Cognition, 7 (1979) 323–331

Does awareness of speech as a sequence of phones arise spontaneously? *

JOSÉ MORAIS
LUZ CARY
JÉSUS ALEGRIA
and PAUL BERTELSON

Université libre de Bruxelles

Abstract

It was found that illiterate adults could neither delete nor add a phone at the beginning of a non-word; but these tasks were rather easily performed by people with similar environment and childhood experiences, who learned to read rudimentarily as adults. Awareness of speech as a sequence of phones is thus not attained spontaneously in the course of general cognitive growth, but demands some specific training, which, for most persons, is probably provided by learning to read in the alphabetic system.

Introduction

Alphabetic writing in first approximation represents speech at the level of units such as phone and phoneme.[1] Both spelling and reading in an alphabetic system imply, in addition to the ability to perceive minimal phonetic distinctions, an explicit knowledge of the phonetic structure of speech. For example, the reader/writer must not only be able to distinguish between *cat* and *bat*, but must also know that *cat* and *bat* consist of three units and differ only in the first.

An important question is how this knowledge is attained. In normal communication, people pay attention to meaning, not to the structural charac-

*Reprints may be obtained from José Morais, Laboratoire de Psychologie expérimentale, Université libre de Bruxelles, 117 av. Ad. Buyl, B-1050 Bruxelles, Belgium.

[1] While the term phone is generally used to indicate the more elementary units of speech that are perceptibly different, there is a considerable disagreement in the literature about the definition of phoneme. In the traditional perspective, the phoneme is any collection of phones whose differences are irrelevant to meaning distinctions; in the generative-transformational perspective, the phoneme is an abstract representation that depends on morphemic information and relates to pronunciation through a set of rules. For a discussion of the distinction between phone and phoneme, from the latter point of view, in relation to the alphabetic system, see Gleitman and Rozin (1977). In the present text we shall refer to analysis into phones rather than into phonemes, because the experimental task simply required our subjects to manipulate different sounds without regard for meaning.

teristics of the speech they hear and produce. However, conscious reflection on language and therefore explicit knowledge of the linguistic structures do occur. Awareness of speech as a sequence of phones, for instance, might appear spontaneously at some age, as a normal outcome of cognitive growth, through maturation and/or linguistic experience. Alternatively, it may require some specific training, which for most children is usually provided by reading instruction itself. The question is important not only from a theoretical point of view but also from a practical one: under the cognitive growth hypothesis, failures in learning to read can best be avoided by adjusting the age at which reading instruction is started to individual rates of development, while under the specific training hypothesis the solution should be sought in the improvement of educational practices.

That the ability to manipulate phones is related to success in learning to read has been largely documented. For instance, Savin (1972) signaled that children who failed to learn to read by the end of the first grade were generally unable to learn Pig Latin. This "secret language" requires the shifting of the initial consonant cluster of each word to the end of the word and the addition of the sound [ei]. This fact, however, may reflect either a delay in the spontaneous acquisition of the ability to analyse speech into phones or the inability to make abstract inferences about the sound system of language from its alphabetic representation.

Some observations on the linguistic behavior of preschool children would suggest that insight into the phonetic structure of language may be possible before formal learning to read and write. Read (1978) could elicit phonetically correct judgments of similarity for vowels in kindergarteners. Slobin's (1978) daughter engaged in rhyming play and noticed sound similarities in her own speech at 3;1: "eggs are beggs; more-bore". Preschool children apply the plural inflection to new words and appreciate the pronunciation of a sound in a word. However, the conscious manipulation of a particular phone or class of phones (like vowels, which are important in rhyme) does not necessarily imply awareness of speech as a sequence of phones. Phones that can be uttered in isolation may be more accessible, i.e., brought more easily to our awareness, than highly encoded ones. Awareness of such phones may be an example of awareness of a linguistic performance, rather than of a linguistic structure. The problem we consider here is how awareness of the phonetic structure, not of this or that phone, is attained.

