ISCA Archive
http://www.isca-speech.org/archive

2nd International Conference on
Spoken Language Processing (ICSLP 92)
Banff, Alberta, Canada
October 12-16, 1992

THE DEVELOPMENT OF LEXICAL EFFECTS ON CHILDREN'S PHONEME IDENTIFICATIONS

Michael S. Hurlburt and Judith C. Goodman

Department of Psychology, University of California at San Diego, La Jolla, California

10.21437/ICSLP.1992-89

ABSTRACT

We conducted two experiments to investigate the extent to which lexical knowledge affects children's interpretation of phonetic input. In Experiment one, five-year-old children and adults identified the initial phoneme in stimuli from 18 acoustic continua varying in voice onset time (VOT). For each continuum, one of the two possible phonemic categorizations formed a word (e.g., *desk*) and the other did not (e.g., *tesk*). Both children and adults tended to identify the initial phoneme as consistent with a word when the tokens were from the middle of a VOT continuum. In addition, children showed a lexical effect at the endpoints of some continua, where acoustic information was clear. Experiment two addressed one possible explanation of the finding that children, but not adults, are influenced by lexical context even when the acoustic input is clear. Children may adopt a strategy to report words, because of the large number of ambiguous targets (i.e., onsets from within the continua). If so, the lexical effect should disappear when all tokens are clear voiced and voiceless stop consonants from the endpoints of the continua. Five-year-olds identified endpoint stop consonants in words and nonwords. The lexical effect was greatly attenuated for these stimuli relative to the endpoint stimuli in Experiment 1, suggesting a strategic component to children's performance. Nonetheless, a small lexical effect remained for these items. The results suggest that five-year-olds engage in very similar processes to adults enabling then to use lexical context to identify phonemes. Under certain conditions, children weigh lexical context more heavily in making phoneme decisions than do adults.

INTRODUCTION

Top-down contextual knowledge interacts with bottom-up input to influence the way in which adults interpret sensory input. This is true in many perceptual domains. For example, consider how one interprets the central letter in the word "T/-\E" and in the word "C/-\T." Readers interpret the same physical input differently depending on the surrounding context. The study we report is concerned with an analogous problem in speech perception. Here too, a percept is determined both by the sensory input and by the context in which the signal occurs [1,2]. Adults are influenced by lexical context in tasks requiring phoneme decisions [3]. They interpret identical acoustic input differently depending on the lexical status of the context in which it occurs. For example, subjects identified an alveolar stop at the boundary of a voicing continuum as a /d/ in the context of [_esk] and as a /t/ in the context of [_est]. Our aim in this study is to examine the extent to which children also use lexical context to identify phonemes.

Lexical context stems from the listener's familiarity with the words of a language. The listener must develop a knowledge base of words in the language if lexical context is to influence phoneme identification. Hence, the use of lexical context in perceptual processing must be a developmental phenomenon. However, it is unclear at what age children will use lexical context to interpret the acoustic input or whether context influences their interpretations in the same manner as it does adults'.

There is little work that directly measures the role of lexical context in phoneme perception in children. The most relevant work [4] found that lexical context enables four- to six-year-old children to better interpret a degraded signal: children correctly reported more phonemes in CVC syllables masked by noise when those syllables formed words than when they formed nonwords. It is not clear from this study how lexical context interacts with acoustic input. Children might use lexical knowledge to fill in phonemes they completely missed or to weigh a decision when they have some partial specification of each phoneme. In our study, we examine whether lexical context affects kindergartners' identification of component phonemes varying along a voicing continuum in clear (unmasked) syllables.

