Perception & Psychophysics
1987, 42 (2), 150-157

Perception of gated, highly familiar spoken monosyllabic nouns by children, teenagers, and older adults

LOIS L. ELLIOTT, MICHAEL A. HAMMER, and KARIN E. EVAN
Northwestern University, Evanston, Illinois

A forward-gating procedure, employing highly familiar monosyllabic words, was used in testing 5-7-year-old children, 15-17-year-old teenagers, and 70-85-year-old adults. Teenagers identified the words at shorter gate durations than either the children or older adults, whose identification performances were nearly identical. Teenagers gave meaningful guesses at shorter durations than children, who, in turn, gave meaningful guesses at shorter durations than adults. The oldest listeners provided the largest number of phonetic guesses, whereas teenagers gave almost none. Individual differences in auditory pure-tone sensitivity did not account for the results. It is hypothesized that both word frequency effects and temporal processing differences were responsible for the findings.

In a study of speech perception of monosyllabic nouns, Elliott et al. (1979) found that 5- and 6-year-old children required higher intensity levels to identify words presented in quiet than did older children or adults, even though the stimuli were within the receptive vocabularies of 3-year-olds. Since an adaptive threshold-tracking procedure was used, stimuli were always presented at levels close to threshold. Elliott et al. suggested that listeners with higher levels of language skill might be more adept at identifying words from partial or limited acoustic information than subjects with less mature language development.

The gating paradigm developed by Grosjean (1980) as a means of examining spoken-word recognition appears to offer a good means of further examining developmental changes in word recognition. In the gating paradigm, portions of words are presented, beginning at either the word onset (Cotton & Grosjean, 1984; Grosjean, 1980) or the word ending (Salasoo & Pisoni, 1985). For example, if 50-msec gates are used and forward gating is employed, one stimulus contains the initial 50 msec of a word, another contains the initial 100 msec, and so on. The listener's task is to identify the word, and this may require considerable guessing when the stimulus duration is brief. Grosjean (1980) used the gating paradigm to replicate influences of word frequency, word length, and sentence context on speech perception that were similar to those that had been reported by others.

The gating paradigm provides one means of presenting limited acoustic information about a stimulus, since the first gates contain only the initial, incomplete portion of the word. Presenting speech stimuli at low-intensity levels, as was done by Elliott et al. (1979), constitutes a different approach to providing limited acoustic information. In the latter case, only phonemes with greatest intensity (e.g., vowels) may be audible at low levels.

Walley (1984) used the gating paradigm to study word identification performance of preschool-aged children and college students. Her data (truncated, forward-gated condition) demonstrated that young children require a longer "gate" (i.e., more acoustic information) for word identification than do college students. She attributed much of the age effect to differences in the elimination of inappropriate guesses. The words used by Walley, however, although reported to be known to a comparable group of young preschoolers, were multisyllabic and far more sophisticated than the monosyllabic words used by Elliott et al. (1979). For example, she included "delicious," "telescope," "uniform," and "dinosaur," whereas the Elliott et al. (1979) stimuli were words such as "dog" and "milk." (Also, unlike Grosjean and others—e.g., Salasoo & Pisoni, 1985—Walley gated words at phoneme boundaries.) Thus, one purpose of this investigation was to compare performances of subjects of different ages on the gating task, using words that were *very* well known by young children, in order to determine whether age differences in correct identification of these familiar stimuli occurred with the gating paradigm as well as with the threshold-tracking task.

It has been well documented that very frequently occurring words are more readily perceived than less frequently occurring stimuli (e.g., Broadbent, 1967; Savin, 1963). Furthermore, Elliott, Clifton, and Servi (1983) demonstrated that young adults and young children with higher language levels were more influenced by word frequency effects than were young children with lower language levels.

If age effects on the gating task were governed exclusively by number of years of experience in listening to

This research was supported, in part, by Grant NS 18556 from NIH (NINCDS). Appreciation is expressed to Jan Wasowicz for preparing the gated stimuli and commenting on an earlier version of the manuscript.
The authors' mailing address is: Audiology Program, Northwestern University, 2299 Sheridan Road, Evanston, IL 60208.

frequently occurring words, one would expect older adults to identify words at shorter gates than teenagers, who in turn would perform better than young children. There is evidence, however, that some types of auditory performance that improve with increasing age, from childhood until the preadolescent or teenaged years, neither continue to improve throughout adulthood nor remain stable, but decline (Elliott, Busse, & Bailet, 1985; Elliott, Busse, Partridge, Rupert, & DeGraaff, 1986; Elliott, Longinotti, Meyer, Raz, & Zucker, 1981; Gordon-Salant, 1986; Patterson, Nimmo-Smith, Weber, & Milroy, 1982). Older adults demonstrate special difficulty in processing brief or temporally distorted auditory signals (e.g., McCroskey & Kasten, 1982). The second purpose of this work, therefore, was to examine older adults' performance on the gating task, with special interest in the question of whether their increased years of experience with the stimulus items enabled them to perform as well as, or better than, the younger subjects.

