Perception & Psychophysics
1977, Vol. 21 (5), 450-454

Perception of syllable-final stop consonants by 2-month-old infants

PETER W. JUSCZYK
Dalhousie University, Halifax, Nova Scotia B3H 4J1, Canada

Previous research has shown that infants are capable of perceiving many phonetic distinctions between initial segments of syllables. The present study demonstrates that 2-month-old infants have the ability to distinguish syllables differing only in their final segments. Infants were found to be sensitive to place-of-articulation differences for stop consonants in final segments of both consonant-vowel-consonant (CVC) and vowel-consonant (VC) syllable pairs. Contrary to previous reports for older infants (Shvachkin, 1973), there was no indication that 2-month-olds have any more difficulty with contrasts of final-stop consonants than they do with initial ones.

Recent investigations have shown the human infant to be capable of discriminating a wide variety of speech sounds differing along such dimensions as voicing, e.g., [ba] vs. [pa] (Eimas, 1975b; Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Lasky, Syrdal-Lasky, & Klein, 1975; Streeter, 1976), place of articulation, e.g., [ba] vs. [da] (Eimas, 1974; Morse, 1972), initial burst cues, e.g., [bu] vs. [gu] (Miller, Morse, & Dorman, in press), and third formant cues, e.g., [ra] vs. [la] (Eimas, 1975a). As a result of these studies with both real and synthetic speech sounds, our knowledge of the basic capacities infants possess for perceiving speech has been greatly increased. Nevertheless, many important questions remain. For instance, just how much of the acoustic information of a word can the young infant process? To date, most studies of infant speech perception have focused largely on the ability to discriminate differences between initial segments of consonant-vowel (CV) syllables (e.g., Eimas, 1974). The issue of whether young infants also are capable of detecting subtle differences between final segments of syllables has not received much attention.

At first glance, it would appear that syllable-final contrasts should pose no particular problem for the infant. After all, if a pair of phones (e.g., [g] and [d]) are discriminable from one another in the syllable-initial position for the infant (Eimas, 1974), then should the same pair not also be discriminable in the syllable-final position? However, Shvachkin (1973) has reported that many older infants (1-year-olds) can distinguish syllables contrasting in initial segments (e.g., l̲ot vs. m̲ot) but not those differing in final segments (e.g., tol̲ vs. tom̲). What could be the source of this difference between syllable-initial and syllable-final contrasts?

There are several reasons why syllable-final contrasts may not be salient for the infant. First of all, it might be the case that syllable-final contrasts are simply less discriminable than syllable-initial ones. However, in research with adult subjects, Malecot (1958) observed that "with releases, the voiced stops b d g in final position are at least as intelligible as they are in initial position." A second possibility for the infant's difficulty with syllable-final contrasts is that he is insensitive to the particular cues for these final distinctions, even though he responds to the cues for the same contrast in initial position. In fact, there are indications that the cues which signal a phonetic distinction in initial positions are not the same as those which mark it in final position. For example, while voice onset time cues play an important role in distinguishing [p] and [b] in initial position (e.g., [pa] vs. [ba]), vowel length is the major cue in final position (e.g., [ap] vs. [ab]). Moreover, using the selective adaptation paradigm, Ades (1974) found that adaptation effects did *not* transfer from the syllable-initial to the syllable-final position or vice versa. This finding suggests that separate mechanisms may be involved in the perception of the same phonetic distinction depending on its location within a syllable. If this is the case, then it is possible that infants might perceive a phonetic distinction in one position but not in the other. Finally, there is also a third possibility for the infant's lack of success with final contrasts. During the initial stages of language learning, the infant may adopt a strategy of attending only to initial segments. Thus, even though he has the necessary perceptual capabilities, the infant may tend to ignore information about final contrasts.

This research was supported by grants to the author from the National Research Council and from the Faculty of Graduate Studies, Dalhousie University. The author wishes to thank Ms. Elizabeth Thompson for her assistance in the collection of the data, and also to express his gratitude to Dr. Alvin M. Liberman for his generosity in making the facilities of the Haskins Laboratories available for the preparation of stimulus materials.