Tasks which do not directly measure phoneme identification also suggest that children's phoneme judgments could be influenced by lexical context. These studies suggest that children under six years of age attend to the global structure of words because they have difficulty with phonetic segmentation [5,6,7]. If this is true, then lexical context may have a large influence when a young child is required to identify a sub-syllabic component. In addition, age-related differences in structural properties of the lexicon point to the possibility that children attend more to the overall sound pattern of a word than to fine acoustic detail. Words in five- to seven-year-olds' lexicons have fewer similar neighbors [8] than the same words analyzed in the adult lexicon, suggesting that children need not conduct as thorough an analysis of acoustic input as do adults. Instead, they might rely on lexical knowledge along with a partial analysis of relatively discriminable input to identify words.

Results from word-recognition measures, however, suggest that children might be limited in their use of lexical context. For example, several studies report that children are less able than adults to recognize words based on partial bottom-up input. Cole and Perfetti [9] reported that adults detect mispronunciations located in the second syllable of words faster than those located in the first syllable, but children do not. This suggests that adults use their lexical knowledge together with a small amount of bottom-up input far more effectively than children do. Gating studies report similar results: Children require more acoustic information to identify a stimulus word than do adults [10,11]. These studies indicate that children are less able to use lexical context to facilitate word recognition. While they do not bear directly on the influence of contextual knowledge on phoneme recognition, they suggest that children might be less likely to use the two sources of information interactively. Still, the relation between context effects on word recognition and on phoneme identification is not clear. A direct measure of whether children are able to use lexical context during phoneme identification is necessary to add to our understanding of the development of mature perceptual processes.

The goal of the present study is to examine how lexical context affects children's identification of phonemes. As discussed above, there are reasons to postulate that lexical context plays a different role for children than adults. It may have more of an influence on children's phoneme identifications because children are poor at segmenting words into component sounds. Alternatively, lexical context may influence children's identifications less if children rely more heavily on the acoustic input to identify the sounds they hear.

EXPERIMENT 1

METHOD

Following a paradigm designed by Ganong [3], children and adults listened to syllables in which the initial phonemes varied along a voice onset time (VOT) continuum. The syllables formed pairs of matched continua. For each pair, one continuum extended from a word to a nonword (e.g. from *desk* to *tesk*) and the other from a nonword to a word (e.g. from *dest* to *test*). For each VOT position, the initial acoustic information in the matched continua was identical. We measured whether the lexical context of a stimulus affected judgments of the identity of the initial phoneme.

Subjects: Twenty 5 - 6 year olds (mean age = 5 years, 10 months; range = 5;5 - 6;9) participated in the experiment. Sixteen of the children were tested at a nearby elementary school while four others, drawn from a local pool of subjects, were tested in our lab. Twenty undergraduates at the University of California at San Diego also participated as control subjects. Only native speakers of English participated in the present study and no subject reported a hearing impairment.

Materials: Each stimulus consisted of one token (100 msec long) from a voice onset time (VOT) continuum spliced onto a phonetic context. Naturally spoken pairs of words and nonwords (e.g. *desk* and *tesk*) were audiorecorded and digitized at 44 KHz using Audiomedia software. The initial 100 msecs of these items were spliced off, yielding a voiced and a voiceless segment. Seven-point VOT continua were formed from the 100 msec segments by replacing approximately 5 msecs of the voiced token with the same amount of aspiration copied from the voiceless token. Each successive token in the continuum was created by replacing a 5 msec segment of the previous stimulus in the continuum with aspiration. The tokens from the VOT continua were then spliced back onto the two different phonetic contexts. For example, each token of the seven-token VOT continuum from /d/ to /t/ was spliced onto the context '*esk*' and onto the context '*est*', resulting in two continua, one extending from *desk* to *tesk* and the other extending from *dest* to *test*. The crucial features of these continua are (a) that at each point along the continua, the initial 100 msecs of the two phonetic contexts are identical, and (b) as the initial phoneme moves from a voiced to a voiceless token, it extends from a word to a nonword in one of the matched continua, and from a nonword to a word in the other.