Most studies that have employed the gating paradigm have placed the gated stimuli in the context of a sentence—at least for some conditions (Grosjean, 1980; Salasoo & Pisoni, 1985). This approach was not adopted in the present study because previous research had demonstrated major age differences in the ability of young children to use sentence context to predict a highly familiar, final word (Elliott, Clifton, & Ferre, 1981). Since Cotton and Grosjean (1984) had demonstrated that when no context is used, no difference in results occurs between successively presented and individually presented gated stimuli, successive presentations were used (i.e., trials for each word began with the briefest gate of the word and continued with increasingly longer gates). For testing children, this had the special advantage of helping them understand the "rules of the game."

As in Grosjean's (1980) original work, the listener was required to give a confidence rating for each response. Grosjean proposed that there are two perceptual stages encountered when gates of a word become successively longer. The first stage occurs when the listener "isolates" an appropriate candidate. This is the point at which the subject actually reports the word correctly but lacks confidence in his or her response. The latter stage occurs when the listener achieves confidence in the response. Grosjean (1980) termed the former stage the "isolation point" and the latter, the "recognition point." The name of the latter was later changed to "total acceptance point" (Grosjean, 1985). The combination of listeners' word responses and confidence ratings permits measurement of both these points.

METHOD

Subjects

Children, teenagers, and older adults were recruited from the local area; they participated for payment, with 20 subjects per group. Age ranges were: for the children, 5 years 7 months to 7 years 0 months (mean = 6 years 3 months); for the teenagers, 15 years 2 months to 17 years 9 months (mean = 16 years 7 months); and for the older adults, 69 years 8 months to 84 years 0 months (mean = 75 years 9 months). Thirteen females participated in the children's group, 10 in the teenaged group, and 11 in the older adult group.[1] Unless the subject preferred the left hand and used the left ear when telephoning, the right ear was the test ear. One person per group was tested in the left ear.

Several procedures were administered to potential subjects in order to control for possible auditory or language problems. Conventional air conduction thresholds were obtained in the test ear at the octave frequencies from 500 through 4000 Hz; all subjects were selected to have auditory sensitivity equal to or better than levels expected for their age group (Table 1). Pure-tone sensitivity screening was conducted in the nontest ear at 20 dB HL (re ANSI, 1970). Conventional audiologic speech reception thresholds were obtained using children's or adults' spondee words. All subjects were required to have normal tympanograms in the test ear. Children and teenagers were also administered the Peabody Picture Vocabulary Test-Revised (PPVT-R, Dunn & Dunn, 1981) and were required to score approximately at age level or above. This was not done for the older adults because of concern that they might take offense and because of certainty that they would be thoroughly familiar with the test vocabulary.

In addition to the 20 subjects per group who completed all procedures, 8 additional subjects were eliminated for various reasons: 3 older adults exhibited auditory sensitivity that was poorer than expected for their ages; 1 teenager scored below age level on the PPVT-R; 1 child had an abnormal tympanogram and was referred for medical treatment of the middle-ear problem; 1 child exhibited severe articulation disorders and was receiving speech therapy; 1 child scored extraordinarily poorly on the experimental task and, when tested on a different, syllable-discrimination task (Elliott et al., 1986), demonstrated abnormal performance; and 1 child complained that she did not like the task.

Stimuli

Forty monosyllabic nouns representing concrete objects and selected from the NU-CHIPS Test (Elliott & Katz, 1980b, 1980c)

Table 1
Test-Ear Average Pure-Tone Thresholds and Speech Reception Thresholds in Decibel Hearing Level

	Children	Teenagers	Older Adults
		500 Hz	
M	4.3	0.8	10.8
SD	5.4	6.1	10.0
Range	−5–15	−5–15	−5–30
		1kHz	
M	0.8	−1.8	12.2
SD	5.2	3.4	12.1
Range	−5–10	−5–5	−5–30
		2 kHz	
M	4.2	1.2	23.2
SD	5.2	2.2	14.3
Range	0–15	0–5	0–60
		4 kHz	
M	5.5	3.8	43.5
SD	6.7	4.6	15.4
Range	0–25	0–15	15–70
		Speech Reception Threshold	
M	8.8	5.0	20.0
SD	6.0	4.3	10.3
Range	0–25	0–15	5–40

Note—Negative thresholds denote thresholds that are better than those of the average non-noise-exposed young adult.

were used as stimuli; two additional monosyllabic nouns from the same instrument served as practice items. NU-CHIPS words were originally developed to be within the receptive repertoires of 3-year-old inner-city children and were derived from a much larger pool of several hundred words that appeared in children's books and in word lists for children.[2] Thus, these nouns may not be considered representative of all English monosyllabic words. The male-talker NU-CHIPS recording was used to derive the gated stimuli.[3]

Since young children were to be tested, it was important to keep test time as short as possible. Therefore, the minimum duration for each monosyllabic stimulus was 120 msec. Each additional gate was 30 msec. Since the shortest word was 390 msec (and the longest, 720 msec; mean word length = 513 msec), the shortest gate was less than one-third the total duration of the shortest stimulus. The 120-msec gate was less than half the mean isolation point reported by Grosjean (1980) for high-frequency monosyllabic words in the no-context condition. Thus, it was assumed (and confirmed in pilot testing) that few listeners would correctly identify any of the words when presented with only the first 120-msec gate.

