XV. SPEECH COMMUNICATION*

Prof. K. N. Stevens	N. Benhaim	J. S. Perkell
Prof. M. Halle	T. H. Crystal	Eleanor C. River
Prof. J. B. Dennis	H. J. Hebert	A. W. Slawson
Prof. J. M. Heinz	W. L. Henke	R. S. Tomlinson
Dr. Paula Menyuk	Y. Kato	J. T. Williams

A. DETECTION OF DSB SIGNALS OCCUPYING THE SAME RF SPECTRUM

1. Introduction

In a recent paper,[1] Bridges and Zalewski have discussed an approximate procedure for reducing the interference of one DSB signal upon another when the signals have overlapping spectra. They compared their scheme with various procedures for taking advantage of bandwidth - SNR tradeoffs – and concluded that it could be used for optimizing bandwidth utilization through the addition and later separation of many DSB signals. Contrary to the results and conclusions of Bridges and Zalewski, we shall show that since DSB is a linear modulation procedure that creates two shifted versions of the modulating signal spectrum, two DSB signals with overlapping spectrum may be perfectly separated by a finite, linear demodulation procedure. Because of the linearity of the modulation and demodulation process, there is no justification for comparison with nonlinear schemes nor for extending this procedure for more signals. The procedure could have important applications for protection of one DSB or AM signal against intentional or chance interference by another.

The ability to separate two overlapping DSB signals is recognized for the singular case of quadrature modulation, when the two signals have the same carrier frequency but one is 90° out of phase with the other. In this case orthogonal detection is used to remove one signal and recover the other. This same technique may be used, even if the carriers are not 90° out of phase, as long as there is some phase difference.

2. Iterative Demodulation Process

The signal to be demodulated r(t) is the sum of two double sideband signals having baseband signals p(t) and q(t), respectively,

$$r(t) = p(t) \cos \omega_p t + q(t) \cos \omega_q t, \tag{1}$$

where $\Delta = \omega_p - \omega_q$ is less than the bandwidth W of q(t), the signal we wish to recover. The phases of the signals have been chosen for mathematical convenience but this imposes

*This work is supported in part by the Air Force Cambridge Research Laboratories, Office of Aerospace Research, U.S. Air Force, under Contract AF19(628)-3325; and in part by the National Science Foundation (Grant GP-2495), The National Institutes of Health (Grants NB-04332-02 and MH-04737-04), and the National Aeronautics and Space Administration (Grant NsG-496).

no limitations on the demodulator, except the possible need for some phase shifters. The carriers are assumed to be available at the receiver.

Following Bridges and Zalewski,[1] the signal that is due to p(t) is removed by orthogonal detection, that is, multiplication by $2 \sin \omega_p t$ and lowpass filtering. What remains is

$$r^{(1)}(t) = q(t) \sin \Delta t, \tag{2}$$

the desired signal multiplied by a low-frequency sine wave. The signal q(t) could be recovered, as attempted by Bridges and Zalewski,[1] through multiplication by the reciprocal of $\sin \Delta t$ if there were no zero crossings. The same results may be achieved with a finite process through repeated orthogonal detection:

$$r^{(2)}(t) = r^{(1)}(t) \cdot \cos \Delta t = q(t) \sin 2\Delta t. \tag{3}$$

By this operation, the frequency of the sine wave has been doubled. Repeating the process N times gives

$$r^{(N)}(t) = r^{(1)}(t)\, 2^N \prod_{i=0}^{N-1} \cos (2^i \Delta t) = q(t) \sin (2^N \Delta t). \tag{4}$$

The purpose of repeating this operation is to make the frequency of the sinusoid multiplying q(t) greater than the bandwidth W of q(t). This will be achieved when

$$2^N \Delta \geqslant W \tag{5}$$

or when

$$N \geqslant \log_2 \frac{W}{\Delta}. \tag{6}$$

There is no upper limit on N. For N as given above, q(t) may be recovered from $r^{(N)}(t)$ by product demodulation and lowpass filtering.