Eight additional pairs of continua were constructed in the same manner, yielding a total of 126 tokens. All words used were highly familiar to both children and adults as determined by adults' and children's word counts [12,13], a child's first dictionary [14] and by pilot work with young children. The tokens in the pairs of continua began with one of three different stop-consonant contrasts: /b/-/p/, /d/-/t/, or /g/-/k/. Three continua pairs began with each contrast.

Procedure: The tasks for adults and children were similar, but not identical. Below we first describe the adults' task and then detail the modifications to the task for the children.

Task for adults. Subjects were seated in a quiet room and stimuli were presented over headphones by a Macintosh IIci computer equipped with a DigiDesign Audiomedia card and software. On each trial, the subject circled one of two letters (the appropriate voiced and voiceless phonemic contrast) on a response sheet that "best corresponded to the first sound in the word or nonword heard". Subjects were told to make their responses quickly because of the limited amount of time between trials (3 seconds). If the subject heard a phoneme which did not correspond to the choices given, s(he) was instructed to write the perceived letter to the side. No such alternative responses were made. Stimuli were arranged in two randomly ordered lists, each list containing all 126 tokens. Adults listened to each list once, thereby giving two responses to each stimulus. Order of list presentation was counterbalanced.

Task for children. Children were tested in a quiet room and stimuli were presented in the same manner as for the adults. Rather than circle a letter in response, however, children were instructed to repeat exactly the sounds they heard. The experimenter explained that a man would say some things through the headphones and they were "to say exactly what he said." It was made clear that some of the items would sound funny, not like words they knew, but to go ahead and repeat them exactly as the man had said them. Each child then practiced on a training list of nine clear nonwords; if s(he) did not repeat back loudly enough for clear recording, the list was presented again. All children repeated the nonwords back as spoken on the recording. Children then responded to the items in the two test lists. Breaks were given halfway through each list; thus, children heard four blocks of trials lasting approximately 3-1/2 minutes each. Order of list presentation was counterbalanced. Responses were audiorecorded and coded by two raters; interrater agreement was 95.5%.

RESULTS

We compared children's and adults' responses to acoustic information in the word - nonword continua to their responses to the identical acoustic input in the nonword - word continua. Of particular interest were responses at subjects' phoneme boundaries -- that is, the range of voice onset times where perception changes from a voiced to a voiceless phoneme -- because lexical context was expected to have the greatest influence at this point. The location of subjects' phoneme boundaries, however, varies across individuals and continua. Following Ganong [3], all subjects' scores were first aligned on the basis of their phoneme boundaries. This involved locating the phoneme boundary for each subject on each pair of continua and shifting the responses from the phoneme boundaries into position 4, the central position. Recall that at each position along a pair of continua, the subject responded to four token presentations; the position at which the subject reported 50% of these stimuli as voiceless was chosen as the phoneme boundary. Other scores were shifted relative to the phoneme boundary but the endpoints (positions 1 and 7) were never shifted. Scores were then averaged across continua at each position, forming curves which trace out the proportion of voiceless responses to the matched lexical contexts.

Figure 1 shows the proportion of tokens identified as voiceless at each VOT position in the word - nonword continua and in the nonword - word continua, collapsing across all 9 pairs of continua. (Note that in the acoustic input, items in position 1 included clear voiced phonemes, and items in position 7 included clear voiceless phonemes.) Subjects in both age groups identified identical acoustic input differently depending on its context. More tokens were reported as voiceless in the word - nonword continua, while more tokens were perceived as voiced in the nonword - word continua. In each case, this resulted in more tokens perceived as compatible with a word. Differences in identification of the same acoustic information are seen most clearly at the phoneme boundary (position 4) where subjects rated many more stimuli in the nonword-word context as voiceless than in the word-nonword context.