In the present study, we investigate aspects of the perception of syllable-final contrasts by 2-month-olds. We examine whether 2-month-olds can detect final stop consonant differences. Moreover, given Shvachkin's (1973) report that 1-year-olds find syllable-initial contrasts easier than syllable-final ones, we also explore this issue with 2-month-olds. Therefore, in one of our experimental conditions, we arrange a test of the same phonetic contrast in both the syllable initial (dab vs. gab) and syllable-final (bad vs. bag) positions.

Since little is known about the infant's ability to distinguish syllable-final contrasts, we also wish to explore a number of factors which might affect the discriminability of these segments. One such factor is the degree of similarity between the two phones to be discriminated. Thus, a pair of phones differing by only a single phonetic feature may be harder to distinguish than a pair differing by a number of features. For this reason, we include a comparison that involves a number of phonetic features (bag and bam) as well as one that involves a single feature, viz, place of articulation (bad vs. bag). Finally, another factor which may affect the infant's ability to discriminate syllable-final contrasts is the overall degree of similarity of the stimuli. The discrimination of a consonant-vowel-consonant (CVC) pair may be more difficult than a corresponding vowel-consonant (VC) pair because of the additional consonantal similarities incorporated into the former pair. Thus, the present study also compares performance on a CVC pair (bad vs. bag) to that on VC pair (ad vs. ag).

METHOD

Procedure

Each infant was tested individually in a small laboratory room. The infant was placed in a reclining seat which faced a rear-projection screen approximately 2 ft away. An image of a man was displayed on the screen for the entire test session. The projection screen was situated just above a loudspeaker through which the test stimuli were played. Each infant sucked on a blind nipple which was held in place by one of the experimenters who wore headphones and listened to recorded music throughout the test session. A second experimenter, in an adjacent room, monitored the apparatus.

The experimental procedure was a modification of the high-amplitude sucking technique developed by Siqueland and DeLucia (1969). For each infant, the high-amplitude sucking criterion and the baseline rate of high-amplitude sucking were established before the presentation of any stimuli. The criterion for high-amplitude sucking was adjusted to produce rates of 15 to 30 sucks/min. After a baseline rate was established, the presentation of stimuli was made contingent upon the rate of sucking. If the time between each criterion response was 1.4 sec or more, then each criterion response produced one presentation of the stimulus, which had an average duration of 575 msec followed by 825 msec of silence. If the infant produced a burst of criterion responses within this 1.4-sec interval, the timing apparatus was automatically reset and the 1.4-sec interval began again.

The criterion for satiation to the first stimulus was a decrement in sucking rate of 25% or more over 2 consecutive minutes compared to the rate in the immediately preceding minute. At this point, the auditory stimulation was changed without interruption by switching channels on the tape recorder. For infants in the experimental conditions, the change resulted in the presentation of a second acoustically distinct stimulus. For infants in the control condition, the channels on the tape recorder were switched, but no acoustic change was made. The postshift period lasted for 4 min. The infant's sensitivity to the change in auditory stimulation was inferred from comparisons of the response rates of subjects in the experimental and control conditions during the postshift period.

Stimuli

In preparing the synthetic speech stimuli for the present experiment, we decided to employ the released version of all final stops. Release cues for final stops consist of a short burst of acoustic energy following closure. These cues have been shown to be very important in the perception of final stops (Malecot, 1958). Malecot noted that "suppression of releases results in a slight but general dimunition of place identification." Consequently, since our interest here was to determine whether infants possess any ability to distinguish syllable-final stops, we chose to examine those conditions under which the maximum amount of information is available to the infant, i.e., released stops.

Synthetic speech stimuli were prepared at the Haskins Laboratories in the following manner. Actual utterances were recorded and converted into synthetic speech stimuli using the OVE III synthesizer. These synthetic utterances were then digitized using the Haskins PCM system (Cooper & Mattingly, 1969). The output from the PCM system was then used to prepare audio tapes. Six pairs of synthetic speech stimuli were recorded for presentation to the subjects.