Each word was low-pass filtered at 4 kHz (Krohn-Hite Model 3343 filter), digitized using a 12-bit analog-to-digital converter with a 5-V dynamic range, sampled at 10 kHz, and stored using a PDP-11/34 computer. Root-mean square intensity was normalized at 72.5 dB SPL for all stimuli. Then, the initial interval of 120-msec from word onset, and subsequent intervals of 30 msec, were marked at the nearest zero crossing to avoid the production of auditory clicks. An interactive waveform editing program was used to locate the interval markings and to truncate each word. When word durations were not an even multiple of 30 msec, the final interval was the difference between the location of the last marked interval and the word ending. Thus, each word was represented by a set of intervals, the size of the set depending on the duration of the word. These were combined into a group of presentations that all shared the same initial 120-msec interval and that then had increasing durations. All gated stimuli were converted from digital to analog form, low-pass filtered at 4 kHz, and recorded onto audiotape for subsequent testing. A 5-sec interval separated each presentation of the same word to allow time for the subject to respond and for the experimenter to record the response. In addition, a 5-msec 1-kHz tone preceded the first (i.e., briefest) presentation of each word, alerting the subject as well as the experimenter to a new stimulus item. The 40 test words were presented in random order (but in the same order to all listeners), with all presentations of the same word completed before the next word began. The practice words were gated in the same manner as the test words.

Two 3×20 in. cardboard-mounted rating scales contained numbers from 1 to 7. For teenagers and adults, the word *confident* was located under "7" and *not confident* was under "1." The rating scale used in testing children contained a diagram of a smiling face and the word *yes* at the "7" end. A diagram of a straight-mouthed face and three question marks were at the "1" end.

Experimental Task

Some of the instructions were identical for all subjects; others were modified according to the listener's age. The experimenter began by saying:

> I am going to play some words for you. They are all words that you know. Your job will be to tell me, first, what you think each word is and, second, how sure you are about that.
>
> You will hear each word several times. In the beginning, it may be hard to decide what the word is, but you *must* give an answer. Then you must also tell me whether or not you think the word you said is right. You should do that by:
>
> (for older adults and teenagers:) calling out the number along this line that represents your confidence in your response. For example, if you think your answer was completely a guess, the number you say should be closer to 1 or 2 (pointing). However, if you are pretty certain you were right, the number you give should be closer to 6 or 7 (pointing). If you are halfway confident about your answer, you should say "4," which is at the middle of the line. Use only whole numbers—for example, do not say 6½.
>
> (for children:) pointing to the place along this line that represents how sure you are. Pointing closer to this end of the line (demonstrate) means that you are sure your answer was right. But, pointing closer to this end (demonstrate) means your answer was a guess. If you are halfway sure and halfway guessing about a word, then you should point to the middle of the line. Always point to a place along the line where there is one of these marks (i.e., numbers). Remember—toward this end (pointing) you are *sure* of your answer and toward *this* end you are not sure.
>
> We will start with a set of practice trials so you can see how this works. (The first practice trial was begun and stopped as needed to reinstruct and answer questions, etc.) Do you have any questions? (If not, the second practice trial was begun.) This is another practice run. (If the subject had no difficulty with the second trial, the experimenter proceeded immediately to the test trials. If the subject needed assistance, the points that seemed to need emphasis were reviewed.)
>
> O.K., we are ready to begin. Remember to try to guess the word as soon as you can.

The experimenter recorded two items in response to each stimulus—the word or partial word and the listener's confidence. If the subject paused too long before responding, the tape recorder was turned off and the response was requested. When the response was not a complete word, it was recorded phonetically. When the experimenter was uncertain whether the response was a word or a nonword, "What does ______ mean?" was asked. Oral responses were required because it was not possible for many of the children to write their responses. Subjects' verbalizations were not tape-recorded because young children sometimes direct more attention to the recorded sound of their voices than to the task. General encouragement, but no direct feedback, was given.

Gated stimuli were presented via headphones at 30-dB sensation level (SL) re each listener's speech reception threshold. (This procedure took account of the higher levels required for older adults' speech recognition thresholds and served to "equate" listening levels for all subjects.) Testing was conducted for the teenagers and the older adults in one 2-h session that included a break. Children reported for two test sessions; the tympanogram and the two practice trials were repeated at the beginning of the second test session.

All testing was conducted in a double-walled, sound-treated chamber. A Qualitone Acoustic Appraiser pure-tone and speech audiometer was used for the auditory test procedures.

Data Analyses

Terminology initiated by Grosjean (1980, 1985) was followed. That is, the "isolation point" (IP) was the word duration at which the subject first correctly[4] reported the stimulus without subsequently changing his or her response. The IP confidence rating (CIP) was the subject's confidence judgment at this duration. The "total acceptance point" (TAP) was the duration at which the subject first correctly identified the word *and* first gave a confidence value of 7 without subsequently giving a lower confidence rating or reporting another word.