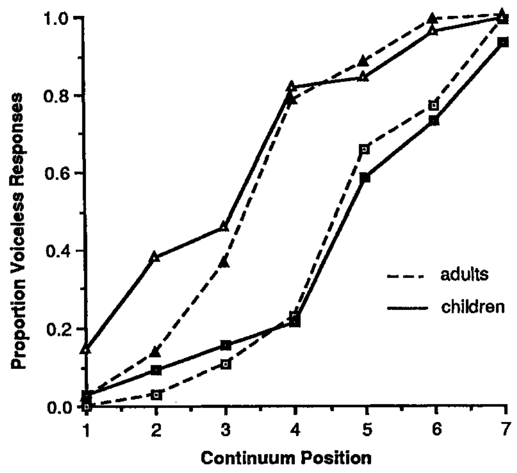

Figure 1. Proportion of voiceless responses by adults and children at each continuum position in word - nonword (squares) and nonword - word (triangles) contexts.

To further analyze the differences between responses in the two continua, we calculated the lexical effect size by subtracting the proportion of voiceless responses to items in each word - nonword continuum from the proportion of voiceless responses to items in the matched nonword - word continuum. We examined the lexical effect size at three points along the continua - the voiced endpoint, the boundary condition, and the voiceless endpoint - for adults and children. The mean lexical effect size for both age groups at these positions can be seen in Table 1. A three-factor mixed Anova was performed using lexical effect size as the dependent variable. Position (3) and Continua pair (9) were the two within-group factors and Age (2) was the between-groups factor. The analysis revealed a main effect of position in the VOT continua (F (2,76) = 205.3, $p < .01$). Pairwise comparisons showed that lexical context had significantly greater impact at the phoneme boundary than at either of the two endpoints. This was true for both adults and children as evidenced by the lack of a significant Age x Position interaction (F (2,76) = .582, $p > .1$).

A main effect of Age revealed that lexical context influenced children's identifications to a greater extent than adults' (F (1,38) = 4.51, $p < .05$). Simple effects comparisons were carried out to determine at which positions adults and children differed. The error term from each individual contrast was used because error was not homogeneous across the three positions; the two endpoint conditions had much lower variance than the boundary condition for both age groups. Simple effects revealed that lexical context had a greater influence on children's interpretations of voiced and voiceless endpoint stimuli than on adults' interpretations (F (1,38) = 11.37, $p < .01$ for voiced endpoints; F (1,38) = 8.42, $p < .01$ for voiceless endpoints). The lexical effect size did not differ significantly between children and adults at the phoneme boundary (F (1,38) = .194, $p > .1$).

Table 1. Mean lexical effect sizes for adults and children at the voiced endpoint, phoneme boundary, and voiceless endpoint.

	Lexical Effect Size		
	Voiced	Boundary	Voiceless
Children	.117	.589	.067
Adult	.022	.558	.008

Effects of the continua pair also appeared, showing that some contrasts resulted in greater lexical effects than others (F (8,304) = 4.06, $p < .01$). In particular, the *gift/kiss* and *bake/page* continua demonstrated particularly large effects. An Age x Continua pair interaction was also present (F (8,304) = 3.44, $p < .01$). Tukey tests revealed that children showed greater lexical effects on the *gift/kiss* contrast and the *bite/pipe* contrast (respectively, F (1,304) = 21.55, $p < .01$; F (1,304) = 4.2, $p < .05$).

DISCUSSION

The larger effect of lexical context at the phoneme boundary than at the continua endpoints indicates that lexical context is likely to have much more influence when the acoustic information does not clearly specify a single phoneme. This was true for both children and adults, indicating that five-year-olds use lexical context much like adults do. Furthermore, the size of the lexical effect at the phoneme boundary was very similar for children and adults, lending additional support to the conclusion that children use lexical context in a similar manner to adults during perceptual processing.

While children and adults showed a similar lexical effect size at the phoneme boundary, lexical context had a greater effect on children's identifications than on adults' at the endpoint positions. In fact, adults rarely showed a lexical effect on any stimuli at the two endpoints. Indeed, their lexical effect sizes at these two points are less than .025 indicating that they reported the clear voiced and voiceless phonemes in the input regardless of the lexical status of their response. Children, on the other hand, occasionally reported words even when the acoustic input supported a nonword response. While the influence of lexical status was small at the endpoints even for children, the difference between children and adults was significant. Two continua pairs were largely responsible for this small, but significant, lexical effect at the endpoints, namely the *gift/kiss* contrast and the *bite/pipe* contrast. Adults' and children's responses differed significantly only on these two contrasts according to pairwise comparisons using a Tukey test.