$$\begin{aligned} q(t) &= \text{L. P. F. } \{r^{(N)}(t) \sin (2^N \Delta t)\} \\ &= \text{L. P. F. } \{q(t) \sin^2 (2^N \Delta t)\}. \end{aligned} \tag{7}$$

The demodulation process for deriving q(t) from $r^{(1)}(t)$ is the multiplication by the finite product D(t), summarized as follows:

$$D(t) = 2^N \prod_{i=1}^{N} \cos (2^{i-1} \Delta t) \sin (2^N \Delta t), \tag{8}$$

where $N \geqslant \log_2 \frac{W}{\Delta}$.

3. Spectrum and Closed Form of D(t)

The spectrum for D(t) may be found by convolving the spectral lines for each of the factors of the product. The first few and the final application of this procedure are shown in Fig. XV-1. The first N spectra come from the repeated multiplication by cosines, and the $(N+1)^{th}$ from the multiplication by a sine (see Eq. 8). From Fig. XV-1 it can be seen that D(t) contains all of the odd harmonics of Δ from the first to the $(2^{N+1}-1)^{th}$. All components have the same phase, so that D(t) may be written as a Fourier sine series.

$$D(t) = \sum_{k=1}^{2^N} \cdot \sin (2k-1)\Delta t \tag{9}$$

By expressing D(t) as two exponential Fourier series, one each for positive and negative frequencies, and then using the closed form of a geometric progression, we find

$$D(t) = \frac{\sin^2 \Delta Kt}{\sin \Delta t}, \tag{10}$$

where $K = 2^N \geqslant \frac{W}{\Delta}$. Because D(t) is always finite, it may be synthesized and used in a product demodulator. Note that there are lower limits but no upper limits on N and K.

The limit of (10) for $k = \infty$ is not obvious, but if one uses the procedure outlined above for (9), on the following equation

$$D_\infty(t) = \sum_{k=1}^{\infty} \sin (2k-1)\Delta t \tag{11}$$

the result is

$$D_\infty(t) = \frac{1}{\sin \Delta t}. \tag{12}$$

Comparison with (2) shows that multiplying $r^1(t)$ by $D_\infty(t)$ would give q(t), thereby demonstrating that the limiting case of D(t) is the same multiplicand used by Bridges and Zalewski.[1]

4. Noise Performance of the Demodulator

The noise performance of the demodulator is difficult to analyze because the demodulation destroys any assumed stationariness of the input-noise signal. This can be seen from the following development. If r(t) in (1) has bandlimited white noise added to it, $r^{(1)}(t)$ obtained from the first stage of demodulation will similarly have lowpass white noise added to it. If the noise added to $r^{(1)}(t)$ is denoted by N(t), the noise output of the demodulator will be

