higher frequency second formant results in a confusion between the two sounds. A similar formant relationship exists between "ā" and "oh," and "ī" and "r" sounds, though to a lesser degree. The implication of this finding is that one might expect to confuse the letter "T" not only with rhyming letters (e.g., "P," "D," "C," etc.) but also with letters such as "Q" and "U."

Since I have access to a psychoacoustics laboratory, I decided to investigate this possibility and at the same time obtain an approximation of the confusibility of all letters that can occur in a situation that as closely as possible simulatesDX'ing.

Method

The 26 letters of the alphabet were recorded in a background of white noise and presented randomly at a rate of one every three seconds to listeners whose task it was to "identify" the letters as they occurred. During each listening session, each letter occurred singly at least three times, and as a member of a letter pair at least twice. Since it was not feasible to represent all 650 permutations of letter pairs in this scheme, it was decided that each letter would be preceded once by the letter "V" and once by the letter "K." An additional five letter pairs were included for their potential confusibility, yielding a list length of 135 items. Four tapes were generated -- two spoken by a male and two by a female voice. One tape by each speaker was created such that the average power of the spoken letter was about 5 dB(SPL) above the total power of the noise (individual letters were allowed to vary in level as they do in natural speech). This resulted in an overall readability which I judged to be between "poor" and "fair." The remaining two tapes were generated such that no letter utterance exceeded the total power of the noise, thereby yielding a "very poor" rating. A total of 24 persons listened to the "poor-fair" tape and an additional 14 persons listened to the "very poor" tape.

Equipment

For those who are interested, the white noise was produced by a random noise generator and combined in an audio mixer with the letters which were spoken into a Shure Commando microphone and passed through a power amplifier. The output of the mixer was then passed through a Krohn-Hite band-pass filter (cut-off frequencies — 200Hz and 2800 Hz) and recorded on a Sony TC-353D tape recorder. The first three subjects listened to the tapes through Grason-Stadler headphones while seated in an Industrial Acoustics chamber. Subsequent subjects listened to copies of the tapes on a General Electric cassette deck in the free field.

Results

Table 1 lists the combined (male-female voices) error rates for the letters of the alphabet that were presented at the "poor-fair" readability level. That is, when the letter "V" was presented, 84% of the time it was perceived to be some other letter. Or, in other words, "V" was mis-IDed 84% of the time.

TABLE 1

V	.84	C	.23
P	.67	X	.11
D	.65	N	.11
B	.60	Y	.06
*Q	.47	M	.05
Z	.40	H	.05
F	.36	R	.03
E	.32	A	.02
T	.31	U	.02
S	.27	J	.01
*O	.27	I	.01
K	.25	L	.00
G	.24	W	.00

* A disproportionate number of errors on this letter occurred with the female voice.

The overall error rate for the "poor-fair" list was 24% with the male voice having an error rate about 5% less than the female's.

Because the ranking of letter error rates is dependent upon both the speaker and the listener, a different sample of speakers and listeners could result in a modification of this list. However, it should be mentioned that an attempt (though feeble) was made to obtain a cross-section of listeners. That is, although a plurality of listeners lived in the Southeast, the scores of listeners from MA, VT, KS, IA, WA, PA, NY, and IN (students) were obtained and included in the list. Nonetheless, this list should be viewed as being approximate.

Table 2 is a listing of the confusions that occurred when the readability was at the "poor-fair" level. It is intended to be entered using the "sounded like" letter. For example, what sounded like an "A" was actually an "A," 57% of the time; a "K," 21% of the time; an "O," 18% of the time; and an "H," 4% of the time. Furthermore, any additional confusions that may have occurred under "very poor" readability conditions are contained within parentheses. Since it is possible to imagine "hearing" almost any sound in a dominant white noise field, those letters included within parentheses indicate those heard by at least two listeners, though at a frequency of less than 1%. Therefore, these should not be viewed as being the only possible mis-ID's.