Occasionally a subject did not succeed in identifying a word, even at the longest gate (this event was also observed by Grosjean, 1980). In this instance, 30 msec was added to the duration of the longest stimulus of the set and the resulting value was used in statistical analyses. A confidence level of 0 was used in calculating CIP. When the word was never identified with a confidence level of 7, 30 msec was added to the duration of the longest stimulus presented and used in statistical analyses of TAP. (Note: A precedent for these procedures derives from clinical audiology where average hearing levels are routinely obtained in cases of failure to respond at specific tonal frequencies by using the maximum audiometer output at that frequency plus 5 dB. This approach is conservative in that it assumes the subject would isolate or recognize the stimulus if only one more gate were presented—an outcome that might not occur.) Because not every word was identified by every subject, the percentage of words correctly identified at the longest duration and the percentage of words that were correctly identified at the highest confidence level (i.e., percentage "accepted") were also recorded.

RESULTS

Word Identification

The average percentages of words that were correctly identified, at least at the longest duration, are shown in Table 2. Teenagers, on the average, identified 7.8% (or approximately three) more words than children and 6.7% more words than older adults. These differences were statistically significant [for arcsin-transformed values, $F(2,57) = 17.70$, $p < .001$]. Teenagers also identified more words at the highest level of confidence (i.e., "words accepted") than children or older adults [$F(2,57) = 5.95$, $p < .01$]. Although there were some differences, the most important feature of Table 2 is its demonstration that most listeners identified most of the stimuli, providing they were sufficiently long. Thus, differences in the IP times for different subject groups were not determined by differences in knowledge of the words.

Table 2
Average Percentages of Words Correctly Identified and of Words Accepted with Highest Confidence

	Children	Teenagers	Older Adults
		Words Identified	
M	89.0	96.8	90.1
SD	6.1	3.6	7.1
Range	75-97.5	87.5-100.0	72.5-97.5
		Words Accepted	
M	80.9	90.6	73.9
SD	20.3	14.7	21.1
Range	2.5-97.5	32.5-100.0	20.0-95.0

Isolation Points

The average IP for older adults was about 3 msec longer than the mean IP for children (Table 3), but both children and older adults had, on the average, approximately 70-msec-longer mean IPs than did teenagers. A mixed-model ANOVA demonstrated significant age effects for IP [$F(2,57) = 23.61$, $p < .001$]. Newman-Keuls tests demonstrated that the IPs for children and older adults were both significantly longer than the IPs for teenagers ($p < .01$). Effects of words were significant[5] [$F(39,2223) = 45.92$, $p < .001$], as was the age × word interaction [$F(78,2223) = 2.65$, $p < .001$].

Significant age effects occurred for confidence ratings at the IP [$F(2,57) = 21.32$, $p < .001$]. Mean confidence ratings at IP were identical for teenagers and older adults and lower than the average ratings for children ($p < .01$). Word effects were also significant [$F(39,2223) = 6.34$, $p < .001$], as was the age × word interaction [$F(78,2223) = 2.59$, $p < .001$].

Total Acceptance Points

Nearly half of the older adults "accepted" (i.e., identified with highest confidence) 80% or fewer of the stimuli. Thirty percent of the children and only 10% of the teenagers "accepted" 80% or fewer stimuli. When *average* percentages of words accepted by each listener were examined, children and teenagers performed similarly (Table 3). Age effects for TAPs were significant [$F(2,57) = 6.95$, $p < .01$], but both children and teenagers attained the highest level of confidence for stimuli of nearly equal durations. The older adults required stimuli that were nearly 50 msec longer (Table 3) than those required by the other two groups before they attained the highest level of confidence ($p < .01$, Newman-Keuls tests). This is a conservative estimate because the older adults awarded the highest confidence rating to fewer words than did the other subject groups (Table 2). Both word effects [$F(39,2223) = 58.61$, $p < .001$] and the age × word interaction [$F(78,2223) = 3.13$, $p < .001$] were also significant.

Isolation and Identification of Individual Words

Large between-words differences occurred in the IP durations. For children, the average IPs ranged from 208.5 msec for "boat" to 531.0 msec for "spoon." Table 4 lists all words in the order of children's average IPs, and gives mean IPs for the other two listening groups. Although some irregularities occurred, there were many similarities in the patterns of mean IPs. For example, teenagers had longer mean IPs than children for only three words—"duck," "soap," and "bike." Older adults, in contrast, whose overall mean IP was nearly identical to children's, had longer IPs than children for 17, or nearly

Table 3
Average Group Isolation Points, Confidence Levels at Isolation and Total Acceptance Points,* and Ranges of Individual Mean Isolation Points, Confidence Levels, and Total Acceptance Points

	Children	Teenagers	Older Adults
		Isolation Point†	
M	343.7	272.5	346.8
SD	38.7	34.5	42.4
Range	288-438	218-339	297-445
		Confidence at IP	
M	4.7	2.9	2.9
SD	1.2	0.9	1.0
Range	0.9-6.0	1.4-4.6	1.1-4.7
		Total Acceptance Point†	
M	403.0	401.2	450.8
SD	52.2	48.7	41.6
Range	318-542	304-524	392-527

*Mean word length was 513.0±77.7 msec (range 390-720 msec). †In milliseconds.

