Perception & Psychophysics
1984, 35 (1), 41-48

The gating paradigm: A comparison of successive and individual presentation formats

SUZANNE COTTON and FRANÇOIS GROSJEAN
Northeastern University, Boston, Massachusetts

Recent use of the gating paradigm (Grosjean, 1980) has shown it to be an effective technique for the study of spoken word-recognition processes. However, because of its "successive" presentation format, questions have been raised regarding the effect of repetition on subjects' performance in the task. In the present study, a subset of the words used in the original experiment were recorded at several gate durations and in two of the three original context conditions. The words were presented to different groups of subjects who heard them at only one gate duration. Thus, the repetitive aspect of the presentation format was eliminated. The results were compared to those obtained in the original study with respect to the number of subjects guessing the words correctly, the confidence ratings, and the error patterns at each test gate in each context condition. The findings indicate that, apart from a slight increase in confidence ratings in the context condition, the successive presentation format does not appear to influence the subjects' performance in the gating task.

Spoken language comprehension has been an important area of research in recent years. Many studies have examined how we process, comprehend, and store spoken language. However, a host of new questions have arisen as well as new techniques with which to answer these questions. In the present study, we will examine the methodology of one technique that has been used to study the spoken word recognition process in language comprehension—the gating paradigm (Grosjean, 1980; see also Pollack and Pickett, 1963, and Ohman, 1966, for earlier versions of the paradigm).

The gating paradigm involves the repeated presentation of a spoken stimulus (in this case, a word) such that its duration from onset is increased with each successive presentation. This is done until the entire word has been presented. After each presentation (or gate), subjects are asked to write down the word being presented and to rate their confidence in each guess. Because there is a series of responses for each word, the experimenter has an abundance of dependent variables to study. He or she can examine response accuracy, error distributions, confidence ratings, and the patterns of these responses over time.

Grosjean's (1980) study replicated several effects obtained with other paradigms: the word-length effect (Mehler, Segui, & Carey, 1978), the word frequency effect (Foss, 1969), and the context effect (Morton & Long, 1976). Furthermore, the results showed that the "isolation times" of words (defined as the amount of the word needed to guess it correctly without subsequently changing that guess) are almost identical to the recognition times obtained in monitoring tasks (Marslen-Wilson & Welsh, 1978) and shadowing tasks (Marslen-Wilson, 1975). In addition, and more importantly, Grosjean used the confidence ratings and the error patterns to propose that word recognition may involve two successive operations: the isolation or "highlighting" of a particular word candidate (based on both the acoustic-phonetic and contextual information) and then, after further acoustic-phonetic input, the "acceptance" or recognition of the candidate. In all, the results emphasized the complexity of the word-recognition process.

One question that arises at this point, however, concerns the successive presentation format of the gating paradigm. It can be argued that the repetition of the initial segments of the word may be affecting subjects' performance in the task. It could be that both the isolation times and the progression of the confidence ratings from unsure to sure (as reported in the 1980 study) were artifactually generated by the paradigm itself. For example, given that subjects heard a sequential presentation of each stimulus series, it is possible (or perhaps likely) that they were using the repeated information to narrow down the set of possible word candidates, and were thereby able to guess the word correctly in less time than in a normal listening situation. If repetition were found to be having this sort of facilitative effect, then Grosjean's measurements of isolation times could have been grossly underestimated.

This study was supported in part by grants from the Department of Health and Human Services (RR 07143 and NS 14923). The authors would like to thank Carlos Soares and Steve Harkins for their advice on the statistical analysis, and Lorene Clark, Harlan Lane, Helen Mahut, and Joanne Miller for their useful comments and suggestions. Their thanks also go to David Pisoni and two unknown reviewers for their constructive advice concerning revisions. The authors' mailing address is: Department of Psychology, Northeastern University, 360 Huntington Avenue, Boston, Massachusetts 02115.

Copyright 1984 Psychonomic Society, Inc.

Furthermore, there is also the possibility that the sequential nature of the task may have influenced subjects' confidence ratings in that they were aware that they would be hearing progressively more of the target word and could therefore expect to be more confident at the later gates. Thus, they may simply have been proceeding along the scale in an orderly fashion without actually judging each gate individually.

The answer to these questions is important because one needs to know whether the results obtained with the gating paradigm are specific (in part, at least) to the presentation format of the task itself. In order to answer this question, we posed one of our own: If one were to present individual gates of words to different groups of subjects, would the results obtained in terms of isolation times, confidence ratings, and error patterns differ significantly from those obtained with the successive presentation format (Grosjean, 1980)?

To answer this question, we modified the original experiment by eliminating its repetitive and sequential aspects and then compared the results of the two studies.[1]

METHOD

Subjects

Eighty individuals, with no reported speech or hearing defects, served individually in a single session lasting 30 min.

