DOCUMENT RESUME

ED 125 005 95 CS 501 411

AUTHOR Rudegeair, Robert E.
TITLE A New Approach to Assessing Speech Sound Perception in Children. Professional Paper No. 31.
INSTITUTION Southwest Regional Laboratory for Educational Research and Development, Los Alamitos, Calif.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
REPORT NO SWRL-PP-31
PUB DATE 9 May 75
NOTE 21p.

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Auditory Perception; *Aural Learning; Educational Research; Elementary Education; Language Skills; *Listening Comprehension; *Speech Skills; *Test Construction

ABSTRACT

Traditional tests of auditory discrimination require children to match successively-presented syllables in short term memory. In the normal course of events perceptual confusions occur as a mismatch between an utterance and stored linguistic data. A speech sound perception task based on matching-to-memory was constructed and administered to 80 children from four age groups. Results show that traditional tests underestimate normal children's perceptual control over phonological oppositions. The data also reveal a novel dimension of perceptual confusion data related to marked and unmarked feature categories. (Author)

SWRL

SWRL EDUCATIONAL RESEARCH AND DEVELOPMENT

Professional Paper 31

9 May 1975

A NEW APPROACH TO ASSESSING SPEECH SOUND PERCEPTION IN CHILDREN

Robert E. Rudegeair

ABSTRACT

Traditional tests of auditory discrimination require children to match successively-presented syllables in short term memory. In the normal course of events perceptual confusions occur as a mismatch between an utterance and stored linguistic data. A speech sound perception task based on matching-to-memory was constructed and administered to 80 children from four age groups. Results show that traditional tests underestimate normal children's perceptual control over phonological oppositions. The data also reveal a novel dimension of perceptual confusion data related to marked and unmarked feature categories.

A NEW APPROACH TO ASSESSING SPEECH SOUND PERCEPTION IN CHILDREN

Robert E. Rudegeair

Many attempts have been made to measure young children's ability to discriminate among speech sounds. The prototype for such experiments was conducted by Travis and Rasmus who published their report in 1931. Their list of test items consisted of 331 contrasting pairs of syllables (e.g., /ta/-/da/) and 35 non-contrasting pairs (e.g., /ta/-/ta/). Subjects were asked to respond "same" or "different" to each pair as it was orally presented. The unwieldy list of all possible English speech sound contrasts has not been employed in subsequent testing. Since the prototypic experiment, testing has been confined to pairs that represent minimal phonemic differences. But the same-different procedure employed by Travis and Rasmus has endured and become standard procedure for testing speech sound discrimination in children. Most studies since that of Travis and Rasmus have been principally concerned with kindergarten and first-grade children. This age group has been of particular interest because speech sound discrimination is typically studied for the purpose of defining its relation to beginning reading or articulatory development or both.

The research attempting to define the relation of speech sound discrimination to other developing language skills is notably inconclusive. Moreover, the overwhelming concern with correlations among various language skills has distracted the research community from addressing fundamental questions about speech sound discrimination per se. The long-standing assumption has been that since speech sound contrasts are presented, speech sound discrimination is being tested. The results of several recent studies call this assumption into question and emphasize the need to reevaluate prevailing notions about speech sound discrimination.

Blank (1968) sought to determine whether differences in auditory discrimination scores between good and poor, seven-year-old readers are a function of failure to discriminate or a failure related to the complex cognitive processes demanded by the same-different task. Her study consisted of three experiments. In the first, good and poor readers exhibited differential scores on a same-different auditory discrimination task. The second experiment, using different subjects, was designed to eliminate the intervening cognitive judgment of same-different by having the child report directly the word pairs he heard (i.e. he had to repeat the pairs). In contrast to the first experiment, none of the children were eliminated because of a failure to understand the task. Blank suggests that, in experiment 1, the need to make a cognitive judgment of same-different posed a problem for the retarded reader which went beyond the perceptual demands of the task. In experiment 2, when same-different judgments were inferred on the basis of what the child reported (i.e., did he report two words that were the same or two words that were different?), the good readers still exhibited significantly lower error rates.