Table 4
Mean Isolation Points (in Milliseconds) for Three Subject Groups

Words	Children	Teenagers	Older Adults
boat	208.5	171.0	189.0
truck	216.0	199.5	306.0
duck	234.0	243.0	292.5
teeth	237.0	213.0	285.0
foot	238.5	237.0	241.5
dog	250.5	228.0	259.5
door	250.5	138.0	244.5
ball	253.5	177.0	292.5
cup	255.0	228.0	238.5
milk	264.0	252.0	331.5
girl	270.0	159.0	264.0
soap	270.0	277.5	283.5
food	276.0	139.5	264.0
bike	280.5	282.0	351.0
bird	288.0	246.0	306.0
bear	306.0	216.0	232.5
shoe	327.0	282.0	357.0
meat	336.0	226.5	270.0
bus	337.5	265.5	339.0
head	337.5	288.0	319.5
school	343.5	298.5	390.0
frog	346.5	237.0	421.5
snake	349.5	249.5	321.0
hair	355.5	277.5	333.0
tongue	360.0	282.0	382.5
mouth	369.0	283.5	354.0
tree	370.5	208.5	360.0
cake	372.0	363.0	351.0
clock	376.5	255.0	355.5
house	376.5	285.0	361.5
dress	379.5	240.0	367.5
shirt	415.5	352.5	382.5
sink	438.0	417.0	435.0
nose	460.5	453.0	525.0
horse	465.0	445.5	535.0
man	465.0	340.5	400.5
smile	475.5	327.0	469.5
train	519.0	351.0	471.0
hand	526.5	369.0	478.5
spoon	531.0	394.5	511.5

Note—Words are in rank order of children's mean isolation points.

half, of the words ("truck," "duck," "teeth," "foot," "dog," "ball," "milk," "soap," "bike," "bird," "shoe," "bus," "school," "frog," "tongue," "nose," and "horse"). "Soap," "duck," and "bike" were the only words for which the children's mean IP was shorter than both teenagers' and older adults'; however, the only case in which the children's IP was *substantially* shorter than the other groups' was "duck," for which children's mean IP was nearly 60 msec shorter than the older adults'.

The correlations between the IP and each of the four measures of pure-tone sensitivity and the speech reception threshold (SRT) across the total group of listeners ($N = 60$) were calculated for each word.[6] For some words, the IPs were significantly related to the subjects' auditory characteristics, but for other words they were not. Furthermore, the lack of any clear pattern suggested that phonemic characteristics of the words did not explicitly determine whether listeners' auditory sensitivity would control performance. For example, the IP for "smile" was significantly related (at $p < .01$) to all four auditory measures, whereas the IP for "sink" was related to none. Both words begin with a sibilant. Also, if auditory sensitivity had been a strongly determining factor, the greatest number of significant correlations with IP should have occured for the pure-tone threshold at 4 kHz, since that is the frequency at which older adults are most likely to have hearing problems. Instead, there were 20 significant correlations at this frequency, whereas there were 23 and 24 for 1 and 2 kHz, respectively.

Durations at which the first meaningful response was given (as opposed to a phonetic imitation of part of the word, a nonsense word, or a failure to respond) were analyzed. Means and standard deviations were 131.0 ± 18.8 msec, 120.6 ± 1.7 msec, and 148.7 ± 18.4 msec for children, teenagers, and older adults, respectively. With few exceptions, the teenagers' first meaningful responses occurred on the first gate (120 msec), whereas the older adults' occurred, on the average, at about the second gate (150 msec). A Kruskal-Wallis rank-order ANOVA showed significant age differences for the three population groups [$H(2) = 17.66, p < .001$]. Also, a t test revealed that the average duration at which the first meaningful word response occurred for children was significantly shorter than for older adults [$t(38) = 3.0, p < .01$].

Table 5 reveals that the majority of incorrect responses were meaningful words. The older adults gave the largest average number of different nonword guesses; teenagers gave almost none. The differences in incorrect guesses often seemed to occur because, when subjects were uncertain, older adults responded by repeating whatever part of the stimulus they had heard (i.e., a "phonetic response"), teenagers produced a meaningful word that incorporated some acoustic characteristic of the stimulus, and children responded with a word that may or may not have shared phonetic characteristics of the stimulus. Rank-order ANOVAs revealed significant age effects for numbers of meaningful words [$H(2) = 6.26, p < .05$] and nonwords [$H(2) = 31.70, p < .001$] guessed. Furthermore, t tests showed that the older adults guessed significantly more different nonwords (but not more different meaningful words) than the children [$t(38) = 3.0, p < .01$].

Table 6 displays correlations between individual mean IPs, average pure-tone sensitivity (.5, 1, 2, and 4 kHz), mean durations of first meaningful responses, and mean numbers of different *incorrect* responses—meaningful and nonwords. Note that mean numbers of different incorrect *meaningful* responses were unrelated to the mean duration of first meaningful responses and to average pure-tone sensitivity. A multiple regression procedure to predict mean IPs resulted in $R = .86$ ($p < .001$), with mean numbers of different incorrect meaningful responses ($p < .001$) and mean duration at first meaningful response ($p < .05$) contributing significantly. Mean

Table 5
Average Numbers of Different Incorrect Responses

	Children		Teenagers		Older Adults	
	M	*SD*	*M*	*SD*	*M*	*SD*
Meaningful Words	2.6	1.1	1.9	.4	2.3	.4
Nonwords	.3	.3	.03*	.04*	.7	.5
Total	2.9	1.1	1.9	.4	3.0	.6

*Values given in hundredths because so few incorrect nonword responses occurred.

numbers of different incorrect nonword responses and average pure-tone sensitivity made no independent contributions.