Materials

Tapes were recorded directly from the original gating stimulus tapes (Grosjean, 1980). A subset of the words were recorded at various gate durations in two context conditions: a no-context (NC) condition, in which the target words are presented alone, and a short-context (SC) condition, in which the target words are preceded by a short sentence.

In the NC condition, 31 words varying in syllable length (one, two, or three)[2] and frequency of occurrence in the language (high or low)[3] were recorded at five gate durations: 30, 120, 210, 300, and 390 msec. The set of words at each gate duration were recorded on a separate tape, for a total of five stimulus tapes in this context condition, each containing 31 stimuli.

In the SC condition, 46 words varying in syllable length[4] and frequency[5] were recorded at five gate durations. Slightly shorter gate durations were recorded in this context condition, since context is known to facilitate word recognition and we were interested in the error patterns prior to the isolation of the target words. Thus, the gate durations were: 30, 90, 150, 210, and 270 msec.[6]

In all, a total of 10 stimulus tapes were used, five in each context condition. On each tape there was an interstimulus interval of approximately 8 sec.

Procedure

Ten groups of subjects, with eight subjects in each group, took part in the experiment.[7] Each group heard 1 of the 10 stimulus tapes. The subjects were assigned randomly to each condition and served individually, listening to the tapes through headphones at a comfortable listening level. They were informed that they would hear words that had been "gated" or cut off. They were instructed to guess the words being presented, write down their guesses in the appropriate place on the answer sheet, and make a slash mark in the sure-unsure scale to indicate their confidence in each guess. The subjects were asked not to leave any blank responses.

Data Analysis

Three dependent variables were examined in this study: the mean number of subjects guessing the words correctly at each test gate in each context condition, the mean confidence ratings at each test gate in each context condition, and the error patterns.

With respect to the first variable, the number of subjects guessing the words correctly, answer sheets were examined to determine how many subjects in each experiment wrote down the target words at a given gate duration (the maximum was eight). To compare the results of this study with those of Grosjean (1980), a mixed between-subjects and within-subjects analysis of variance was performed in each context condition with presentation format and gate duration as fixed effects, and subjects and words as random effects (Erlebacher, 1977).

For the second variable, the confidence ratings, a mean rating (possible range: 0 to 10, measured in centimeters on the scale) was calculated for each word at each gate in each context condition. To compare the two studies, a mixed between-subjects and within-subjects analysis of variance was again performed in each context condition with presentation format and gate duration as fixed effects, and subjects and words as random effects.

Finally, the error patterns were analyzed with respect to the shared guess types, the phonotactic configurations of the erroneous candidates in the two experiments, and the effects of context on the number and types of erroneous candidates.

RESULTS AND DISCUSSION

Number of Subjects Correctly Guessing the Target Words

The results of the original experiment (which we will call the successive presentation experiment) and the present experiment (which we will call the individual presentation experiment) were compared with respect to the mean number of subjects guessing the words correctly as a function of gate duration, in both the NC and SC conditions. The results are presented in Figure 1.

As can be seen, similar findings were obtained in the two studies in each of the two context conditions: as the duration of the gates increased, the mean number of subjects isolating a word also increased, and did so at the same rate for the two presentation formats. Two ANOVAs confirmed this observation. In the NC condition, a main effect for gate duration was found [$F'(4,62) = 77.86$, $p < .01$], indicating that as the gate durations increased, more subjects were able to guess the words correctly. No main effect was found for presentation format, however [$F'(1,14) = 0.68$, n.s.], confirming that repetition did not have an effect on subjects' performance. Finally, there was no significant presentation format × gate duration interaction [$F'(4,62) = 0.87$, n.s.]. In the SC condition, the results showed basically the same pattern: a main effect was found for gate duration [$F'(4,58) = 94.11$, $p < .01$]; no effect was found for presentation format [$F'(1,45) = 0.31$, n.s.], and there was no significant presentation format × gate duration interaction [$F'(4,58) = 0.86$, n.s.].