While it appeared that the cognitive judgment of same or different did not in itself cause the problem for the poor readers, an analysis of the kinds of errors made by the two groups still indicated that task strategy factors rather than differential perceptual abilities accounted for differential group performance. Poor readers showed an overwhelming tendency to perseverate pairs (i.e., match the second member to the first). Good readers did not exhibit this pattern in their error responses. The high frequency of perseverations suggests lack of attention across pair members. Thus, the third experiment was designed to remove the conditions for perseveration bias and still test accuracy of perception. Another group of children was required to repeat single words. First they were presented, one at a time, the first members of pairs employed in experiments 1 and 2, followed by the words that were originally second members. Each group exhibited a mean rate of correct imitation of 85% suggesting that the groups are not differentiated as a function of speech sound perception abilities. Cognitive demands of the task, extraneous to speech sound perception, appear to account for differential performance on a measure of speech sound discrimination.

In an experiment similar to Blank's, Beving and Eblen (1973) assessed the influence of the concepts "same" and "different" on speech sound discrimination performance of 30 children. A 25-item discrimination test using the same-different format was administered to three groups of ten four-year-old, ten six-year-old, and ten eight-year-old children. Three days later, the same children were given the same items and asked to repeat the syllable pairs in lieu of responding same or different. From the data in the second task, as in the Blank study, same-different judgments were inferred on the basis of what was said in repetition. Results showed that while all three groups exhibited significantly different error rates on the judgment test, performance on the repetition test was equivalent among groups. (Error rate = approximately 8%). But only for the four-year-old group was a significant difference found between the judgment scores and the repetition scores, judgment errors being significantly more numerous. Thus, it was concluded that while four-year-olds are able to discriminate among speech sounds, as shown by their performance on the repetition task, they appear to be unable to work with the concepts "same" and "different" (in the study by Blank, nine children were eliminated from experiment 1 because they did not understand how to respond same and different).

Other studies have also questioned the value of standard speech sound discrimination tests as measures of perceptual processes. The Wepman test of Auditory Discrimination (1958), for example, has been shown to yield significantly lower error rates when six-year-olds repeat the test (Rudegeair & Kamil, 1970). Berlin and Dill (1967) report similar findings when feedback and positive reinforcement is provided on the second test administration. Vellutino, DeSetto, and Steger (1972) report findings that indicate a response bias inherent in tests like the Wepman because of the imbalance of "different" and "same" pairs. Others have shown that errors on the Wepman test can be accounted for by dialect factors (Coller et al., 1968) or social class factors (Elenbogen & Thompson, 1972). These various research reports combine to suggest that errors on the standard speech sound discrimination tests may be a function of everything except the perceptual confusion of speech sounds. Yet such tests have been used to diagnose otherwise normal children as having inadequately developed speech sound discrimination.

In the present study a novel approach was taken in an attempt to measure perceptual confusability of consonant sounds in the context of a speech sound discrimination test. It was felt that, based on speech sound repetition data like that presented by Beving and Eblen (1973), children four years of age and older would show insignificant discrimination problems under normal hearing conditions. Such a hypothesis could not be realistically entertained so long as traditional test results were taken at face value. Error rates of 20% to 40% are commonly reported for 40-item tests administered to normal samples of first-graders.

Obviously, discriminability must be measured by an instrument radically different from the traditional test. Yet it must retain face validity as a test of speech sound discrimination. It must be designed to circumvent the memory and attention problems shown to accompany paired-word presentations and yet be simple enough to elicit the responses of preschool children.

Traditional tests require the child to discriminate two stimulus syllables spoken by someone else. In making a perceptual confusion a child, in the normal course of events, mismatches something someone says with something in his memory. For example, the utterance skirt may be perceived as shirt, cert, or even as a meaningless element such as /čət/. A plausible test of discriminability can be constructed on the basis of this normal process of matching-to-memory. Toward this end, a list of minimal consonant contrasts were selected and a familiar word which could be associated with a familiar object was selected to represent each speech sound in the list. All contrasts were distinguished by one, and sometimes two, acoustic distinctive features (Chomsky & Halle, 1968). Words were subjectively judged by the experimenter to be in the spoken lexicon of preschool children and a pilot study was conducted to check these judgments. For example, thumb was selected to represent /θ/. To test the contrast /θ/-/f/, the child was shown a thumb and asked, "Is this a /fəm/?" The child simply responded "yes" or "no." At some point in the test the child was shown the thumb and asked, "Is this a /θəm/?" Each test item was queried in this manner since one measure of the tests validity is the child's ability to answer "yes" in legitimate instances. Another control on test validity was achieved by having each child name all the test objects before the actual discrimination trials commenced. This provided some assurance that the words used were familiar to him.