The few studies in which the development of the ability to make an explicit analysis of utterances into phones has been investigated do not permit one to choose between the cognitive growth and the specific training hypotheses. In one of those studies (Zhurova, 1973), children were shown dolls with colored jackets and told, for instance, "the boy with the

yellow jacket is Yan, the boy with the green jacket is Gan, the boy with the white jacket is Whan", etc... . Then, they were tested for the retention of names and questioned about other dolls with colored jackets that had not been shown before (pink, violet, etc...). The rule for new jackets was used successfully by 12%, 39% and 100% of the children in the 4 to 5, 5 to 6 and 6 to 7 years age groups. In another study (Liberman, Shankweiler, Fischer and Carter, 1974), children were asked to play a tapping game, in which segments of a word spoken by the experimenter had to be indicated by the number of taps. The segments were either syllables or phones. The authors found that none of the nursery school children (mean age: 4 years 10 months) could segment by phone (i.e., reach a criterion of six consecutive errorless trials) while 46% could segment by syllable. The percentage of children who were able to segment by phone increased in the other groups: 17% of the kindergarteners (mean age: 5 years 10 months) and 70% of the first graders (mean age: 6 years 11 months).

In both the Russian and the American studies the most dramatic progress in segmentation performance occurred between ages 5 and 6. As the Haskins workers pointed out, this increase "might result from the reading instruction that typically begins between ages five and six. Alternatively it might be a manifestation of cognitive growth not specifically dependent on training" (Shankweiler and Liberman, 1976). A test of the issue, they suggested, would be provided by a developmental study of segmentation skills in children learning to read in a logographic system, such as Chinese, which does not demand explicit phonetic analysis. However, such a study, they pointed out later (Liberman, Shankweiler, Liberman, Fowler and Fischer, 1977), can no longer be carried out in China, because children now learn to read alphabetic text before they start studying the logographic characters.

Fortunately, testing readers of non-alphabetic systems is not the only possibility. In communities where the writing system is alphabetic, there remains a minority of adults who either have never been taught to read or have dropped out of school at a very early stage. Illiterate people should be unable to perform tasks requiring conscious phonetic analysis, if the improvement observed between ages 5 and 6 is related to reading instruction. On the contrary, if the improvement is the result of some cognitive growth process, independent of reading, they would, of course, succeed.

Method

The present experiment was run in a poor agricultural area of Portugal (Mira de Aire, district of Leiria). Subjects were all of peasant origin, but

most were now working in the textile industry. Thirty illiterate people (*I* subjects) and 30 people who learned to read beyond the usual age (*R* subjects) were tested. *I* subjects, 6 males and 24 females, were aged 38 to 60 and *R* subjects, 13 males and 17 females, were aged 26 to 60. Among *I* subjects, twenty had never received any instruction at all, four had been taught by their children to identify letters, and six had been in school for 1 to 6 months in childhood (some of them could "draw" their names). *R* subjects had attended classes for illiterate people organized by the government, by the Army or by industry. All were at that time 15 years old or more. Twenty-two, as a result, had received some kind of certificate and eight had failed to obtain any.

Two tasks were administered. In the "deletion" task, the subject had to delete the first phone from an utterance provided by the experimenter. In the "addition" task, he had to introduce an additional phone at the beginning of the utterance. Half the subjects in each group worked with one of the two tasks. For each task, five subjects worked with the phone [p], five with the phone [ʃ], and five with the phone [m]; three different groups of consonants (plosives, fricatives and nasals) were thus represented in the experiment. The test consisted of 15 introductory trials to illustrate the rule, and 20 experimental trials. The subjects were told that their task was to add (delete) one "sound" to the utterances produced by the experimenter. In the introductory trials, these utterances were non-words which became words by adding (deleting) the phone assigned to the subject. For instance "alhaço" became "palhaço" (clown) and "purso" became "urso" (bear). A correction procedure was used at that stage: when the subject failed to produce the correct response, the experimenter provided it. The experimental trials were of two types: in *W* trials, the experimenter uttered a word which, by the transformation rule, would become another word, for instance "uva" (grape) became "chuva" (rain), and vice-versa; in *NW* trials, the experimenter uttered a non-word which would become another non-word, for instance "osa" became "posa", "chosa" or "mosa" depending on phone condition. In both types of experimental trials, no information was provided after the subject's response. The subject had been told beforehand that on some experimental trials the correct response might be a non-word. All the words were of current use and, in all probability, were known by the subjects.