Thus, the same number of subjects guessed the words correctly at each test gate in each context con-

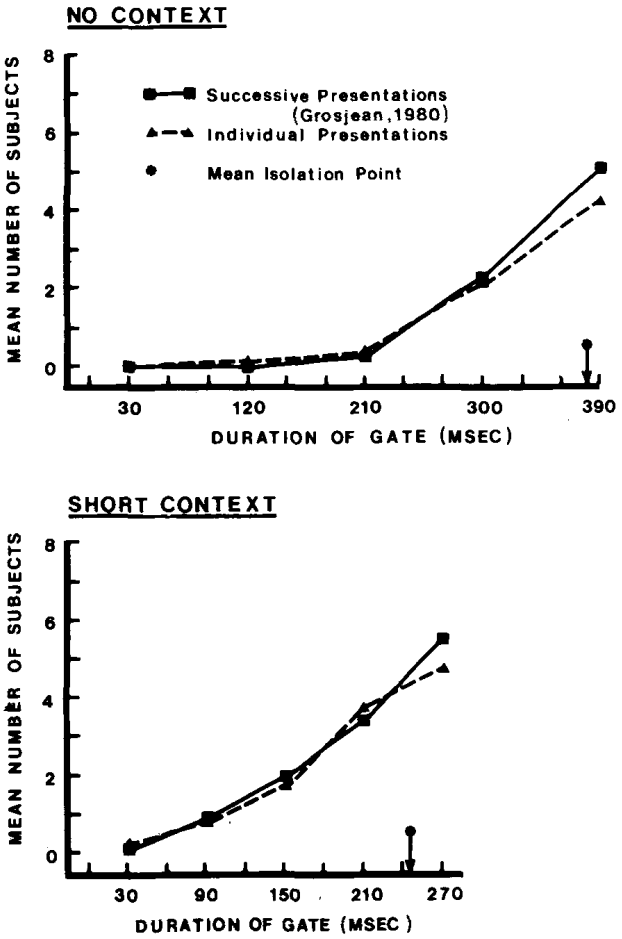

Figure 1. Mean number of subjects guessing the words correctly as a function of gate duration in both the original study (Grosjean, 1980; successive presentations) and the present study (individual presentations). The upper graph shows the results in the NC condition (each point is based on 31 words); the lower graph shows the results in the SC condition (each point is based on 46 words). The asterisk on the X axis of the top graph (NC condition) indicates the mean isolation time obtained by Grosjean (1980) for the 31 words tested in this study. The corresponding mean time for the SC condition (bottom graph) is also indicated.

dition, and this number increased as subjects heard more of the word. It seems safe to conclude, therefore, that repetition was not aiding the subjects in Grosjean's study and that the isolation times obtained in his study were not specific to the presentation format.[8]

It is worth recalling here that Grosjean (1980) obtained word isolation times that were very similar in value to word recognition times inferred by Marslen-Wilson (Note 1) from data obtained with totally different tasks: word monitoring and shadowing. This would seem to confirm the validity of the gating paradigm as used by Grosjean.

Confidence Ratings

Figure 2 presents the mean confidence ratings, measured in centimeters, as a function of gate duration. In this case, a rating of 0 indicates perfect confidence and a rating of 10 indicates complete uncertainty. Different patterns emerge in the NC condition and the SC condition. In the NC condition, no significant effect for presentation format emerged [$F'(1,14) = 1.142$, n.s.], but a strong main effect for gate duration [$F'(4,58) = 300.19$, $p < .01$] was found. This indicates that subjects were more confident of their guesses when longer gate durations were presented, regardless of the presentation format. A small, but significant, presentation format × gate duration interaction was also found [$F'(4,58) = 7.90$, $p < .05$]. In the SC condition, however, significant effects were found for both presentation format [$F'(1,49) = 167.65$, $p < .01$] and gate duration [$F'(4,61) = 93.73$, $p < .01$], as was a presentation format × gate duration interaction [$F'(4,61) = 16.24$, $p < .05$].

A possible explanation for the different presentation format patterns found in the SC condition could be that subjects in the individual presentation experi-