Overall, these confusions listed in Table 2 are in general agreement with predictions derived from the phoneme literature. Significant deviations from prediction are noted.

TABLE 2

Sounded like	Was	P
A	A	.57
	K	.21
	O	.18*
	H	.04
	(AJ, D, T)	

G21-2-2

* This value is somewhat inflated. On the "poor-fair" tape a disproportionate number of A-O confusions occurred with the female voice. On the "very poor" tape this confusion occurred more equally on the male and female tapes.

B	B	.45
	V	.28
	D	.15
	E	.04
	C	.03
	Z	.02
	G	.02
	(U, T)	

C	C	.74
	B	.07
	T	.06
	P	.05
	Q	.05
	Z	.02
	E	.02
	(V, K, H)	

D	D	.35
	B	.19
	V	.19
	E	.08
	G	.07
	Z	.06
	C	.04
	P	.02

E	E	.60
	V	.12
	B	.08
	D	.05
	T	.04
	P	.03
	G	.03
	C	.02
	Z	.02
	(U, Q, M, A)	

F	F	.83
	S	.11
	X	.06

G	G	.60
	D	.15
	V	.09
	Z	.05
	E	.05
	B	.03
	Q	.02
	U	.01
	(C, HE, P, T)	

/Some phoneme research suggests that "Z" should be most frequently confused with "G" but the above ranking was obtained on all four tapes.

H	H	.97
	K	.02
	A	.01

I	I	.93
	Y*	.04
	R	.03
	(L)	

*This proportion increases when "Y" is preceded by "W," that is, "WY."

J	J	.97
	G	.02
	K	.01

K	K	.97
	J	.02
	A	.01

L	L	.99
	O	.01
	(M)	

M	M	.85
	N	.14
	(S, C, B)	

N	N	.94
	M	.05
	(B, A)	

O	O	1.00
	(N, K, Z)	

P	P	.45
	T	.24
	Q	.08
	B	.07
	E	.05
	C	.05
	D	.04

Q	Q	.86
	P	.06
	T	.04
	E	.02
	KU	.01
	(C, D, B, U, O, KO)	

R	R	.96
	I	.04
	(Y, F)	

S	S	.75
	F	.19
	X	.06

T	T	.40
	P	.37
	Q	.16
	C	.04
	G	.02
	D	.01
	(B, E, U, I)	

U	U	.71
	E	.07
	Z	.07
	Q	.05
	V	.04
	G	.02
	D	.02
	P	.01
	(B, M)	

V	V	.25
	Z	.25
	B	.19
	D	.19
	G	.09
	E	.03
	(U, C)	

W	W	1.00

(Two listeners confused EE with W when readability was "very poor").

X	X	.71
	F	.16
	S	.13
	(M)	

Y	Y	1.00
	(I, R)	

Z	Z	.60
	V	.12
	G	.08
	D	.08
	B	.06
	C	.04
	E	.02
	(T)	

Those errors that occurred with letter pairs typically were the same confusions that occurred when the letters were presented singly. However, a couple of selected letter pairs resulted in confusions that did not follow this rule. The following letter pairs were confused sufficiently often to be noted below.

AJ — HA
HE — AG

In conclusion, the degree to which the confusion information contained in this article is valid in a given situation will depend upon the similarity of the speakers, listeners, and type of interference to those in this study. However, it is hoped that these results are sufficiently general to be relatively widely useful.

Notes: The tapes used were generated after-hours at the Psychoacoustics Laboratory of the Institute for the Advanced Study of the Communication Processes, University of Florida.

Selected References:

Miller, G. A. & Nicely, P. E. An analysis of perceptual confusions among some English consonants. Journal of the Acoustical Society of America, 1955, 27, 338-352.

Miller, R. L. Auditory tests with synthetic vowels. Journal of the Acoustical Society of America, 1953, 25, 114-121.

Stevens, K. N. & House, A. S. Speech Perception. In J. V. Tobias, Foundations of Modern Auditory Theory (Vol. 2). New York: Academic Press, 1970.