Pair 1 (bad-bag). Each stimulus had a total duration of 576 msec. Both stimuli had identical starting frequencies for all three formants as well as identical state values of 702, 1,916, and 2,502 for the first, second, and third formants, respectively. However, the terminal values for the second and third formants of *bad* were 1,695 and 2,557, respectively, while the comparable values for *bag* were 2,045 and 2,345. In addition, both stops were followed 94-msec later by release bursts with durations of 52 msec. Information about the release characteristics of each stop is dis displayed in Table 1.

Pair 2 (bag-bam). The *bag* stimulus was identical to the one described above. Once again, the stimuli were matched for overall duration and the starting frequencies of the first three formants. However, since the vowel in *bam* was nasalized, it had steady state values of 608, 1,916, and 2,413 for the first, second, and third formants, respectively. Moreover, the terminal values for the second and third formants of bam were 1,007 and 2,431, respectively.

Pair 3 (ad-ag). In terms of the important stimulus parameters of steady state and terminal frequencies of all three formants, Pair 3 was identical to Pair 1. The chief difference between the two pairs was that Pair 3 lacked the initial formant transitions present in Pair 1. Moreover, the steady state of Pair 3 was lengthened slightly to bring its overall duration (563 msec) more closely in line with the other pairs. In all other respects, Pair 3 was identical to Pair 1.

Table 1
Frequency (in Hz) Characteristics of the Release Burst of Each Final Stop

Formant		B	D	G	M
First	Onset	329	359	359	303
	Offset	479	541	541	511
Second	Onset	1477	1809	2030	1107
	Offset	1543	1822	2105	1297
Third	Onset	2396	3000	2449	2502
	Offset	2484	2576	2690	2594

Pair 4 (dab-gab). Each stimulus was 576 msec in length. The starting frequencies for the second and third formants of *dab* were 1,744 and 2,594, respectively, while the comparable values for *gab* were 2,016 and 2,912. Both stimuli had steady state values of 702, 1,822, and 2,413 for the first, second, and third formants, respectively. The terminal values for all three formants and the release burst of the final stop were identical for both stimuli.

Pair 5 (bad-bad). Both stimuli were identical in all respects to the *bad* stimulus of Pair 1.

Pair 6 (bag-bag). Both stimuli were identical in all respects to the *bag* stimulus of Pair 1.

Design

All subjects were seen for two sessions. (Mean interval between sessions was 6 days; the range was 2 to 14 days.) In one session, all subjects heard the bad-bag tape. The other sessions differentiates the four groups of subjects. Subjects in Group I, since this session was a control condition, heard either the bad-bad or the bag-bag tape. By contrast, subjects in Group II heard the bag-bam tape, those in Group III the ad-ag tape, and those in Group IV the dab-gab tape. The order of test conditions and the order to stimuli within a session were each counterbalanced.

Apparatus

A blind nipple was connected to a Grass PT5 volumetric pressure transducer which was coupled in turn to a Type DMP-4A physiograph. A Schmitt trigger provided a digital output of criterial high-amplitude sucking responses. Additional equipment included a Sanyo four-track stereo tape recorder with a Sanyo speaker, a Lafayette 100-sec timer, a power supply, two relays, a counter, and a Physiograph dc-ac preamplifier. Each criterion response activated the 100-sec time for a 1.4-sec period or restarted the period. Auditory stimulation at a level of 72 ±2 dB (A) SPL (approximately 15 dB above the background noise level caused by the ventilation system) was available to the infant whenever the timer was in an "active" state.

Subjects

The subjects were 32 infants, 12 males and 20 females. Mean age was 7.2 weeks (range: 6 to 10 weeks). In order to obtain complete data on 32 infants, it was necessary to test 57. Subjects were excluded from the study for the following reasons: falling asleep (30%) or crying (20%) prior to shift, failure to maintain a minimum criterion sucking rate of 10 responses/min during the satiation period (12%), failure to keep the second appointment (30%), and experimenter error (8%).