DISCUSSION

Outcomes of this study in some ways replicate and in other ways differ from the results obtained by Walley (1984). In Walley's study, children required more acoustic information to identify stimuli than did students in an introductory-level college psychology course, who were probably only slightly older than the oldest teenagers of this study. Similarly, the children in this experiment (and the 70–85-year-old adults) had longer IP durations than the 15–17-year-old teenagers. Walley reported no age differences in the durations (she actually measured number of phonemes rather than milliseconds) of the first meaningful guess, whereas, in this study, children's first word responses occurred at longer durations than did teenagers'. Also, the mean numbers of different incorrect responses for both of Walley's subject groups were considerably greater than those that occurred here (and there were, apparently, no significant between-age-groups effects in numbers of different guesses). It is likely that many of the differences between Walley's work and this study are attributable to the differences in stimulus words. Those used by Walley, although probably known by the younger subjects, were not nearly as familiar as those used in this study. Furthermore, she used words of two and three syllables, whereas only monosyllables were employed in this study.

Walley, Smith, and Jusczyk (1986) reported that young children's perception of speech sounds is "more global" than older children's, that children require perception of more of a word's "constituents" for a response, and that responding to single phonemes is difficult for them. This interpretation corresponds with results of the present study, where children's IPs were longer than teenagers' and where children occasionally gave meaningful incorrect guesses that did not share acoustic phonetic characteristics with the onset of the gated stimulus word. Although the task used by Walley et al. (1986) was quite different from this gating procedure, their data suggest that the same processes may have operated in both experimental procedures.

The multiple regression procedure revealed that guessing meaningful words in response to short gates and guessing fewer different incorrect meaningful words were both associated with shorter IPs. This outcome corresponded with the finding that teenagers guessed fewer different words than did the other groups (Table 5) and more frequently guessed a meaningful word in response to the initial gate.

The significant word × age interaction for IP was primarily attributable to words that had IPs of intermediate duration (i.e., neither the longest nor shortest IPs). For example, when mean word IPs were rank ordered, the five words for which children had longest IPs (ranks 36–40) were in rank positions 32, 31, 34, 36, 37 for teenagers and 32, 35, 36, 37, 38 for older adults. This represents an overall range of seven rank-order positions for teenagers and older adults; the overall range for children was five rank-order positions. Furthermore, rank orderings of words that had longest IPs were similar for all three ages. In contrast, the five words in rank-order positions 11–15 for children ranked 3, 22.5, 2, 25, and 17 for teenagers (a range across 23 rank-order positions) and 8.5, 11, 8.5, 21.5, and 15.5 for older adults (a range of 21 rank-order positions). These significant word × age effects may be attributable to differences in the frequency with which the subject groups had encountered the words or to other, unknown influences.

Elliott et al. (1979) suggested that developmental effects such as age-related changes in pure-tone detection (Elliott & Katz, 1980a; Yoneshige & Elliott, 1981) and in identification and discrimination of simple speech

Table 6
Correlations Between Performance Measures and Mean Pure-Tone Sensitivity (*N*=60)

	Mean Isolation Point	Mean Pure-Tone Sensitivity	Mean Duration of First Meaningful Responses	Mean Number of Different Incorrect Meaningful Responses
Mean Isolation Point	1.00			
Mean Pure-Tone Sensitivity	.50*	1.00		
Mean Duration of First Meaningful Response	.43*	.55*	1.00	
Mean Number of Different Incorrect Responses:				
Meaningful	.65*	.15	−.15	1.00
Nonwords	.60*	.66*	.82*	.11

**p* < .01, one-tailed test.

stimuli (Elliott, Longinotti, Clifton, & Meyer, 1981; Elliott, Longinotti, Meyer, et al., 1981; Elliott et al., 1986) might help explain the age effects in their data. In the present research, however, pure-tone sensitivity measures related to IPs for only some words, there was no particular pattern of word-to-word relations, and pure-tone sensitivity did not contribute significantly to the regression analysis. Furthermore, even for those correlations that were significantly different from zero, pure-tone sensitivity accounted for a very small percentage of the total IP variance. Thus, it was concluded that the listeners' auditory sensitivity did not play a major role in producing the age-related differences of this study. This finding was of special interest since presenting stimuli at equivalent sensation levels can never truly "equate" hearing levels.

The confidence rating data produced mixed results: (1) Children had a mean IP that was longer than teenagers'; (2) children had a mean confidence at IP that was higher than teenagers'; (3) children had a mean total acceptance point (implying highest level of confidence) that was essentially identical to teenagers'; and (4) children accepted with highest confidence 10% fewer words than did teenagers. The children's pattern of assigning confidence ratings differed from teenagers' in complex ways; they did not demonstrate a simple bias toward higher confidence ratings.

