Prosody, Memory Load, and Memory for Speech

BURTON S. ROSNER, ESTHER GRABE, HANNELE B. M. NICHOLSON, KEITH OWEN AND ELINOR L. KEANE

1 Introduction

Previous experiments indicate that prosody may facilitate memory for spoken utterances. Memory load, however, has not been systematically manipulated and may interact with such facilitating effects. Therefore, subjects were asked to listen to connected prose passages between one to five intonation phrases long and to repeat as much of each narrative as possible. Stimuli were normal productions or else had been deprived of intonational variation, pauses, or both. Performance on both the normal and the prosodically altered passages decreased as passage length grew. Errors were overwhelmingly omissions and occurred mainly in the middle of a passage. The prosodic manipulations affected memory for the longest passages. These results alongside previous findings suggest that prosody facilitates memory for spoken stimuli that are relatively difficult to process. Lengthy material, infrequent words, unrelated items, citation form prosody, or thematic or grammatical anomalies may create such difficulties.

2 Background

Numerous experiments show that prosody affects spoken word recognition, the computation of syntactic relationships, and the processing of discourse structure (see Cutler, Dahan, & van Donselaar, 1997, for a review). Other work indicates that disruption of prosody can interfere with memory for spoken language (Darwin, 1975; Wingfield, 1975a, 1975b; Wingfield & Klein, 1971; Stine & Wingfield, 1987; Wingfield, Lahar, & Stine, 1989; Paris, Thomas, Gilson, & Kincaid, 2000) In these studies, however, memory load was not systematically manipulated. Since speech contains numerous redundant cues, prosody may be unimportant for remembering relatively simple, short utterances. As processing demands grow, however, prosody may increasingly affect memory for connected prose. To test this hypothesis, we examined the influence of intonation and of pauses on memory for continuously spoken passages of differing lengths.

Leonard (1974), O'Connell, Turner, and Onuska (1968) and Zurif and Mendelsohn (1972) all reported that monotone speech produced decrements in memory for spoken items. We therefore eliminated all intonational variation in the first of our three conditiona. Using a resynthesis technique, F0 was held constant throughout an entire narrative.

The results of Huttenlocher and Burke (1972), Frankish (1985, 1989), Saito (1998), and Martin (1968) indicate that pauses, appropriate or otherwise, do affect memory for speech. Furthermore, manuals on public speaking (e.g., Mandel, 1993; Berry, 1994) discuss the importance of putting pauses of the right lengths at the right places, in order to help the listener's comprehension. Presumably, comprehension of running speech would call on mnemonic resources. For our second experimental condition, then, we removed all naturally occurring pauses from continuously spoken prose. In a third condition, we eliminated both intonational variation and pauses from spoken narratives. This procedure seemed to promise the greatest disruption of memory for prose.

3 Method

Stimuli. The stimuli were based on six-word utterances. Each basic utterance comprised one intonation phrase (IP). A complete passage contained between one and five such basic items and formed a connected narrative about a single topic. One six-word constituent of a passage was drawn from the revised Harvard Sentences (Institute of Electrical and Electronic Engineers, 1965). We added whatever further narrative material was necessary. When a passage contained two or more basic utterances, the insertion of an initial 'and' made the final part seven words in length. The 1-IP passages therefore had six words, the 2-IP passages had 13 words, the 3-IP passages 19 words, the 4-IP passages 25 words, and the 5-IP passages 31 words.