Results

In interpreting the results account must be taken of the fact that only *NW* trials provide unambiguous information regarding segmentation and fusion

abilities. In *W* trials, the correct response might be found by searching the lexicon for a similarly sounding word. *W* trials yielded in fact better performances than *NW* ones. On *NW* trials, *I* subjects gave a very poor performance and *R* subjects quite a good one: mean correct responses were respectively 19% and 72%. The pattern of results is nearly identical for the two tasks (Table 1).

Table 1. *Mean percentages of correct responses for each type of trial, task, and group of subjects. In parentheses, the percentage of subjects who attained 100% of correct responses.*

		Task			
		Addition		Deletion	
	Trials	*W*	*NW*	*W*	*NW*
Subjects	*I*	46 (13)	19 (0)	26 (7)	19 (0)
	R	91 (33)	71 (13)	87 (47)	73 (27)

Fifty percent of *I* subjects failed on all *NW* trials, while no *R* subject did. More than 50% of *R* subjects and only one of the *I* subjects gave 8 correct responses or more on the 10 *NW* trials (Figure 1).

I subjects failed whatever the target phone: mean correct responses on *NW* trials were 17%, 19% and 20%, for [p], [ʃ] and [m] respectively. *I* subjects who had been in school for some time in childhood or who had been taught the names of letters ($n = 10$) performed somewhat better on *NW* trials (30%) than the remaining subjects (13%). The difference approached significance at $p < 0.05$ by a one-tailed t test ($t = 1.696$; df = 28).

Within the *R* group, the mean percentage of correct responses on *NW* trials was 55% for the 8 subjects without a course certificate and 79% for the other 22. The difference is significant at $p < 0.025$ ($t = 2.41$; df = 28). On the other hand, *R* subjects who learned to read before age 25 ($n = 10$) did not perform significantly better than those who learned beyond that age (75% and 71% respectively; $t = 0.384$; df = 28).

The analysis of errors on *NW* trials revealed that only 19% of the incorrect responses made by *I* subjects involved the correct deletion or addition of the required phone plus some other transformation, while these kinds of responses represented 56% of the *R* subject's errors.[2] A tendency to produce

[2] An example is the response *pili* instead of *pécli*.

words in response to non-words was present in both *I* and *R* groups and accounted for, respectively, 46% and 32% of the errors; however, the proportion of wrong responses that both were words and involved the required phone[3] was much smaller in group *I* (6%) than in group *R* (28%). The great majority of errors made by *I* subjects can thus be linked to lack of awareness of phonetic structure, while an important portion of the errors made by *R* subjects were apparently due to some other cause.

Figure 1. *Number of subjects at the different levels of performance in the* I *and* R *groups (for* NW *trials only).*

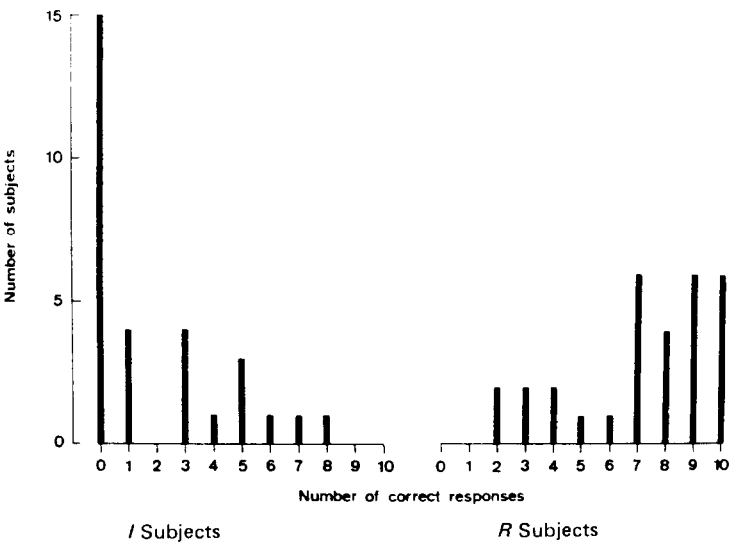

Table 2 shows the errors that occurred twice or more (over a maximum of five) in *NW* trials for each combination of group, task and phone. It should be noticed that the most frequent errors were generally words (except *bli*, *go* and the repetitions *mosa* and *maçuto*). The items in italics are those for which the phone to be deleted (or added) has not been deleted (or added). It should be noticed that this more frequent type of error was made by the subjects of group *I*, not by those of group *R*.