Gordon-Salant (1986) reported that elderly listeners performed less cautiously than college-aged subjects on a word-identification-in-noise task—a finding that seems to disagree with the confidence rating data of this study. In contrast, Botwinick, Brinley, and Robbin (1958/1969) found that older subjects waited to obtain more sensory information before responding—an outcome more similar to the results of this study than to those of Gordon-Salant (1986). Differences in the task requirements, however, may account for the different research outcomes. Since this experiment was not designed to investigate this issue in depth, an explanation of these confidence-rating results requires further study.

Differences in frequency of having heard the words of this study may explain the performance differences between the children and teenagers. Word frequency effects, however, cannot explain why older adults performed so much more poorly than the teenagers. If the teenagers had experienced the acoustic representations of, for example, "ball" and "house" more often than the 5-to-7-year-olds, the older adults had surely experienced these words many times more frequently than the teenagers. Yet, their isolation points were consistently longer than those of the teenagers.

One possible reason for older adults' performing more poorly than teenagers pertains to their frequently cited difficulties in processing temporally varied stimuli (e.g., Konkle, Beasley, & Bess, 1977; McCroskey & Kasten, 1982; Newman & Spitzer, 1982; Price & Simon, 1984; Robin, Royer, & Gruhn, 1985). Initial gates were briefer than entire words. Furthermore, even though stimuli were gated in a manner that did not produce audible clicks, the initial gated stimuli did not always sound like natural speech (presumably because the gating procedure "interrupted" a phoneme).

Thus, several different types of processes may underlie the age differences that were observed. Word frequency effects may account for much of the difference between children's and teenagers' mean IPs, whereas differences in ability to process temporally altered stimuli may underlie performance differences between teenagers and older adults. It is known that processing of temporal auditory information improves from the age of 3–5 years to adolescence or young adulthood (e.g., Davis & McCroskey, 1980; Morrongiello, Kulig, & Clifton, 1984). The only other auditory temporal processing procedure that appears to have been administered to young children as well as older adults was a modified auditory fusion task. McCroskey and Kasten (1980) reported that both very young and very old listeners exhibited less efficient temporal processing than did intermediate-aged subjects. Consequently, final attribution of the outcomes of this gating procedure to word frequency and temporal processing effects must await additional research.

SUMMARY

Results may be summarized in five main points:

1. Teenagers exhibited better performance on the gating task than did young children or older adults in terms of the percentage of words identified, the percentage of words totally accepted (i.e., identified at highest level of confidence), and the average duration at which words were first identified (i.e., average isolation point).

2. Teenagers gave their first meaningful guesses at average gate durations that were shorter than children's, which were shorter than older adults'. Older adults gave the largest number of different phonetic guesses, whereas teenagers gave almost no nonword guesses.

3. Measures of auditory sensitivity did not explain the isolation point outcomes; the scattered significant correlations between sensitivity and the experimental measures of the gating paradigm occurred as frequently for children and teenagers as they did for the older adults.

4. Both teenagers and children exhibited better average total acceptance points (i.e., they first identified words with highest confidence at shorter durations) than did older adults. This outcome, in conjunction with differences in mean confidence at average isolation point and in percentages of words that were eventually identified with highest confidence, indicated major differences in patterns of confidence ratings between children and teenagers.

5. The hypothesis was advanced that teenagers' isolation points occurred at shorter gate durations than children's because they had experienced the stimulus words more frequently and made better use of the limited acoustic information. It was proposed that the still greater experience that older adults had had with these words was not sufficient to counterbalance the impact of their difficulties in processing brief, temporally altered stimuli.

REFERENCES

BOTWINICK, J., BRINLEY, J. F., & ROBBIN, J. S. (1969). The interaction effects of perceptual difficulty and stimulus exposure time on age differences in speed and accuracy of response. In F. I. M. Craik (Ed.), Application of signal detection theory to studies of aging, *Interdisciplinary Topics in Gerontology*, **4**, 147-157. (Original work published 1958)

BROADBENT, D. (1967). Word-frequency effect and response bias. *Psychological Review*, **74**, 1-15.

COTTON, S., & GROSJEAN, F. (1984). The gating paradigm: A comparison of successive and individual presentation formats. *Perception & Psychophysics*, **35**, 41-48.

DAVIS, S. M., & MCCROSKEY, R. L. (1980). Auditory fusion in children. *Child Development*, **51**, 75-80.

DUNN, L. (1965). *Peabody Picture Vocabulary Test*. Circle Pines, MN: American Guidance Service.

DUNN, L. M., & DUNN, L. M. (1981). *Peabody Picture Vocabulary Test-Revised*.

ELLIOTT, L. L., BUSSE, L. A., & BAILET, L. (1985). Identification and discrimination of consonant-vowel syllables by younger and older adults. *Perception & Psychophysics*, **37**, 307-314.

ELLIOTT, L. L., BUSSE, L. A., PARTRIDGE, R., RUBERT, J., & DEGRAAFF, R. (1986). Adult and child discrimination of CV syllables differing in voicing onset time. *Child Development*, **57**, 628-635.

ELLIOTT, L. L., CLIFTON, L. B., & FERRE, J. M. (1981). *A note on semantic integration in young children*. Unpublished manuscript available from the first author.

ELLIOTT, L. L., CLIFTON, L., & SERVI, D. (1983). Word frequency effects for a closed-set word identification task. *Audiology*, **22**, 229-240.