The larger lexical effect children show at the endpoints of the continua could be due to the nature of the particular stimuli in which the result appeared or to a task bias present for children but not for adults. Perhaps the large number of somewhat ambiguous tokens (i.e., those from within the VOT continua) were confusing for the children and led them to report more words. To test the possibility that the result was due to a task bias, a second experiment was performed in which children responded to a new set of words and nonwords. None of the tokens in the second experiment came from within the VOT continua. That is, all were clearly articulated and unedited. If the original result was due only to a task bias resulting from the large number of ambiguous tokens, children in Experiment 2 should would have no difficulty repeating back the nonwords when they are embedded in the list of other clear tokens. Instead, their responses should look like those of adults in Experiment 1.

EXPERIMENT 2

METHOD

Subjects: Nine kindergartners from a local elementary school participated in the study. All were native speakers of English with no hearing impairments.

Materials: Ninety stimuli were recorded in two randomly ordered lists. The stimuli consisted of two words and two nonwords for each of eight of the continua pair endpoints from the original study, one new set of endpoint items, and 54 additional filler words and nonwords. The filler items were designed to have many qualities similar to the endpoint stimuli; all began with the same set of stop consonants and ended with sounds similar to those in the endpoint stimuli. The lists were recorded on cassette and presented over headphones to children.

Procedure: The task was identical to that in the first experiment. Subjects listened to one of the lists of stimuli and repeated back exactly what they heard. Children performed the task in two blocks, each block lasting approximately 3-1/2 minutes. Responses were recorded on tape and coded.

RESULTS

We calculated the proportion of responses which included an initial phoneme that differed from the input for voiced and voiceless stimulus items for both the words and the nonwords. Table 2 presents the proportion of incorrect phonemes in each of the four conditions. Subtracting the proportion of incorrect responses to words that began with a voiced phoneme from the proportion of incorrect responses to nonwords that began with a voiced phoneme yielded the lexical effect size, analogous to that in experiment 1. Similarly, we calculated the lexical effect size for items beginning with voiceless phonemes. The mean lexical effect size for Experiments 1 and 2 are presented in Table 3.

Table 2. The proportion of incorrect responses to words and nonwords beginning with voiced and voiceless phonemes.

Proportion of Incorrectly Reported Phonemes		
Onset Phoneme	Word	Nonword
Voiced	.006	.062
Voiceless	0.000	.012

Planned comparisons were conducted on the proportion of incorrect responses to words and nonwords with voiced and voiceless onsets. Slightly more phonemes were incorrectly reported in nonwords than in words for items beginning with voiced phonemes (F (1,8) = 5.14, p = .053), but not for items beginning with voiceless phonemes (F (1,8) = 1, p > .1). The same trend was found in Experiment 1 where the lexical effect size was larger at the voiced than the voiceless end of a phoneme continuum. Note, however, that the lexical effect size is smaller in Experiment 2 than in Experiment 1, suggesting that the large number of ambiguous tokens in Experiment 1 contributed to the lexical effect size.

Table 3: Lexical effect sizes for children at voiced and voiceless endpoints in Experiments 1 and 2.

Lexical Effect Size		
Onset Phoneme	Exp 1	Exp 2
Voiced	.117	.056
Voiceless	.067	.012

DISCUSSION

As predicted, the lexical effect size in Experiment 2 is smaller than that in Experiment 1. The lexical effect size for voiced stimuli is half that of children's lexical effect size in Experiment 1 and for voiceless stimuli is roughly one-fifth as large. Thus, the large number of ambiguous stimuli from within the voicing continua in Experiment 1 led children to interpret the phonemes at the beginnings of clear nonwords as consistent with words. With this biasing situation removed in Experiment 2, the lexical effect for nonwords with clearly articulated phonemes diminished, but it did not disappear completely.