Using a specific word for each speech sound provided an opportunity to test a novel dimension of speech sound discrimination which can be referred to as directionality. For example, misarticulation data indicate that children's consonant substitutions tend to occur in only one direction; /f/ is substituted for /θ/ yielding utterances like "my two front teef," but /θ/ is not typically substituted for /f/. Thus, it is legitimate to ask if unidirectionality is a characteristic of the developing phonological system as a whole, or if it is a characteristic confined to emerging articulatory processes.

Williams, Cairns, Cairns, and Blosser (1970), in an analysis of young children's misarticulation data, were able to account for the unidirectional nature of consonant substitution patterns in terms of markedness theory. In the data they analyzed, substitutions exhibited a movement from a phoneme marked on a given feature to a phoneme unmarked on that feature. If the same principle governs perceptual processes during the developmental period, it can be hypothesized that, in the present tasks, unmarked distortions are likely to be considered acceptable variants of the target while marked distortions will be rejected. In other words, less complex, first-order approximations of the target will elicit "yes" responses more frequently than more complex, first-order approximations. Presenting <u>thumb</u> and asking, "Is this a /fəm/?" tests the /f/-/θ/ confusion in one direction. Presenting <u>fish</u> and asking, "Is this a /θɪš/?" tests it in the other. It is possible that word familiarity represents a confounding factor in this regard, but control for this is presumed on the basis of the object naming task that preceded the discrimination trials.

METHOD

<u>Stimulus materials</u>. Fifteen minimal consonant contrasts involving 13 different consonants were selected for testing. Since each contrast was tested in two directions, the test list consisted of 30 contrastive words in addition to the 13 non-contrastive words (real words). Table 1 shows the consonants and the words selected to represent them as well as the 30 contrastive nonsense words. Concrete

objects corresponding to each of the test words were used in the object-naming task and in the discrimination trials themselves. The 43-word test list was recorded stereophonically, thus transmitting the identical signal over both channels. The item-form used was "Is this an X?" (where X is one of the 43 test words).

Table 1

Familiar Words Selected to Represent Initial Consonant Sounds and Real and Artificial Words Used to Test Discrimination of Phonetically Similar Consonants

Sounds		Words	Test Items							
1.	/p/	pipe	1.	pipe	14.	/taip/	27.	/kaip/	39.	/fəip/
2.	/t/	towel	2.	towel	15.	/paul/	28.	/kaul/	40.	/θaul/
3.	/k/	cup	3.	cup	16.	/pəp/	29.	/təp/		
4.	/b/	ball	4.	ball	17.	/dɔl/	30.	/gɔl/	41.	/mɔl/
5.	/d/	dog	5.	dog	18.	/bɔg/	31.	/gɔg/	42.	/nɔg/
6.	/g/	girl	6.	girl	19.	/bəl/	32.	/dəl/		
7.	/f/	fish	7.	fish	20.	/θɪš/	33.	/pɪš/		
8.	/θ/	thumb	8.	thumb	21.	/fəm/	34.	/səm/	43.	/təm/
9.	/s/	sock	9.	sock	22.	/θak/	35.	/šak/		
10.	/š/	shoe	10.	shoe	23.	/su/	36.	/ču/		
11.	/č/	chair	11.	chair	24.	/šer/				
12.	/m/	mouse	12.	mouse	25.	/naus/	37.	/baus/		
13.	/n/	nose	13.	nose	26.	/moz/	38.	/doz/		