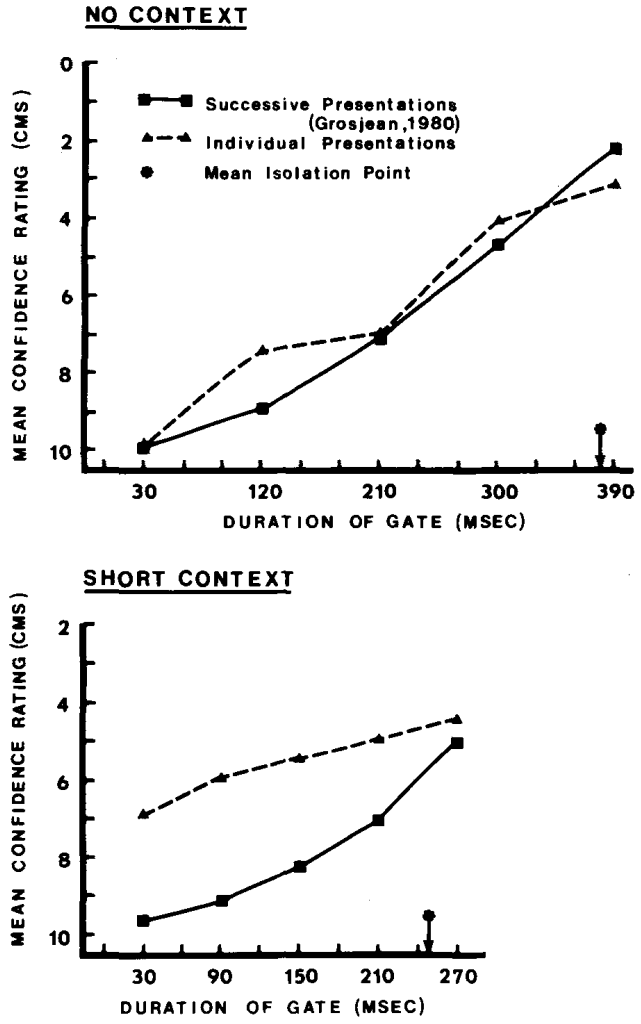

Figure 2. Mean confidence ratings, measured in centimeters, as a function of gate duration in each of the two context conditions and with two presentation formats. A rating of 0 indicates perfect confidence; a rating of 10 indicates complete uncertainty. See the figure caption of Figure 1 for information concerning the asterisks.

ment give greater weighting to the semantic-syntactic context of the sentence. That is, they give more weight to the fit of their candidate with the context than to the fit of the candidate with the acoustic-phonetic information. This in turn produces lower values (higher confidence), since there is more semantic-syntactic information than acoustic-phonetic information. Subjects in the successive presentation format, on the other hand, pay more attention to the bottom-up (acoustic-phonetic) information—at least at the first gates—and hence reflect the NC pattern in their ratings. This pattern reflects complete uncertainty at the first gate (30 msec of the word) and then a slow progression towards greater certainty. This greater reliance on the bottom-up information could itself be due to the fact that subjects in the successive presentation format have learned to be conservative in their ratings; they may have been led down a few garden paths in previous word sets, for example.

Evidence for this explanation comes from an examination of the mean confidence ratings at the third gate (90 msec) of the first four words in the two experiments. In the successive presentation format, the values given by subjects increase (i.e., subjects become less confident) as new words are presented: 4:14 for the first word (Bog), 8.51 for the second (Pharmacist), 9.19 for the third (Stranger), and 8.99 for the fourth word (Gull). In the individual presentation experiment, however, the values remain stable at a middle value: 4.88, 6.61, 6.70, and 5.60, respectively. This seems to indicate that the two formats lead to different confidence rating strategies, but only, it should be noted, when the words are in context and when early gates are presented. At later gates, the ratings are practically identical. For example, at the last test gate (270 msec) in the SC condition, the two groups of subjects produce almost identical ratings: 5.50 for the successive presentation format and 4.78 for the individual format ($t = 1.83$, n.s.).

Another similarity between the two experiments is the increase in confidence as the gates increase in duration. This is clearly evident in the NC condition. As for the SC condition, a post hoc analysis (Tukey HSD, $p < .05$; Kirk, 1967) reveals that subjects in both experiments are more confident of their guesses at the last test gate than they are at the first gate.

In conclusion, we can state that a presentation format effect is present only at early gates and in context conditions. At later gates and in isolation, the presentation formats give identical results. In addition, and irrespective of the format, subjects increase their confidence in their guesses as more and more of the stimulus word is heard.

Error Patterns

An illustration of the error patterns found in the two studies is presented in Figure 3.

In this figure, the candidates proposed for the target word CAPTAIN are shown for both the NC and SC conditions. Candidates that are proposed at only one gate are depicted with a dot; continuous lines represent those candidates that are proposed at two or more consecutive gates. The thickness of the lines is representative of the number of subjects proposing a particular candidate—the thicker the line, the more subjects proposing it at that point in time.

Grosjean noted three specific effects that context has on the candidates proposed prior to the isolation of the target word. One effect is the reduction in the number of candidates proposed by subjects. In this example, 15 candidates were proposed in the NC condition in the original experiment, whereas only 10 candidates were proposed when the target word was presented in context (top part of figure). The same effect is seen in the present study; 18 candidates were proposed in the NC condition but only 8 in the SC condition—a marked and similar reduction.

A second effect is the point at which the target word begins to be proposed by subjects. In isolation, the target word is rarely a candidate at onset, whereas in context the target word is proposed very early, possibly even at the first presentation. In this example, CAPTAIN was proposed at the first gate in the SC condition in both the original experiment and the present one. It did not appear as a candidate in the NC condition, however, until much later—120 msec in the original study and 210 msec in the present study.