RESULTS

Figure 1 displays the mean number of high-amplitude sucking responses as a function of minutes and experimental groups. For purposes of statistical comparisons, we examined each subject's rate of high-amplitude sucking during five intervals: baseline minute, third minute before shift, average of Minutes 1 and 2 before shifts, average of Minutes 1 and 2 after shift, and average of all 4 min after shift. Difference scores were calculated for each subject for the following rate of comparisons: (1) acquisition of the sucking response, third minute before shift less baseline; (2) satiation, third minute before shift less average of last 2 min before shift; (3) release from satiation, average of first 2 min after shift less average of last 2 min before shift; and (4) release from satiation for full 4 min after shift, average of the 4 min after shift less average of last 2 min before shift.

Wilcoxon matched-pairs signed-ranks tests (Siegel,

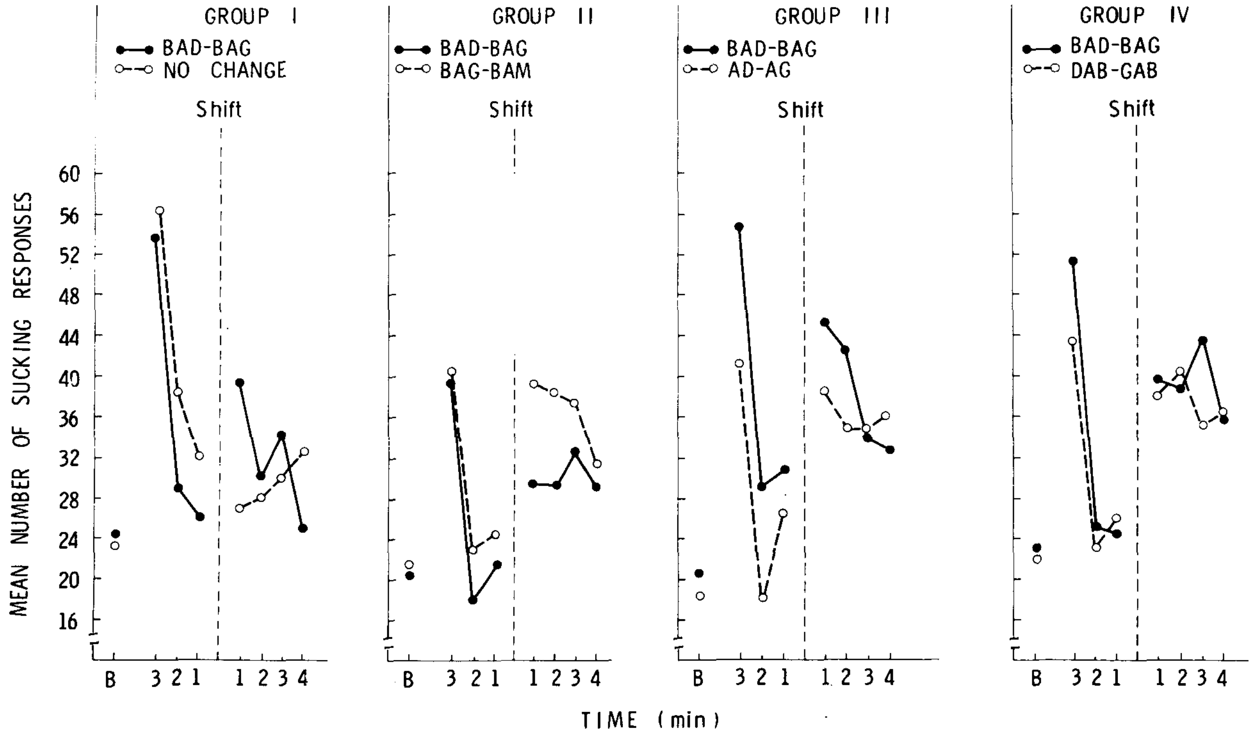

Figure 1. Mean number of high-amplitude sucking responses as a function of time and experimental group. Time is measured with reference to the moment of the stimulus shift, marked by the vertical dashed line. The baseline rate of sucking is indicated by the letter "B."