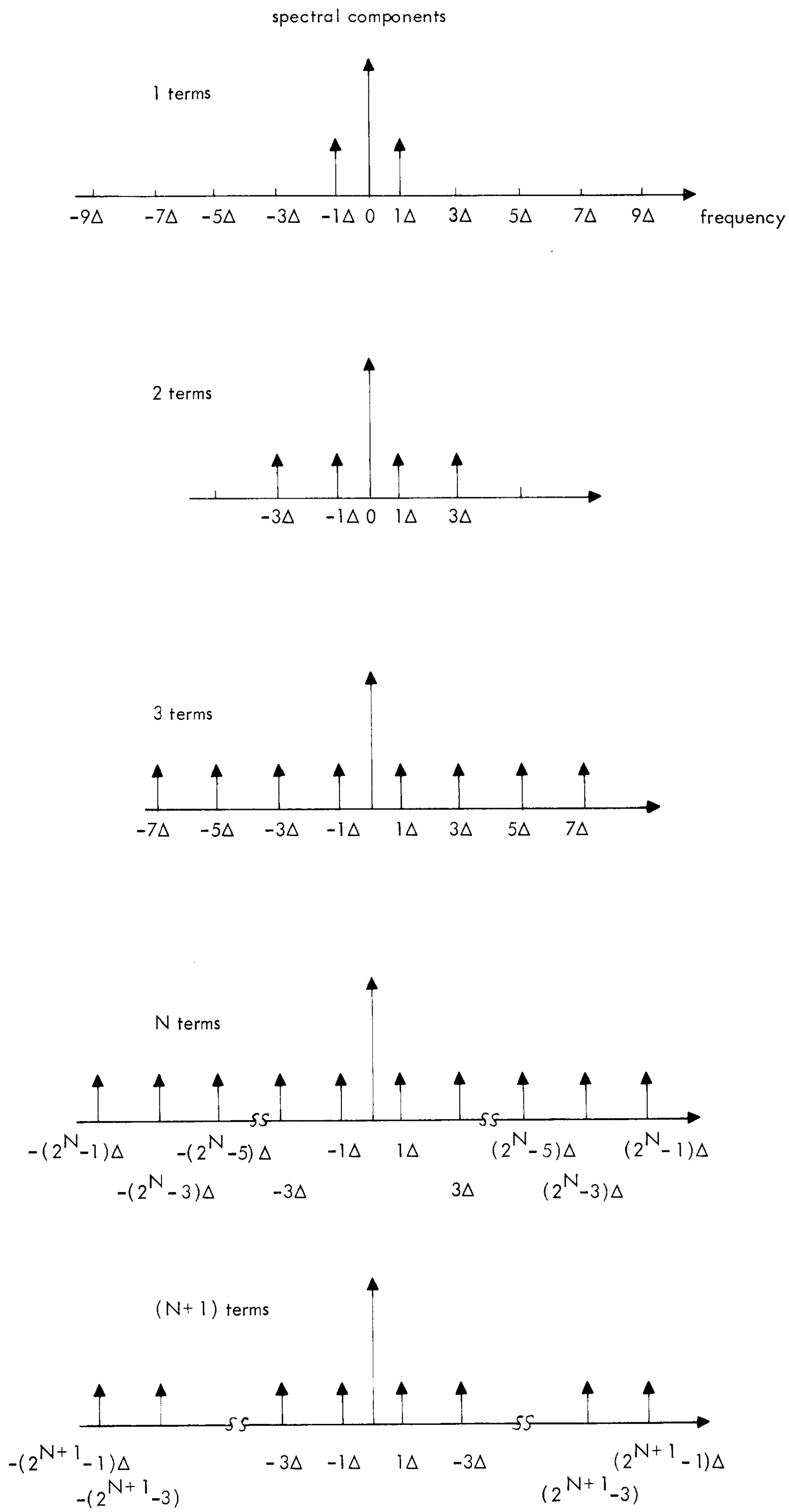

Fig. XV-1. Demodulator spectra.

$$n_o(t) = D(t) \cdot n(t). \tag{13}$$

One statistic which may be computed is the noise power, $\langle n_o(t)^2 \rangle$ which, since n(t) is independent of D(t), may be expressed as

$$N_o = \langle n_o(t)^2 \rangle = \overline{D^2(t)} \cdot \langle n^2(t) \rangle . \tag{14}$$

The last factor of (14) is the input noise power, N_i. Dividing both sides of the equation by this factor gives the ratio of output noise power to input noise power:

$$\frac{N_o}{N_i} = \overline{D^2(t)}. \tag{15}$$

This ratio is equal to the power of $D^2(t)$ and may be computed from the power of the individual components in (9):

$$\frac{N_o}{N_i} = \sum_{k=1}^{2^N} \frac{1}{2} = 2^{N-1}. \tag{16}$$

This indicates that the noise power increases exponentially with N and that the minimum N consistent with (6) would give the smallest output noise power. Combining this result with that shown in Eq. 5 gives

$$\frac{N_o}{N_i} \geq \frac{1}{2} \frac{W}{\Delta}. \tag{17}$$

Because (16) does not take into account the lowpass filtering which follows the multiplication by D(t), (17) is probably a good estimate of over-all demodulator performance for all values of K above the minimum.