[3] An example is the word *poda* instead of the non-word *posa*.

Table 2. *Frequent errors in* NW *trials for each combination of group, task and phone. The first item is the stimulus and the second the response. The first number inside the brackets indicates the number of occurrences of the response; the second number indicates the total number of errors in the trial.*

	[p]	[ʃ]	[m]
Deletion task			
I Subjects	Puada - Ada (2/5) *Pobli - Pobre* (2/4) *Pecli - Pé* (3/4)	Chuada - Ada (2/5)	*Muada - Amuada* (3/5) *Mobli - Móvel* (3/5)
		Chube - Chuva (2/4)	
		Chimá - Má (3/5)	Mimá - Má (3/5) *Mosa - Mosa* (2/5)
		Chigó - Ó (3/5)	*Migó - Amigo* (3/5) *Maçuto - Maçuto* (2/5)
		Chabatá - Batata (2/5)	Mabatá - Batata (3/5)
R Subjects	Puada - Ada (2/2)	Chuada - Ada (2/3) Chobli - Bli (2/2) Chimá - Má (3/3) Chigó - Gó (3/4) Chabatá - Tá (2/3)	
Addition task			
I Subjects	*Imá - Irmã* (2/5) *Abatá - Batata* (2/5)		
R Subjects			Imá - Mãe (2/4)
	Açuto - Poço (2/3)	Açuto - Chuto (2/4)	

Discussion

Illiterate adults were unable to delete or add a phone at the beginning of a non-word, while adults from the same environment who learned to read in youth or as adults had little difficulty. It is interesting to note that the performance of the *I* subjects was slightly inferior to that of Belgian first graders aged 6 years who were tested in the third month of the school year with similar tasks (18% correct responses for deletion, 29% for addition). The performance of the *R* subjects was at about the same level as that of Belgian second graders aged 7 years and tested in the fourth month of the school year (73% correct responses for deletion and 79% for addition) (Alegria and Morais, 1979).

The extremely poor performance of the *I* subjects cannot be explained in terms of some general inability to manipulate speech segments or to under-

stand an inductive instruction. Cary and Morais (1979) have tested a group of 12 illiterates, from the same origin as those of the present experiment, with a more complex task which consisted in reversing the order of either phones or syllables (for instance, *cha* for *ach,* or *chave* for *vechá,* respectively) after inductive training. In the reversing phones condition the mean percentage of correct responses was 9% (ranging from 0% to 20%), while in the reversing syllables condition it was much higher: 48% (ranging from 13% to 93%).

The present results clearly indicate that the ability to deal explicitly with the phonetic units of speech is not acquired spontaneously. Learning to read, whether in childhood or as an adult, evidently allows the ability to manifest itself. Thus, it is not right to say that awareness of the phonetic structure of speech is a precondition for starting learning to read and write. The precondition for the acquisition of these skills is not phonetic awareness as such but the cognitive capacity for "becoming aware" during the first stages of the learning process. Of course, the present results do not mean that cognitive growth plays no part in the development of phonetic awareness. Specific training may not be effectual before some critical developmental stage. If awareness depends on instruction, it does not follow it necessarily. Successful instruction, on the other hand, depends on awareness. There is a reciprocal relationship between learning to read and the developmental changes in phonetic awareness.

Two important questions should now be examined. The first is to what extent phonetic awareness can be provoked by other stimulating experiences. Although for most children learning to read constitutes the exercise that renders the analysis of speech into its phonetic elements imperative, it is not necessarily unique to that function, and other kinds of training might presumably achieve the same effect.