1956) were used to analyze performance within each type of session. In all sessions, subjects acquired the conditioned high-amplitude sucking response and satiated to the first stimulus prior to shift. This was not surprising, since subjects were selected for having met these criteria. The result of the analyses for subjects' performance after shift are presented in Table 2. Subjects in all experimental sessions displayed a reliable increase in sucking after shift. Moreover, the release from satiation exhibited by subjects in the experimental sessions was reliable for the full 4-min period after shift. By contrast, subjects in the control session gave evidence of a significant decrease in sucking for both the average of the first 2-min and the full 4-min periods after shift. Further inspection of the data by Kruskal-Wallis one way analyses of variance (Siegel, 1956) confirmed that there were no reliable differences between the BAD-BAG sessions of Groups I, II, III, and IV [$\chi^2(3)$ ranged from 1.12 to 5.12].

Comparisons of performance across sessions for subjects in each group were made with randomization tests for matched samples (Siegel, 1956). Only Group I subjets displayed any reliable tendency to respond more after shift in one condition as opposed to the other one. Group I subjects were shown to have significantly higher response rates for both the first 2-min ($p < .01$) and full 4-min ($p < .05$) periods after shift in the BAD-BAG condition than they did in the control condition. For Groups II, III, and IV, there was no indication that subjects performed reliably better in one condition than in the other.

Table 2
T Values for Wilcoxin Matched-Pairs Signed-Ranks Test

	Release from Satiation First 2 min After Shift			Release from Satiation Full 4 min After Shift		
	T	Mean	SD	T	Mean	SD
Group I						
Bad-Bad	2*	7.4	6.09	1*	5.1	4.39
Control	1*	−7.6	8.55	1*	−5.0	5.92
Group II						
Bad-Bag	3*	9.3	9.82	0**	10.2	9.70
Bag-Bam	0**	14.7	6.83	0**	13.0	6.70
Group III						
Bad-Bag	0**	14.4	10.90	2*	10.2	9.77
Ag-Ad	0**	14.6	10.27	2*	14.0	12.99
Group IV						
Bag-Bad	2*	15.8	13.24	3*	14.7	15.20
Dab-Gab	2*	13.9	12.72	0**	13.1	10.21

*$p < .05$ **$p < .01$

DISCUSSION

The present study demonstrates that infants as young as 2 months old can discriminate syllable-final stops. There was no evidence to suggest that 2-month-olds perceive initial-stop differences (d̲ab vs. g̲ab) any better than they do final-stop differences (bad̲ vs. bag). Thus, the present findings stand in contrast to Shvachkin's (1973) but parallel those of Malecot (1958), who reported that adults were as adept at identifying *released* final-stops as they were at identifying initial ones. Moreover, in the present study, it was also observed that a contrast by a single phonetic feature (bad vs. bag) was as detectable for infants as one by several features (bag vs. bam). Additionally, there was no indication that infants found a CVC pair (bad vs. bag) any more difficult to distinguish than a corresponding VC pair (ad vs. ag). Consequently, the presence of an additional consonant similarity incorporated into the former pair did not appear to affect the infant's ability to detect final-stop differences.

The failure to find differences between sessions for any of the experimental groups other than Group I (where performance during the "bad-bag" condition exceeded that for the control condition) deserves comment. First, it should be noted that there was a considerable amount of between-subject variance for each of these groups. Second, while the magnitude of recovery after shift is similar to that observed by others (e.g., Eimas, 1974; 1975a; Eimas, Siqueland, Jusczyk, & Vigorito, 1971), it is possible that a ceiling effect has occured. Perhaps, it is unrealistic to expect that infants will greet a change from [bag] to [bam] with any more enthusiasm (as reflected in their sucking rates) than they will a change from [bag] to [bad]. After all, in both cases, a shift is made to a novel stimulus. Thus, it may be that the infant notices that the magnitude of change is greater in one case than in the other, but that the high-amplitude sucking procedure is not sensitive enough to indicate it.