Two sets of passages, labelled set I and set II, were constructed. Each set was ordered to begin with two 1-IP sentences. They were followed consecutively by one 2-IP narrative, a 3-IP narrative, a 4-IP narrative, two 5-IP narratives, one 4-IP narrative, a 3-IP narrative, a 2-IP narrative, and finally two 1-IP sentences. Each set of 12 narratives therefore contained four 1-IP passages and two of each of the

2-IP, 3-IP, 4-IP, and 5-IP types. Example sentences, with accented words in upper case, are: 1-IP) Men STRIVE but seldom get RICH; 2-IP) The GIRL gave no clear RESPONSE, and the MEDICS went straight to WORK; 3-IP) SOME ads serve to cheat BUYERS; read EACH with VERY great care, and look CLOSELY at any SMALL print; 4-IP) Pink CLOUDS floated with the BREEZE, the SUN was setting at NIGHTFALL, the SKY slowly turned deep BLUE, and EVERY street light began to GLOW; 5-IP) The MAP shows where we ARE; the BAG contains something to EAT; your CLOTHES are in the SUITCASE; the TENT is in the BOOT; and the CAR is full of PETROL. After the experiment was under way, we found that one 3-IP passage in set II contained only 18 rather than 19 words, while a 4-IP passage in that set contained 26 rather than 25 words.

A male native speaker of Southern British English produced all 24 passages, which were recorded on a CD. In the Appendix, words accented during recording appear in capitals. (See Ladd, 1996, Ch. 6, for a discussion of accentuation.) For recording, we used an Audio-Technica AT4031 cardioid microphone and a Symetrix SX202 phantom power supply and preamplifier. The preamplifier output drove a HHB Communications CDR-850 compact disk recorder.

The Cool96 editing program (Syntrillium Software Corporation, 1996) running on a PC was used to digitize each spoken passage at 22050 16-bit samples per second and to normalize its maximum amplitude to 100 per cent. Each normal passage was then stored to the computer hard disk as a separate .WAV file. Next, we employed PRAAT 3.8 (Boersma & Weenink, 1996), a speech processing program, to make three variants of each normal passage: monotone, pause-free, and monotone-pause-free. The monotone variant was produced by PSOLA resynthesis (Carpentier & Moulines, 1990) of the passage, with F0 fixed throughout at 120 Hz. To produce the pause-free variant, we manually edited out all pauses from the normal passage. Finally, the pause-free variant was subjected to a PSOLA resynthesis with F0 again fixed at 120 Hz, generating the monotone-pause-free variant. At the end of these procedures, we had 96 .WAV files, with a normal, a monotone, a pause-free, and a monotone-pause-free version of each of the 24 spoken passages.

Subjects. Subjects were 24 undergraduates or graduate students at the University of Oxford. All were native speakers of Southern British English. Nine were male and 15 female. They were recruited by the experimenters and were paid for participating in one experimental session of about 45 min duration.

Design. Each subject heard two sets of passages. One had normal prosody, while the other was one of the three variants. Half the subjects heard the normal passages first, and the other half heard either the monotone or the pause-free or the monotone-pause-free passages first. This counterbalanced the normal/variant order. Of the 12 subjects in each of those two subgroups, six heard the normal passages in set I and variant passages in Set II, and the other six subjects heard the normal passages in set II and the variant passages in Set I. This counterbalanced normal and variant passages across sets. Altogether, the design yielded 12 conditions: two orders x three variants x two distributions of sets I and II. Consequently, two subjects underwent a given condition. One of those subjects heard the narratives in each set in the order shown in the Appendix. The other subject heard those narratives in the reverse order.

Procedure. Listeners were tested individually in a sound-attenuated recording booth. The subject faced a VDU and held a keyboard on his or her lap. A PC outside the booth controlled the experiment through a program written in C++. The subject was instructed that the experiment would be split into two parts, with a short rest period between them. In one part, a series of normally spoken passages would be heard. In the other, passages altered by a computer would be heard. The subject was asked to listen to each passage and, after the passage had finished, to repeat aloud what had been heard, as accurately as possible. The subject was encouraged to get as many words as possible, even if some words were felt to have been missed. Stimuli were presented at approximately 65 dBA over Sennheiser HD-320 earphones.