The second question is to what extent the procedures used in recognizing and producing speech can be affected by awareness of speech as a sequence of phones. The fact that illiterates are not aware of the phonetic structure of speech does not imply, of course, that they do not use segmenting routines at this level when they listen to speech. But that fact should remind us of the risk we may incur in studying the mechanisms of speech perception through tasks that require conscious, explicit segmentation. Under the pressure of modern developments in linguistics and phonetics some psychologists were led to consider the so-called "psychological reality" of, for example, transformational grammars, or phones and phonemes. It is not always clear whether this kind of inquiry concerns implicit (tacit) or explicit knowledge (cf., a discussion of this point by Seuren, 1978). If the question concerns how we perceive speech, by first segmenting it either in phones

(phonemes) or in syllables – the question apparently considered by Savin and Bever (1970) and other authors – then it refers to tacit knowledge. The present results with illiterates are irrelevant to this question, but they urge us to distinguish between the prevalence of such or such a unit in segmenting routines at an unconscious level and the ease of access to the same units at a conscious, metalinguistic level.

References

Alegria, J. and Morais, J. (1979). Le développement de l'habileté d'analyse phonétique consciente de la parole et l'apprentissage de la lecture. *Archives de Psychologie,* in press.

Cary, L. and Morais, J. (1979). A aprendizagem da leitura e a consciência da estrutura fonética da fala. *Revista Portuguesa de Psicologia,* in press.

Gleitman, L. R. and Rozin, P. (1977). The structure and acquisition of reading. I: Relations between orthographies and the structure of language. In A. S. Reber and D. L. Scarborough (Eds.), *Toward a Psychology of Reading.* Hillsdale, Lawrence Erlbaum Associates.

Liberman, I. Y., Shankweiler, D., Fischer, F. W. and Carter, B. (1974). Reading and the awareness of linguistic segments. *J. Exper. Child Psychol., 18,* 201–212.

Liberman, I. Y., Shankweiler, D., Liberman, A. M., Fowler, C. and Fischer, F. W. (1977). Phonetic segmentation and recoding in the beginning reader. In A. S. Reber and D. L. Scarborough (Eds.), *Toward a Psychology of Reading.* Hillsdale, Lawrence Erlbaum Associates.

Read, C. (1978). Children's awareness of language, with emphasis on sound systems. In A. Sinclair, R. J. Jarvella and W. J. M. Levelt (Eds.), *The Child's conception of language.* Berlin, Springer-Verlag.

Savin, H. B. (1972). What the child knows about speech when he starts to learn to read. In J. F. Kavanagh and I. G. Mattingly (Eds.), *Language by ear and by eye.* Cambridge, Mass., MIT Press.

Savin, H. B. and Bever, T. G. (1970). The non-perceptual reality of the phoneme. *J. Verb. Learn. Verb. Behav., 9,* 295–302.

Scuren, P. (1978). Grammar as an underground process. In A. Sinclair, R. J. Jarvella and W. J. M. Levelt (Eds.), *The Child's conception of language.* Berlin, Springer-Verlag.

Shankweiler, D. and Liberman, I. Y. (1976). Exploring the relations between reading and speech. In R. Knights and D. J. Bakker (Eds.), *The neuropsychology of learning disorders: Theoretical approaches.* Baltimore, University Park Press.

Slobin, D. I. (1978). A case study of early language awareness. In A. Sinclair, R. J. Jarvella and W. J. M. Levelt (Eds.), *The Child's conception of language.* Berlin, Springer-Verlag.

Zhurova, L. Y. (1973). The development of analysis of words into their sounds by preschool children. In C. A. Ferguson and D. I. Slobin (Eds.), *Studies of child language development.* New York, Holt, Rinehart and Winston, Inc.

Résumé

Un groupe d'adultes analphabètes a été incapable de soustraire ou d'ajouter un phone au début d'un non-mot, mais ces tâches ont été facilement effectuées par un groupe de personnes dont l'environnement et l'expérience pendant l'enfance étaient similaires et qui ont appris à lire de façon rudimentaire à l'âge adulte. La prise de conscience de la parole comme une séquence de phones n'est donc pas acquise spontanément au cours du développement cognitif mais exige un entraînement spécifique, lequel, pour la plupart des personnes, est fourni probablement par l'apprentissage de la lecture dans le système alphabétique.