A third effect is the type of candidates proposed in and out of context. The sentence in our example (Figure 3) is "Margaret told the . . .," and, in both experiments, subjects proposed words that were generally appropriate to that context in the SC condition, whereas their guesses in the NC condition reflected only the acoustic-phonetic information. It is also interesting to note that, because of the coarticulatory information in the SC condition, all the candidates began with the correct initial consonant, /k/, and this in both experiments. Subjects proposed KIDS, CAMP, COW, and CAT in both the original experiment and the present one in the SC condition. In the NC condition, however, the correct initial consonant did not begin to appear in the guesses until about 120 msec. But once the initial consonant had been correctly perceived in the NC condition, the erroneous candidates were similar in the two experiments; subjects in both studies proposed CAT, CAP, COUNT, CALCULATE, and CAPTIVE.

In order to show the similarity of the erroneous candidates in the two experiments in general, we calculated the percent shared candidates across the studies. To do this, we counted the number of shared guess tokens at each gate and for each word and divided this by the total number of candidate tokens at that gate and for that word (that is, by 16, as there were eight subjects in each of the two studies). This ratio was then multiplied by 100 to give a percentage. For example, at gate 10 (300 msec), the following

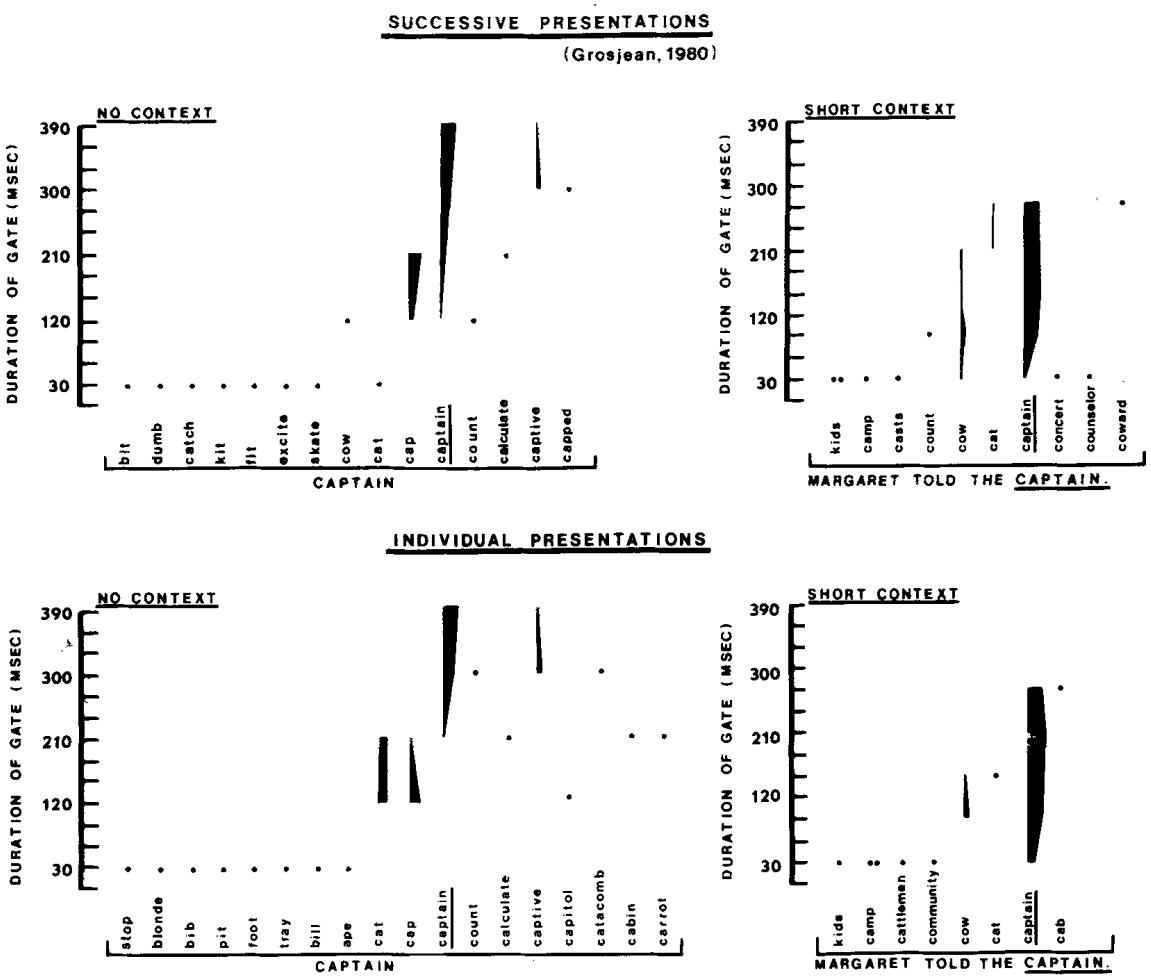

Figure 3. The candidates proposed for the target word CAPTAIN in both the no-context and short-context conditions, in the original study and the present one. Candidates proposed at only one gate are depicted with a dot; continuous lines indicate those proposed at several consecutive gates. The thickness of the lines is representative of the number of subjects proposing a particular candidate–the thicker the line, the more subjects.