5. Conclusion

Our discussion has shown that since a pair of overlapping DSB or AM signals (1) may be created by a linear of modulation process and (2) has an RF bandwidth that is at least as great as the sum of the baseband bandwidths, the combined signal may be demodulated and the two modulating signals recovered exactly. The demodulation process does, however, increase the noise power and make the noise probability distribution a periodic function of time.

T. H. Crystal

References

1. J. E. Bridges and R. A. Zalewski, Orthogonal detection to reduce common channel interference, Proc. IEEE 52, 1022-1028 (1964).

B. CUES USED IN SPEECH PERCEPTION AND PRODUCTION BY CHILDREN

In the study of language acquisition and development we would like to discover the particular cues in the language environment which are used by children to determine the linguistically distinct elements, and we would like to know the developmental course of this process. The process occurs from babbling to sentence formation within a short span of time. It seems reasonable to suppose that children have, at an early age, learned to classify speech sounds into grammatically meaningful elements. Also, it seems evident that there are particular cues that are used by children to discover linguistically distinct elements. Some experiments have been undertaken to explore some questions about the phonological component of the grammar which seem relevant to these hypotheses.

Experiment 1

Peterson and Barney[1] found in their experiment concerning vowel identification that children's productions were more difficult to identify than adult productions, and that some vowels are better understood than others. Characteristic shifts in vowel identification were found. They also hypothesized that language experience influences both the production and perception of sounds.

We have undertaken an experiment to explore the effects of the following factors on vowel perception and production: (i) immediate imitation of a model on vowel production by children: (ii) correlations of age of child, sex of child, and sex of adult presenting the stimulus materials on vowel production by children, word context versus nonsense syllable context on vowel production of both children and adults; and (iii) the age of the producer and effect of context on the identification of vowels.

The subjects were a boy and girl, both aged 4, and a boy and girl, aged 10 and 9, respectively. The list of words and the nonsense syllables that were used are shown in Table XV-1.

Table XV-1. Stimulus materials in Experiment 1.

Words	Nonsense Syllables
bid	bεv
bored	pɪf**
bard	buv
bit*	bɑrv
bead	bʊv
spit**	spɪf**
bird	bæv
bed	bɪv
beat*	bɔrv
booed	bɩv
bud	pɩf*
bad	bɜv
pit**	

*Included to examine vowel lengthening.

**Included to examine aspiration.

Each item on the lists was presented to each child for immediate repetition by an adult female, then by an adult male. The order of presentation in each list was changed in each presentation. Presentations and responses were recorded from an anechoic chamber. The tapes are now undergoing spectrographic analysis and will then be presented to children and adults for identification.

Experiment 2

An experiment has been designed to explore the questions of whether or not children can learn to distinguish equally well between members of a set of nongrammatical nonsense syllables and a set of grammatical nonsense syllables, and whether or not this capacity changes with age. The nongrammatical set is composed of initial consonant clusters that are not permissable in English, and the grammatical set is composed of initial clusters that are. Examples are given in Table XV-2.

Table XV-2. Sample stimulus materials in Experiment 2.

Nongrammatical	Grammatical
gzæk	glæk
pfʌm	klʌm
dlɛv	drɛv
srut	sput
dlɛd	krɛd
srɪk	blɪk
pfudʒ	prudʒ
gzos	glos

The task has three parts: (i) to learn to associate a nonsense syllable with a colored circle, each circle having the same color and having the name of one member of the set, and all circles presented simultaneously in a horizontal array; (ii) to immediately repeat all members of a different pair of sets after the experimenter; and (iii) to repeat each member of a set, one by one, after the experimenter. The first part of the task may give us information about differences, if any, in the perception and storing of nongrammatical and grammatical speech material, the second, differences in perecption, storing and reproduction of materials, and the third, differences in capacity to articulate the materials.