The presence of a small lexical effect for endpoint stimuli even in the context of clearly spoken words and nonwords suggests that children use lexical context slightly more than adults do to interpret acoustic input. Children incorrectly reported a marginally significant number of nonword endpoint stimuli as words even in the context of other clearly spoken words and nonwords. Adults in Experiment 1, on the other hand, showed virtually no lexical effect at the acoustic endpoints even in an experimental session that included many other less clear tokens.

Together, Experiments 1 and 2 show that children, like adults, can use lexical context to constrain identification of the component sounds of a syllable. Both children and adults interpret the same acoustic information as two different phonemes when presented in different phonetic contexts. Children in these tasks showed a slightly greater lexical effect than did adults for some clearly articulated items. In Experiment 2, no particular item was responsible for this finding. It is possible, however, that familiarity with certain words or structural properties of the lexicon influences the degree to which children rely on lexical context. These possibilities need further examination.

REFERENCES

[1] A. G. Samuel. "Phonemic Restoration: Insights From a New Methodology," *Journal of Experimental Psychology: General*, vol. 110, pp. 474 - 494, 1981a.

[2] R. M. Warren. "Perceptual Restoration of Missing Speech Sounds," *Science*, vol. 167, pp. 392 - 393, 1970.

[3] W. F. Ganong. "Phonetic Categorization in Auditory Word Perception," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 6(1), pp. 110 - 123, 1980.

[4] S. Nittrouer and A. Boothroyd. "Context Effects in Phoneme and Word Recognition by Young Children and Older Adults," *Journal of the Acoustical Society of America*, vol 87(6), pp. 2705 - 2715, 1990.

[5] B. Fox and D. K. Routh. "Analyzing Spoken Language Into Words, Syllables, and Phonemes: A Developmental Study," *Journal of Psycholinguistic Research*, vol. 4(4), pp. 331 - 342,1975.

[6] R. Treiman and A. M. Breaux. "Common Phoneme and Overall Similarity Relations Among Spoken Syllables: Their Use by Children and Adults," *Journal of Psycholinguistic Research*, vol. 11(6), pp. 569 - 598, 1982.

[7] A. C. Walley, L. B. Smith, and P. W. Jusczyk. "The Role of Phonemes and Syllables in the Perceived Similarity of Speech Sounds for Children," *Memory & Cognition*, vol. 14(3), pp. 220 - 229, 1986.

[8] J. Charles-Luce and P. A. Luce. "Some Structural Properties of Words in Young Children's Lexicons," *Research on Speech Perception*, Rept. 10, pp. 323 - 344, 1984. Indiana University

[9] R. A. Cole and C. A. Perfetti. "Listening for Mispronuciations in a Children's Story: The use of Context by Children and Adults," *Journal of Verbal Learning and Behavior*, vol. 19, pp. 297 - 315, 1980.

[10] L. L. Elliott, M. A. Hammer, and K. E. Evan. "Perception of Gated, Highly Familiar Spoken Monosyllabic Nouns by Children, Teenagers, and Older Adults," *Perception & Psychophysics*, vol. 42(2), pp. 150 - 157,1987.

[11] A. C. Walley. "Spoken Word Recognition by Young Children and Adults," *Cognitive Development*, vol. 3, pp. 145 - 167, 1988.

[12] H. Kucera and W. Francis. *Computational Analysis of present-day American English.* Providence, RI: Brown University Press. 1967.

[13] J. M. Wepman and W. Hass. *A Spoken Word Count: (Children Ages 5, 6, and 7).* Chicago: Language Research Associates, 1969

[14] *Macmillan First Dictionary*. New York: Macmillan Publishing Company. 1990.