candidates were given for the stimulus word SMOG in the NC condition:

Successive Presentation	Individual Presentation
flop	smug
smock	smock
smock	smock
smile	small
small	small
slide	fog
snug	swallow
smog	slug

There are 16 total candidate tokens of which 7 are shared (i.e., 4 *smock*s and 3 *small*s). Thus, the ratio becomes 7/16, or 43.8%, shared responses. The mean percent shared candidates at each test gate are presented in Table 1.

Several points can be made based on the table. The first is that the percent shared candidates is lower at the first gates than at the last gates. This is quite simply due to the fact that at early gates there are many more candidates for a stimulus word than could be proposed by the eight subjects in each experiment. As a consequence, only a few candidates are shared by the two groups of subjects. This pool-of-candidates phenomenon is clearly illustrated if we take a sample set of candidates proposed by the first four subjects

Table 1
Mean Percent Shared Candidates at Each Test Gate in the Two Context Conditions

No Context		Short Context	
Gate Durations	Percent Shared Candidates	Gate Durations	Percent Shared Candidates
30	3.4	30	34.1
120	26.0	90	45.3
210	39.9	150	56.4
300	62.7	210	76.8
390	72.2	270	82.2

Note–Gate durations given in milliseconds.

in Grosjean's experiment and compare them with those proposed by the second four subjects. For 10 words and at each of the test gates, we found the following percentages: 7.5%, 17.5%, 27.5%, 43.7%, and 65.0% in the NC condition, and 41.2%, 37.5%, 42.5%, 55.0%, and 55.0% in the SC condition. Thus, the rather low values at the early gates presented in Table 1 are really a reflection of candidate sampling by subjects and not of the differences between the two presentation formats. It is encouraging that at the last test gates the percent shared candidates was found to be 72.2% in the NC condition and 82.2% in the SC condition, indicating great similarity in candidates.

The second point is that, as expected, context leads subjects to propose a greater percentage of common candidates at earlier gates: for example, at 30 msec, there are 34.1% shared candidates in the SC condition but only 3.4% in the NC condition.

Therefore, not only are the subjects guessing the target words correctly at the same time in both presentation formats (see Figure 1), but their overall responses in both tasks prior to isolation are very similar.

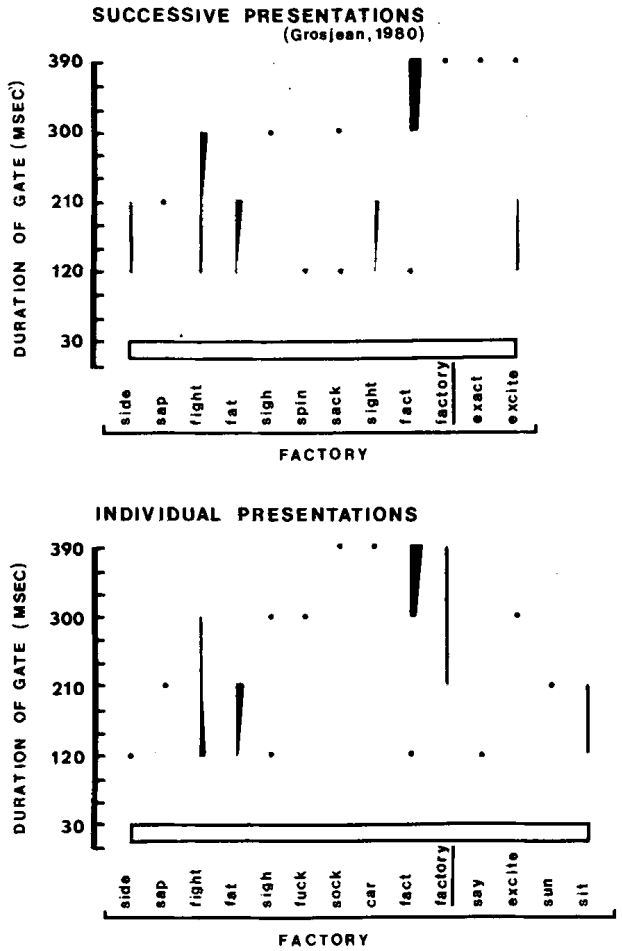

Figure 4. The "word from a word" garden path: candidates for the word FACTORY in the no-context condition in the successive and the individual presentation experiments. For lack of space, the candidates proposed at the 30-msec gate are not shown.