By showing that 2-month-olds perceive place-of-articulation differences for initial stops in CVC pairs, the present results replicate those previously observed with CV pairs (Eimas, 1974; Miller et al., in press; Morse, 1972). More importantly, the present study extends previous developmental findings by demonstrating that infants are sensitive to place-of-articulation information in both syllable-initial *and* syllable-final positions.

If 2-month-olds have the perceptual capacity to distinguish syllable-final stops, why did Shvachkin (1948) find that many 1-year-olds have difficulty with these contrasts? One possibility is that Shvachkin's measure, which required children to distinguish objects according to their names, was not as sensitive as the present technique. A second possibility resides in the strategies which his subjects adopted. Although Shvachkin's subjects had the capacity to detect final contrasts, they may have been attending only to initial segments of the words. Indeed, Shvachkin may have biased his subjects in this direction, since evidently most of his test items involved initial contrasts. It would be interesting to know whether Shvachkin's findings are merely an artifact of his procedure or if they are indicative of a general strategy of 1-year-olds to attend to initial segments when first learning words.[1] While the present research cannot tell us about what strategies 1-year-olds are employing, it, and other studies of infant speech perception, help lay the foundation for answering this question. Certainly, the nature of the infant's speech perception capacity constrains the kinds of strategies he can use to learn words.

REFERENCES

Ades, A. E. How phonetic is selective adaptation? Experiment on syllable position and vowel environment. *Perception & Psychophysics*, 1974, **16**, 61-66.

Branigan, G. Syllable structure and the acquisition of consonants. *Journal of Psycholinguistic Research*, 1976, **5**, 117-133.

Cooper, F. S., & Mattingly, I. G. A computer-controlled PCM system for the investigation of dichotic speech perception. *Journal of the Acoustical Society of America*, 1969, **46**, 115(A).

Eimas, P. D. Auditory and linguistic processing of cues for place of articulation by infants. *Perception & Psychophysics*, 1974, **16**, 513-521.

Eimas, P. D. Auditory and phonetic coding of the cues for speech: Discrimination of the [r-l] distinction by young infants. *Perception & Psychophysics*, 1975, **18**, 341-347. (a)

Eimas, P. D. Speech perception in early infancy. In L. B. Cohen & P. Salapatek (Eds.), *Infant perception*. New York: Academic Press. Pp. 1975. (b)

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. Speech perception in infants. *Science*, 1971, **171**, 303-306.

Lasky, R. E., Syrdal-Lasky, A., & Klein, R. E. VOT discrimination by four and six and a half month old infants from Spanish environments. *Journal of Experimental Child Psychology*, 1975, **20**, 215-225.

Malecot, A. The role of releases in the identification of released final stops. *Language*, 1958, **34**, 370-380.

Miller, C. L., Morse, P. A., & Dorman, M. F. Cardiac indices of infant speech perception: Orienting and burst discrimination. *Quarterly Journal of Experimental Psychology*, in press.

Morse, P. A. The discrimination of speech and nonspeech stimuli in early infancy. *Journal of Experimental Child Psychology*, 1972, **14**, 477-492.

Moskowitz, A. I. The two-year-old stage in the acquisition of English phonology. *Language*, 1970, **46**, 426-441.

Shvachkin, N. K. The development of phonemic speech perception in early childhood. Reprinted in C. A. Ferguson & D. I. Slobin (Eds.), *Studies in child language development*. New York: Holt, Rinehart, & Winston, 1973.

Siegel, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.

Siqueland, E. R., & DeLucia, C. A. Visual reinforcement of nonnutritive sucking in human infants. *Science*, 1969, **165**, 1144-1146.

Streeter, L. A. Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. *Nature*, 1976, **259**, 39-41.

NOTE

1. Studies of the young child's first productions of words support the notion that initial segments may be learned prior to final segments. A number of researchers (e.g., Branigan, 1976; Moskowitz, 1970) have commented on the stability of initial segments as opposed to final ones in the young child's first words.

(Received for publication September 20, 1976;
revision accepted January 25, 1977.)