Procedure. Children were tested individually in a mobile trailer that was located adjacent to their classrooms. Testing consisted of three phases: object naming, task training, and discrimination trials. The 13 objects were placed on a table before each child. The child was instructed to lift each one and say what it was. No child had difficulty naming the objects used, probably because difficult items were eliminated after the pilot study. Task training consisted of asking the child his name. If the child said his name was Robin, he was asked, "Is your name Lobin?"--"Is your name Wobin?"--"Is your name Robin?" This was a simple means of verifying that the child was capable of rejecting a distorted utterance. No child failed to respond appropriately. Each child was then given the following instructions by the experimenter:

> I am going to put earphones over your ears and play a tape. The man on the tape is going to try to say the names of the things on the table--just like you did. Sometimes he says the names right and sometimes he says the names wrong. I want you to tell me when he says them right. Say "yes" or "right" when he says them right and say "no" or "wrong" when he says them wrong. What would you say if the man asks, "Is this a bable?" [experimenter points to the table] . . . what would you say if the man asks, "Is this a table?"

Again no child showed any evidence of not comprehending the task. The instructions were followed immediately by the discrimination trials, again transmitted via the stereo headset. In conjunction with each query, the experimenter held the appropriate object up before the child. Responses were recorded by the experimenter on prepared data sheets. The test was administered twice to each participating child, once on each of two successive school days.

Participants. Eighty children were randomly selected from four age groups to participate in the study. Three-year-olds (N=12) and four-year-olds (N=18) were selected from their respective classes at a public day care center. Kindergarteners (N=25) and first-graders (N=25) were selected from their respective classes at a public elementary school. All children were monolingual speakers of English with no known hearing loss or articulation disorder.

RESULTS

In responding to the perception test all children demonstrated an ability to distinguish real words from distortions. Mean rates of accepting real words were above 90% for each age group; mean rates for accepting distortions were below 20% for each age group. These data are shown in Table 2.

The 30 nonsense words that represent distortions of real words contain 15 contrastive consonant sounds tested in two directions. In order to test reversibility of consonant confusions, it is possible to group contrastive sounds according to their relation to the target sound in terms of markedness theory. Since the distortions in the present study have initial consonants that differ from the target with regard to one distinctive feature, each distortion can be classified as marked or unmarked in relation to the target to which it

Table 2

Mean Rates of Accepting ("Yes" Responses) Real vs. Distorted Words for Each Age Group

Words	Age Groups			
	3 yrs.	4 yrs.	5 yrs.	6 yrs.
real words n = 13	92.0	95.3	98.6	98.0
nonsense words. n = 30	19.0	11.3	7.0	4.2

corresponds. Thus /moz/, a distortion of nose, is classified as a marked distortion since /n/ is unmarked on the feature coronal, while /m/ is marked on that feature. No other feature distinguishes /m/ and /n/. The reciprocal distortion /naʊs/ for the target mouse is an unmarked distortion. In Table 3, the 30 distortions presented in the perception test are categorized according to their marked or unmarked status.

The mean number of false positive responses, as well as standard deviations, for each of the four age groups appear in Table 4. The error rates are also presented. These rates are assumed to reflect perceptual confusion and, as in previous developmental studies of speech sound discrimination, decrease as a function of increasing age. To compare false positive response frequency across the two types of distortions as well as test the significance of the differences between

Table 3

Thirty Test Distortions According to Whether They are Marked or Unmarked in Relation to the Target With Which They Contrast

Contrast	Unmarked Distortions	Marked Distortions
p vs. t	/taip/ for pipe	/paʊl/ for towel
b vs. d	/dɔl/ for ball	/bɔg/ for dog
θ vs. f	/fəm/ for thumb	/θɪš/ for fish
m vs. n	/naʊs/ for mouse	/moz/ for nose
k vs. t	/təp/ for cup	/kaʊl/ for towel
g vs. d	/dɚl/ for girl	/gɔg/ for dog
k vs. p	/pəp/ for cup	/kaɪp/ for pipe
g vs. b	/bɚl/ for girl	/gɔl/ for ball
š vs. s	/su/ for shoe	/šak/ for sock
θ vs. s	/səm/ for thumb	/θak/ for sock
š vs. č	/ču/ for shoe	/šer/ for chair
θ vs. t	/təm/ for thumb	/θaʊl/ for towel
n vs. d	/doz/ for nose	/nɔg/ for dog
m vs. b	/baʊs/ for mouse	/mɔl/ for ball
f vs. p	/pɪš/ for fish	/faip/ for pipe

mean error rates exhibited by the four age groups, a two-way, mixed-model ANOVA was performed on the error data. The between-subjects factor was Age Group (3-, 4-, 5-, and 6-year-olds), and the within-subjects factor was Item Type (marked and unmarked distortions). The results of this analysis appear in Table 5.