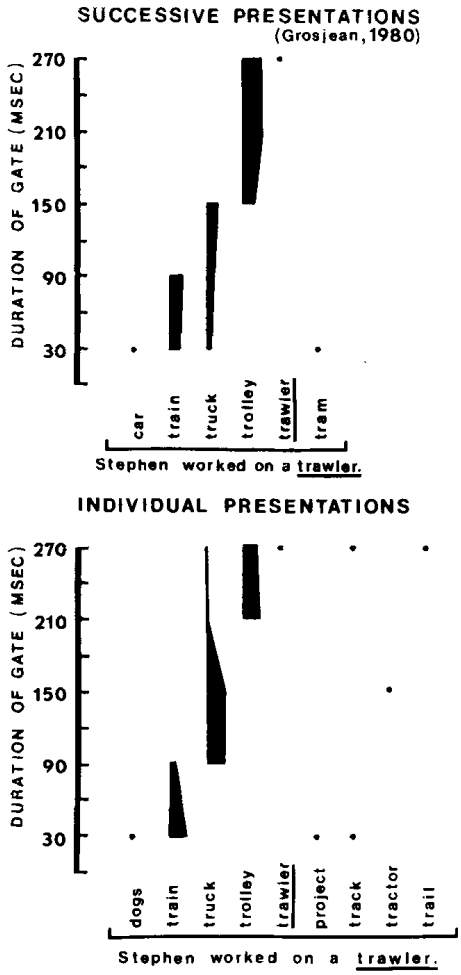

Figure 5. The frequency garden path: candidates for the word TRAWLER in the short context condition in the two experiments (successive and individual).

Another type of phenomenon that was delineated in the original study with respect to error patterns is the "garden path" phenomenon. Subjects are said to go down a garden path if a number of them propose an erroneous candidate over one or several gates. Three types of garden paths were proposed: the "word from a word," the frequency, and the semantic garden paths.

In both Figure 3 (see above) and Figure 4, we see examples of the "word from a word" garden path.

In this type of garden path, subjects propose a word that is contained within the target word for several gates until enough acoustic-phonetic information arrives to convince them that another syllable will follow. As can be seen in Figure 3, subjects proposed CAP for CAPTAIN in the NC condition, and in Figure 4, we note that subjects propose FACT for FACTORY. This garden path occurred in both the successive presentation format and the individual presentation format. The "word from a word" garden path almost never appears in the SC condition because the word is not usually appropriate to the semantic content of the context sentence. In our two examples, CAP was not proposed in the SC condi-

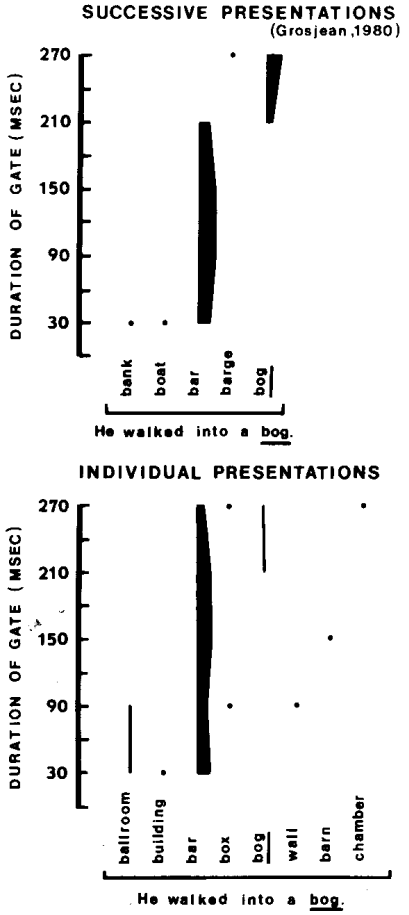

Figure 6. The semantic garden path: candidates for the word BOG in the short context condition in the two experiments.

tion as it did not fit the sentence "Margaret told the . . ."; nor was FACT proposed in the SC condition because it did not fit the sentence "John entered the. . . ."

The frequency garden path occurs when a target word is a low-frequency word and subjects propose a high-frequency word that is appropriate both to the acoustic-phonetic information and, in context, to the semantic/syntactic information. In Figure 5, we present the candidates for the low-frequency target word TRAWLER. Subjects in both experiments proposed TRAIN, TRUCK, and TROLLEY. These words are all higher frequency words than the target word, they begin with the correct initial consonant cluster /tr/, and they are appropriate to the sentence "Stephen worked on a. . . ."

performance in the successive presentation format and can conclude that both the successive and individual presentation formats yield similar results.