In a preliminary test of the stimulus materials, with two subjects, aged 5 and 9, used, it was found that only 4 members in a set for the first part of the task was too simple for both children, independently of the kind of material presented. Learning occurred after one or two trials. The second part of the task was very difficult, even with only 4 members in a set. Perfect repetition of the nongrammatical set was not accomplished after many trials by either child and long after exhaustion and disinterest had occurred. A sample study will be undertaken with small groups of children, aged from 3 to 7 years, with an expanded set for the first part of the task. Unlike the preliminary test, the two kinds of materials will be presented one week apart to counteract both learning effects and fatigue.

Further Research

An analysis of some available data on the perception and production of consonants by both adults and children has been undertaken. We have postulated that by describing the behavior observed in terms of a distinctive feature matrix[2] more information could be derived about cues used in the perception and production of consonants by children. The features looked at were: grave, diffuse, strident, nasal, continuant, and voice. We felt that this kind of analysis might possibly lead to productive research with children.

The data looked at were confusion in the perception of consonants by adults,[3,4] the consonant substitutions of children with functionally deviant speech[5], speech sounds which have been found to be most frequently defective in analyses of children's and adults' speech[6], and the acquisition and proportion of usage of consonants from age 3 months to 8 years[7,8] and adults.[9]

There has been very little research on the perception of consonants by children, and practically none in the age period before school. Also, it is difficult to determine, whether the consonants described as being acquired in developmental studies are in the ear of the hearer or the mouth of the child, since the data were gathered by phonetic transcriptions of either recorded or nonrecorded speech, rather than by spectrographic analysis. Also, a distinction must be made between productive use of consonants (as in morpheme formation) and production of consonants.

In this preliminary examination, we found that in general the features voicing and nasality were largely resistant to perceptual and productive confusion, and weighed heavily in the proportion of usage of consonants from age 3 months to 4 years. The importance of place (gravity and diffuseness) in proportion of usage differed over the age range of 3 months to adulthood. Gravity weighed more heavily during the early age periods of from 3 months to 30 months, and again of from 2 years to 8 years; then, diffuseness assumed a greater importance. In both the consonant substitutions by children and the perceptual confusions of adults, however, place was the feature that was least maintained. Consonants with the features stridency and continuancy are used proportionately less by both children and adults. They are consonants which are acquired last by children and they compose the list of those consonants most frequently found defective in the analyses that have been carried out.

As a possibility for future research, taking into account the results of the completed analysis of the data, one could look at the question of rank ordering of features in terms of ease of discrimination by young children, and even, possibly, infants.

Paula Menyuk

References

1. G. Peterson and H. L. Barney, Control methods used in a study of vowels, J. Acoust. Soc. Am. 24, 175-184 (1952).

2. R. Jakobson, M. Halle, and C. G. M. Fant, Preliminaries to speech analysis, Technical Report 13, Acoustics Laboratory, M.I.T., June 1955.

3. G. A. Miller and P. E. Nicely, Analysis of perceptual confusions among some English consonants. J. Acoust. Soc. Am. 27, 338-352 (1955).

4. W. Wicklegren, Psychology Department, M.I.T., Personal communication, 1964.

5. Paula Menyuk, Comparison of grammar of children with functionally deviant and normal speech. J. Speech Hearing Res. 7, 109-121 (1964).

6. M. H. Powers, Functional Disorders of Articulation, Handbook of Speech Pathology, L. E. Travis, (ed.)(Appleton, Century, Crofts, New York, 1957), pp. 707-769.

7. O. C. Irwin, Infant speech: Consonant sounds according to manner of articulation, J. Speech and Hearing Disord. 12, 397-401 (1947).

8. D. McCarthy, Research in language development: Retrospect and prospect, Monograph, Society for Research on Child Development, Vol. 24, pp. 3-25, 1959.

9. P. B. Denes, On the statistics of spoken English, J. Acoust. Soc. Am. 35, 892-905 (1963).