Table 4

Mean Number of False Positive Responses, Standard Deviations, and Corresponding Error Rates on Marked and Unmarked Distortions for Each of Four Age Groups

Age Groups	Item Type	Mean # of errors	N	SD	Error rate (%)
3 yrs.	unmarked	6.83	12	3.48	22.77
	marked	4.58	12	2.72	15.27
4 yrs.	unmarked	4.22	18	2.39	14.25
	marked	2.55	18	1.42	8.33
5 yrs.	unmarked	2.36	25	2.17	7.86
	marked	1.84	25	1.56	6.13
6 yrs.	unmarked	1.28	25	1.28	4.26
	marked	1.24	25	0.90	4.13

Table 5

Source Table from a 2-way ANOVA (Groups X Stem Type)

Source	df	MS	F
Between Subjects			
Groups	3	137.32	21.33*
Error	76	6.43	______
Within Subjects			
Item Type	1	45.78	27.67*
Item Type x Groups	3	9.44	5.71*
Error	76	1.65	______

*p < .01

The analysis indicates a significant effect due to groups, a significant effect due to Item Type and a significant Group X Item Type interaction. Comparisons among group and interaction means were made according to a procedure specified by Winer (1962, p. 378). These comparisons showed, with regard to the group main effect, that three- and four-year-old scores are significantly different from one another as well as from the five- and six-year-old scores. The difference between five- and six-year-old scores was not statistically significant.

The Item Type effect reflects significantly less false positive responses to marked distortions. Subsequent pairwise comparisons of interaction means revealed that three- and four-year-olds exhibit less false positive responses to marked distortions while five- and six-year-olds do not respond differentially to unmarked and marked distortions.

False positive responses to distortions involving two particular contrasts, /f-θ/ and /m-n/, had the effect of greatly inflating overall group error rates. Relatively speaking /f-θ/ and /m-n/ distortions had unusually high error rates. Of all false positive responses made by six-year-olds, 65% are accounted for by /f-θ/ confusions (/θɪš/ for fish and /fəm/ for thumb). In fact, /f-θ/ errors represent a major portion of false positive responses for each age group: 27.7% of three-year-old errors, 47.5% of four-year-old errors, and 50.4% of five-year-old errors. This preponderance of /f-θ/ errors is a result of unusually high error rates for each age group in response to /θɪš/

and /fəm/. Each group accepted these items at least half the time with the exception of six-year-old responses to /fəm/. Error rates on these particular distortions were:

	3 yrs	4 yrs	5 yrs	6 yrs
/fəm/ (unmarked)	75%	72%	46%	26%
/θɪš/ (marked)	83.3%	88.9%	60%	56%

These data represent the only instances of a marked distortion eliciting higher error rates than its unmarked counterpart.

In contrast, /m-n/ confusions only appear problematic in response to the unmarked distortion /naʊs/ for mouse. False positive responses to /moz/ for nose are in line with false positive response frequencies to all of the other distortions. Error rates on distortions involving /m-n/ contrasts were:

	3 yrs	4 yrs	5 yrs	6 yrs
/naʊs/ (unmarked)	62.5%	30.5%	26%	16%
/moz/ (marked)	20.8%	8.3%	6%	0%

Like /m-n/ confusions, all other confusions tested showed a preponderance of false positive responses to unmarked distortions, hence the effect reported earlier.