There are, however, two major advantages to the successive presentation format of the gating paradigm: the ease of stimulus presentation and preparation, and the ability to call on a much smaller subject pool. These two advantages make the successive presentation format the more useful of the two approaches.

The semantic garden path occurs when the semantic content of the context sentence leads the subject to propose a particular erroneous candidate. An example of this is shown in Figure 6. The context sentence "He walked into a . . ." led subjects in both studies to almost unanimously propose BAR instead of the target word BOG.

If we tabulate all the garden paths in the two experiments (in which a garden path is a particular erroneous candidate proposed by four or more subjects at one or more gates), we find that there are 65% shared paths in the NC condition and 72% shared paths in the SC condition—another indication of the similarity of the two presentation formats.

CONCLUSION

We have replicated in this study the results of Grosjean's original (1980) study with respect to both the number of subjects guessing the target words correctly and the error patterns obtained. As for the confidence ratings, we have shown that it is only at the early gates and in context that the presentation format has any effect. At later gates and in isolation, the ratings are similar. In addition, subjects become more confident of their guesses as they hear more of the word, irrespective of the presentation format.

We have been able to show, therefore, that repetition does not have a facilitative effect on subjects'

REFERENCE NOTES

1. Marslen-Wilson, W. *Sequential processes during spoken word recognition.* Paper presented at the Annual Meeting of the Psychonomic Society, San Antonio, Texas, 1978.
2. Salasoo, A., & Pisoni, D. *Sources of knowledge in spoken word identification* (Research on Speech Perception Progress Report, Vol. 8, pp. 105-145). Bloomington: Department of Psychology, Indiana University, 1982.

REFERENCES

ERLEBACHER, A. Design and analysis of experiments contrasting the within- and between-subjects manipulation of the independent variable. *Psychological Bulletin*, 1977, **84**, 212-219.

FOSS, D. Decision processes during sentence comprehension: Effects of lexical item difficulty and position upon decision times. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 457-462.

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

KIRK, R. *Experimental design: Procedures for the behavioral sciences.* Belmont, Calif: Brooks/Cole, 1967.

MARSLEN-WILSON, W. Sentence perception as an interactive parallel process. *Science*, 1975, **189**, 226-228.

MARSLEN-WILSON, W., & WELSH, A. Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, 1978, **10**, 29-63.

MEHLER, J., SEGUI, J., & CAREY, P. Tails of words: Monitoring ambiguity. *Journal of Verbal Learning and Verbal Behavior*, 1978, **17**, 29-35.

MORTON, J., & LONG, J. Effect of word transitional probability on phoneme identification. *Journal of Verbal Learning and Verbal Behavior*, 1976, **15**, 43-51.

OHMAN, S. Perception of segments of VCCV utterances. *Journal of the Acoustical Society of America*, 1966, **40**, 979-988.
POLLACK, I., & PICKETT, J. The intelligibility of excerpts from conversation. *Language and Speech*, 1963, **6**, 165-171.

NOTES

1. It is not the aim of the present study to extend existing theories of word recognition or to discuss the value of the gating paradigm as a "true reflector" of the underlying operations leading to lexical access in real-time processing. Both these issues are critical, but our aim here is simply to learn more about the gating paradigm by testing its presentation format.

2. In the no-context condition, there were 4 one-syllable words, 11 two-syllable words, and 16 three-syllable words. All of the 31 words in this condition were included in the list of 46 words in the short-context condition.

3. There were 15 high-frequency words and 16 low-frequency words in the no-context condition.

4. In the short-context condition, there were 14 one-syllable words, 16 two-syllable words, and 16 three-syllable words.

5. There were 23 high-frequency words and 23 low-frequency words in the short-context condition.

6. Thus, the NC and SC conditions differ in number, frequency, and length of words. But because the study was aimed at comparing presentation formats *within* a condition and not between conditions, no effort was made to equate the two conditions.

7. Although more subjects would have been preferable, we wanted to keep to the number used by Grosjean (1980).

8. In a recent study, Salasoo and Pisoni (Note 2) describe the use of a modified version of the gating paradigm: Words in a sentence are replaced by envelope-shaped noise and, in consecutive presentations of the sentence, 50-msec increments of the original speech waveform replace either the initial or the final segment of the noise replacing the words, and this is incremented until the entire words are presented. As part of their second experiment, the authors tested the presentation format of the paradigm (successive vs. individual presentations). They found no presentation format effect, for either the forward-gated or backward-gated words, thus confirming that the use of a repeated presentation method does not affect subjects' performance in the gating task.

(Manuscript received May 16, 1983;
revision accepted for publication October 5, 1983.)