At the start of each part of the experiment, the subject's name and date of birth were entered into a data file, along with the current date and time. The filenames of the passages that the subject would hear were read in their desired order from an input file and were written to the data file. Then a message on the VDU told the subject to signal readiness by hitting the 'Enter' key on the keyboard. Upon sensing that action, the PC sent a new message to the VDU. This asked the subject to initiate a passage by pressing the 'Enter' key. When the subject did so, the VDU went blank and a passage was played after a 500 ms pause. At the end of the passage, a message on the VDU asked the subject to repeat the passage and to hit the 'F' key on the keyboard when finished. The subject's spoken response was recorded on a CD, using the system described above. After the subject pressed the 'F' key, a pause of 1 s occurred before the beginning of the next trial. The passages within each half of the experiment increased and then decreased in length, as shown in the Appendix. This seemed to make the subject feel comfortable. At the start or the end of the experiment, the WAIS-R digit span subtest was administered to the subject.

Scoring. Using the data file for a given subject, two experimenters scored the responses recorded on CD. The number of words reproduced in the correct order was counted first, even when some intervening words had been missed. Then words reproduced out of order were identified. Finally, incorrect intrusions or substitutions were noted. The two experimenters had to agree on the scoring for each passage. This often required them to listen several times to the subject's attempted repetition of a given passage, particularly when the passage was long and the repetition became hesitant.

4 Results

Intrusions or substitutions were infrequent. They will receive no further attention. Two scores were determined for a subject's response to each passage. One measure, $P(C_O)$, was the proportion of words reproduced in the correct order and represented order information. The other measure was the proportion of words reproduced in total without regard to order, $P(C_I)$, representing item information.

Relatively few items were actually reproduced out of order. Consequently, the two measures differed very little and were often identical. Nonetheless, we subjected both measures to the same statistical treatments. Apart from one analysis, no differences emerged. We therefore report only the results on $P(C_O)$, the proportion of words reproduced in the correct order, except for the single analysis where the two measures differed somewhat. For some analyses, we combined data across the monotone, pause-free, and montone-pause-free conditions. Otherwise, comparisons would have depended on a small number of degrees of freedom, opening the way to errors of type II.

4.1 Shorter (1-IP and 2-IP) passages

The $P(C_O)$ scores for normal passages showed that the 1-IP stimuli were reproduced with no errors. The 2-IP normal passages occasioned a total of just three errors, each due to a different subject. The geometric mean for $P(C_O)$ was .995.

Across all 96 attempts at reproducing prosodically altered 1-IP narratives, the geometric mean for $P(C_O)$ was .968. One subject completely failed to reproduce one altered 1-IP passage, generating six errors, and made an error on another altered 1-IP narrative. Twelve more errors were distributed across six subjects, giving a total of 19 errors. The altered 2-IP passages again produced just three errors, distributed over two subjects, and gave a geometric mean of .995 for $P(C_O)$.

In short, the data show virtually perfect memory for 6-word and 13-word passages of connected prose. Results on both the normal and the prosodically altered 1-IP and 2-IP passages therefore were dropped from further analysis, due to lack of variance.

4.2 Longer (3-IP, 4-IP, and 5-IP) stimuli

One subject failed to reproduce anything after hearing a normal 4-IP passage and a normal 5-IP passage in set I. Another subject dried up similarly on a normal 4-IP passage in set II. Finally, one subject gave no response to a 5-IP pause-free passage. We counted each of these four failures as a response of zero. Non-zero responses were therefore made to practically all the normal and the prosodically altered stimuli.

Overall, perfect performances occurred on about 15 per cent of the 3-IP and 4-IP passages. None occurred on the 5-IP stimuli. About two-thirds of the partial recalls of the 3-IP, 4-IP, and 5-IP passages arose from omission of words in the middle of a narrative. The remaining minority of imperfect performances were largely due to omission of material either at the beginning or at the end of a passage. Rarely did a listener recall words exclusively from the middle of a passage.