DISCUSSION

The data from the speech sound perception task support several conclusions relevant to a more sensitive appraisal of developmental speech sound perception. In the first place, the low rates of false positive responses suggest that normal children are capable of adult-level discriminations among most minimally distinct initial consonant sounds at a relatively early age. Six-year-olds rejected distortions, with the exception of /f-θ/ substitutions, with at least 92% accuracy. Five-year-olds' responses, while more variable, were not significantly different from those of the older group. These results contradict the conclusion based on traditional assessment procedures that children in this age range (5-6 years) have difficulty discriminating minimally different sound contrasts. The data from a study by Templin (1957) are typical in this regard. In her investigation of developmental trends in speech sound discrimination, the analysis showed significant improvement from six to seven years old, indicating that some form of development was taking place across these age groups. Results from the present study reveal no reliable difference between the means when five and six-year-old scores are compared. Coupled with the fact that these groups exhibited near-perfect (/f/ - /θ/ substitutions aside, accuracy in judging distortions always exceeds 95%) discrimination accuracy, the conclusion can be drawn that age five is a more reasonable ceiling in the development of speech sound perception skills than age seven. The implication, then, is that Templin's data reflect development of skills necessary for performing a task involving successively presented spoken syllables and same-different judgments rather than any significant increase in sound discrimination ability.

The problem in successfully discriminating successively presented syllable pairs is a problem of matching in short-term memory. Accurate perception of speech sounds is a function of matching to memory. The traditional failure to respect this distinction has resulted in a distorted view of the typical child's control over phonological oppositions. The present findings support this conclusion and the technique employed demonstrates that tasks confined to testing the linguistic skills at issue are not precluded.

The finding that unmarked distortions elicit more false positive responses from three- and four-year-olds conforms to the predictions based on articulatory substitution data. Marked distortions are not likely substitutes for related phonological segments, nor are they likely candidates for confusion with related segments. This result strongly suggests that articulatory behavior and perceptual behavior are governed by the same central mechanism. In other words, emerging perceptual and productive control over contrasting segments is a function of a single, central processor. Such a conclusion supports the general belief that perception and production of phonological units are inextricably tied to one another, and that a model of developmental phonological competence will essentially account for either level of behavior. Only problems at the periphery will distort the predictive impact of such a model.

REFERENCES

Berlin, C. L., and Dill, A. C. Effects of feedback and positive reinforcement on the Wepman auditory discrimnation test scores of lower-class Negro and white children. Journal of Speech and Hearing Research, 1967, 10, 384-389.

Beving, B., and Eblen, R. E. "Same" and "different" concepts and children's performance on speech sound discrimination. Journal of Speech and Hearing Research, 1973, 16(3), 513-517.

Blank, M. Cognitive processes in auditory discrimination in normal and retarded readers. Child Development, 1968, 39, 1091-1101.

Chomsky, N., and Halle, M. The sound pattern of English. New York: Harper and Row. 1968.

Coller, A. R., Coleman, R., and Schwartz, S. Initial and final consonant relationships in speech-sound tests: a discrimination or response set problem. ERIC Doc. No. ED 032 131, New York Institute for Developmental Studies, New York University, 1968.

Elenbogen, E. M., and Thompson, G. R. A comparison of social class effects in two tests of auditory discrimination. Journal of Learning Disabilities, 1972, 5(4), 209-212.

Rudegeair, R. E., and Kamil, M. L. Assessment of phonological discrimination in children. Technical Report No. 118, 1970, Wisconsin Research and Development Center for Cognitive Learning, University of Wisconsin, Madison, Wisconsin.

Templin, M. Certain language skills in children. Minneapolis: University of Minnesota Press, 1957.

Travis, L. E., and Rasmus, B. The speech sound discrimination ability of cases with functional disorders of articulation. Quarterly Journal of Speech, 1931, 17, 217-226.

Vellutino, F. R., DeSetto, L., and Steger, J. A. Categorical judgment and the Wepman test of auditory discrimination. Journal of Speech and Hearing Disorders, 1972, 37(2), 252-257.

Wepman, J. M. Auditory discrimination test (Manual of Directions). Chicago: Language Research Associates, 1958.

Williams, F., Cairns, H., Cairns, C., and Blosser, D. Analysis in the production errors in the phonetic performance of school-age standard-English-speaking children. Final Report. Center for Communication Research. Austin, Texas, December, 1970.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1962