ELLIOTT, L. L., CONNORS, S., KILLE, E., LEVIN, S., BALL, K., & KATZ, D. (1979). Children's understanding of monosyllabic nouns in quiet and in noise. *Journal of the Acoustical Society of America*, **66**, 12-21.

ELLIOTT, L. L., & KATZ, D. R. (1980a). Children's pure tone detection. *Journal of the Acoustical Society of America*, **67**, 343-344.

ELLIOTT, L. L., & KATZ, D. (1980b). *Development of a New Children's Speech Discrimination Test*. St. Louis: Auditec.

ELLIOTT, L. L., & KATZ, D. (1980c). *The Northwestern University-Children's Perception of Speech Test (NU-CHIPS)*. St. Louis: Auditec.

ELLIOTT, L. L., LONGINOTTI, C., CLIFTON, L., & MEYER, D. (1981). Detection and identification thresholds for consonant-vowel syllables. *Perception & Psychophysics*, **30**, 411-416.

ELLIOTT, L. L., LONGINOTTI, C., MEYER, D., RAZ, I., & ZUCKER, K. (1981). Developmental differences in identifying and discriminating CV syllables. *Journal of the Acoustical Society of America*, **70**, 669-677.

GORDON-SALANT, S. (1986). Effects of aging on response criteria in speech recognition tasks. *Journal of Speech & Hearing Research*, **29**, 155-162.

GROSJEAN, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics*, **28**, 267-283.

GROSJEAN, F. (1985). The recognition of words after their acoustic offset: Evidence and implications. *Perception & Psychophysics*, **38**, 299-310.

KONKLE, D. F., BEASLEY, D. S., & BESS, F. H. (1977). Intelligibility of time-altered speech in relation to chronological aging. *Journal of Speech & Hearing Research*, **20**, 108-115.

MCCROSKEY, R. L., & KASTEN, R. N. (1980). Assessment of central auditory processing. In R. Rupp & K. Stockdell (Eds.), *Speech protocols in audiology*. New York: Grune & Stratton.

MCCROSKEY, R. L., & KASTEN, R. N. (1982). Temporal factors and the aging auditory system. *Ear & Hearing*, **3**, 124-127.

MORRONGIELLO, B. A., KULIG, J. W., & CLIFTON, R. K. (1984). Developmental changes in auditory temporal perception. *Child Development*, **55**, 461-471.

NEWMAN, C. W., & SPITZER, J. B. (1982). Prolonged auditory processing time in the elderly: Evidence from a backward recognition-masking paradigm. *Audiology*, **22**, 241-252.

PATTERSON, R. D., NIMMO-SMITH, I., WEBER, D. L., & MILROY, R. (1982). The deterioration of hearing with age: Frequency selectivity, the audiogram, and speech theshold. *Journal of the Acoustical Society of America*, **72**, 1788-1803.

PRICE, P. J., & SIMON, H. J. (1984). Perception of temporal differences in speech by "normal-hearing" adults: Effects of age and intensity. *Journal of the Acoustical Society of America*, **76**, 405-410.

ROBIN, D. A., ROYER, F. L., & GRUHN, J. J. (1985). Age-related changes in auditory temporal processing. *Journal of the Acoustical Society of America*, **77**, S37.

SALASOO, A., & PISONI, D. B. (1985). Interaction of knowledge sources in spoken word identification. *Journal of Memory & Language*, **24**, 210-231.

SAVIN, H. (1963). Word-frequency effect and errors in the perception of speech. *Journal of the Acoustical Society of America*, **35**, 200-206.

WALLEY, A. C. (1984). *Developmental differences in spoken word identification*. Unpublished doctoral dissertation, Indiana University, Bloomington.

WALLEY, A. C., SMITH, L. B., & JUSCZYK, P. W. (1986). The role of phonemes and syllables in the perceived similarity of speech sounds for children. *Memory & Cognition*, **14**, 220-229.

YONESHIGE, Y., & ELLIOTT, L. L. (1981). Pure-tone sensitivity and ear canal pressure at threshold in children and adults. *Journal of the Acoustical Society of America*, **70**, 1272-1276.

NOTES

1. Six-year-old boys in the local area seem less interested in earning money than in playing; thus, equal numbers of male and female children could not be recruited for testing. Teenaged boys responded very differently! Females outnumber males among those aged 70-80 years, so the nearly equal representation of males among the older adults was considered particularly good.

2. A lengthy, iterative process identified only 67 monosyllabic nouns that could be represented by pictures and that were within the receptive vocabularies of 3-year-old inner-city children (Elliott et al., 1979).

3. In the original development of the NU-CHIPS stimuli, attention was given to the talker as well as to the words. One male talker could not be understood by many preschool-aged youngsters (Elliott et al., 1979). The talker who made the final recordings was highly intelligible to listeners of this age group (as well as to older listeners).

4. A "correct response" was always identical to the stimulus, even though some other words might share the same phonetic onset. This scoring procedure has been applied by other users of the gating paradigm.

5. Because the words were not randomly selected, it was appropriate to use a mixed ANOVA model with words as fixed effects.

6. The table of these correlations is available from the authors.

(Manuscript received September 22, 1986;
revision accepted for publication March 7, 1987.)