Inspection of the data from both the normal and the combined prosodically altered conditions revealed some exponentially shaped or bimodal samples. No monotone transformation could render those results Gaussian. Therefore, Monte Carlo nonparametric statistics (500,000 samples per test) were used for analyses of differences. We set α at a conservative .025 for all statistical tests.

4.3 Aberrant passages and condition orders

As stated earlier, set II proved to have one 3-IP passage containing only 18 rather than 19 words and a 4-IP passage containing 26 rather than 25 words. We compared performance on each aberrant narrative against performance on its counterpart in set II with the same number of intonation phrases. This yielded two one-tailed Wilcoxon tests for $P(C_O)$ for the normal passages and two more for data combined across prosodically altered passages. None of the four tests (N=12 in each case) was significant. Accordingly, we pooled data across both 3-IP passages and across both 4-IP passages within set II as well as within set I.

Data on $P(C_O)$ were compared across condition orders (normal passages first or altered passages first) within each passage size and within each set, for the normal passages and separately for the prosodically altered stimuli combined. This resulted in 12 two-tailed Mann-Whitney tests (N=12 in total in each case) over the 3 passage sizes, 2 sets, and normal or altered passage type. None of the 12 tests proved significant, so we pooled data across condition orders.

4.4 An unexpected result: differences between sets

Box plots of $P(C_O)$ indicated that the stimuli in set I were harder to remember than those in set II. The upper panel in Figure 1 contains the data for normal 3-IP, 4-IP, and 5-IP passages, and the lower panel shows data pooled across altered passages. Given our within-subjects experimental design, we had envisioned comparing results across the two sets in order to measure the effects of the prosodic manipulations. Before making any such comparisons, however, we first had to detour into an examination of the apparent differences between sets I and II.

Two-tailed Mann-Whitney tests (N =24 for each test) showed that $P(C_O)$ for the normal 3-IP and 4-IP passages differed significantly between sets, $z = 2.410$, $p < .025$, and $z = 3.820$, $p < .001$, respectively. For data combined over the prosodically altered stimuli, the 3-IP and 4-IP passages also yielded significant differences between sets, $z = 2.410$, $p < .025$, and $z = 3.820$, $p < .001$, respectively. In keeping with these four significant differences, perfect performances occurred over 20 per cent of the time on the 3-IP and 4-IP passages of set II, normal or altered. Only some 3 per cent of the 3-IP and 4-IP passages of set I produced perfect performances. The difference between sets was not significant for the normal or for the altered 5-IP stimuli.

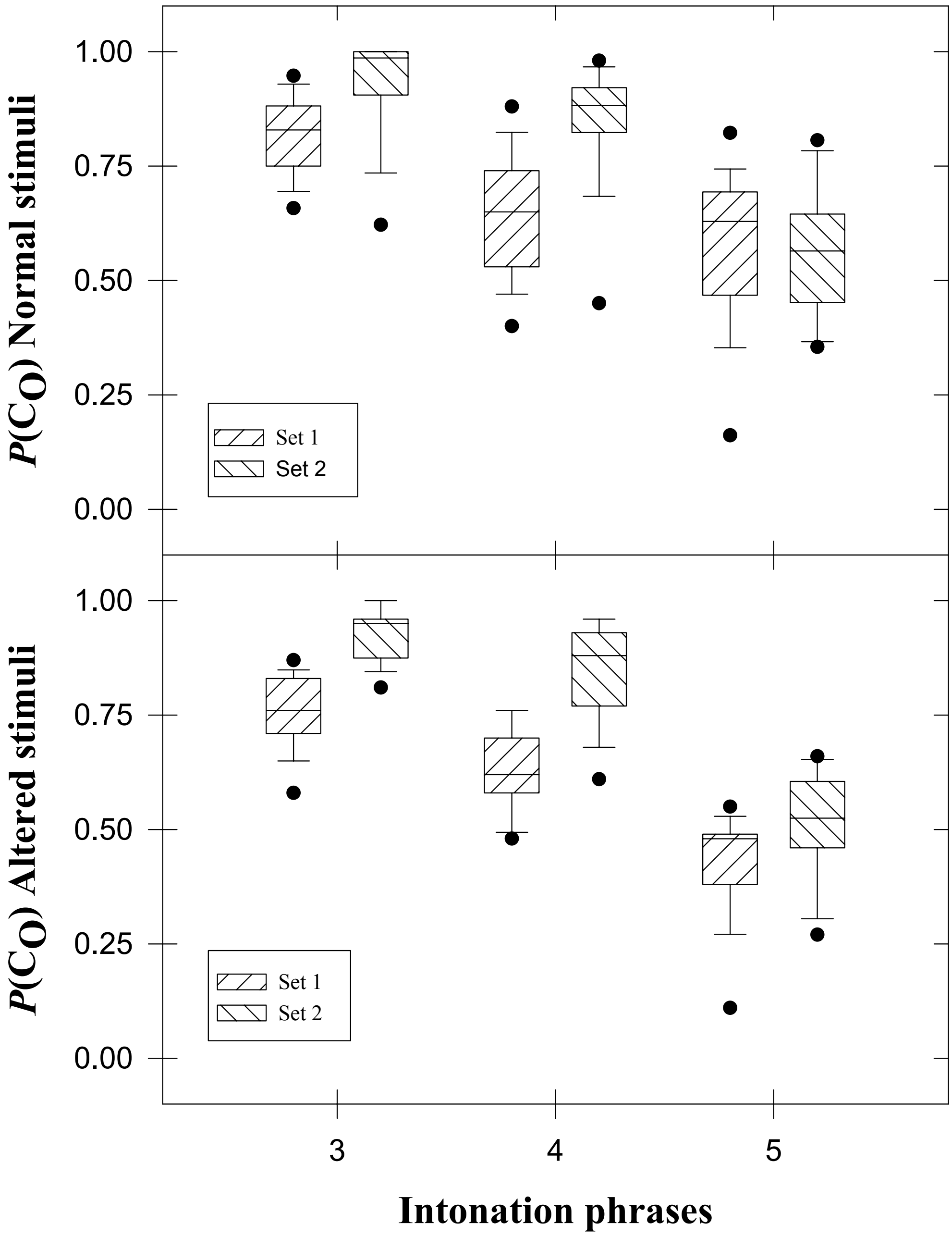

Fig. 1. Box plots of $P(C_O)$ for normal passages (upper panel) and for prosodically altered passages (lower panel) with different numbers of intonation phrases. Results are shown separately for stimuli in set I and in set II. Each box runs from the 25th to the 75th percentile. The line through the box shows the median. Whiskers delineate the 10th and 90th percentiles, and black points indicate outliers.

Figure 1 also shows that longer passages imposed a greater memory load, as expected. Confirmation came from four Monte-Carlo Friedman one-way ANOVAs on $P(C_O)$, carried out for each set of normal stimuli and each set of prosodically altered stimuli combined. The differences between sets forced the use of separate ANOVAs. Each ANOVA ($N = 12$) had three levels (3-IP, 4-IP, and 5-IP passages). For normal sets I and II, the ANOVAs gave significant results, $\chi^2(2) = 15.167$, $p < .001$, and $\chi^2(2) = 18.167$, $p < .001$, respectively. Across the prosodically altered stimuli in sets I and II, the ANOVAs yielded $\chi^2(2) = 18.167$, $p < .001$, and $\chi^2(2) = 20.667$, $p < .001$, respectively.

The subjects who had heard the normal passages of set I heard the altered passages of set II and vice-versa. Nonetheless, the 3-IP and 4-IP stimuli of set I, normal or altered, were always more difficult than those of set II. These differences between the two sets therefore arose from a discrepancy in inherent difficulty and not from differences between subjects.

The syntactic structures of the narratives in sets I and II provided no obvious reasons for the better recall of set II. Within each IP, structures were principally SVO or VO. The only embedded clause occurred in a 4-IP narrative in set II. Each set had a 4-IP passage with a dependent connection between two successive intonation phrases. One VSO structure occurred in set I. If syntactic complexity affected recall of the stimuli, then this more complex 4-IP passage should have been harder than its counterpart in set I. The normal passage with the VSO structure, however, yielded a geometric mean of .90 for $P(C_O)$, while its counterpart produced a value of .48. Across the altered prosody conditions, the two passages gave geometric means of .61 and .64, respectively.

Word frequency, however, did contribute to the difference between sets. We used a frequency dictionary (Kilgarriff, 1997) based on the 'demographic' spoken part of the British National Corpus. Over 4.2 million spoken tokens had been tagged grammatically and counted to make the frequency dictionary. It lacked two items that occurred in our stimuli ('botany' and 'unburnt'). Each was assigned a count of 0.5, and the total number of items in the corpus was increased by 1. For each word in the individual 3-IP, 4-IP, and 5-IP passages, we obtained a frequency count from the dictionary and converted it to a log-probability [$\log(p)$]. Minimum $\log(p)$, which was intended to indicate how unusual or surprising a passage might be, was the measure that we found best related to performance. It was lower for longer passages and for the stimuli in set I.

4.5 Effects of prosodic alterations.

Eight subjects heard the stimuli with a given type of prosodic alteration: monotone, pause-free, and monotone and pause-free combined. For a given subject, subtracting $P(C_O)$ for the altered passages of a given type and length from $P(C_O)$ for equally long normal passages would presumably quantify the effect of the prosodic manipulation. The difference in difficulty between sets I and II, however, would confound the results for the 3-IP and for the 4-IP passages.

To try to overcome this confound, we compensated for the unequal difficulty of sets I and II before testing the effects of the prosodic manipulations. Consider first the 3-IP passages in sets I and II. To equate for difficulty, we increased each $P(C_O)$ for the relatively hard 3-IP passages in set I by multiplying it by a factor greater than unity. The multiplier was simply the ratio of the geometric mean $P(C_O)$ across subjects and passages for the easier 3-IP stimuli in set II to the geometric mean score for the harder 3-IP stimuli in set I. A separate multiplier was calculated for the normal passages and for the prosodically altered conditions as a whole. This increased all individual $P(C_O)$ scores for set I. Similar adjustments were carried out with the data for the 4-IP passages of the two sets. The 5-IP passages in the two sets showed no signs of unequal difficulty, so data on them needed no adjustment.

After the adjustments, we employed one-tailed Monte Carlo Wilcoxon tests ($N = 8$ in each instance) to evaluate the differences between scores on the normal passages and on each type of prosodically manipulated passage. For each type of manipulation, this resulted in 3 tests, one for each passage length. We therefore calculated 9 tests in all.

Two of the 9 tests were significant. The 3-IP pause-free passages produced a significant decrease in $P(C_O)$, $z = 2.366$, $p < .01$. The scores for the 5-IP passages were significantly lower for the monotone than for the normal stimuli, $z = 1.965$, $p < .025$. In addition, the Wilcoxon test for the 5-IP pause-free passages produced a marginally significant result, $z = 1.823$, $p < .05$.

Figure 2 shows the data. The dependent variable is the adjusted difference between performances on the normal and altered passages. Each individual plot contains individual results for 8 subjects. Performance differences are shown for the 3-IP, 4-IP, and 5-IP passages under each type of prosodic manipulation. Each horizontal dashed line indicates a null difference. Results that yielded significant Wilcoxon tests are marked with an asterisk ($p < .025$) or a double asterisk ($p < .01$); a question mark indicates a marginally significant effect. Figure 2 makes apparent the lack of *consistent* significant differences between performances on the normal and the prosodically altered 3-IP and 4-IP passages. Indeed, the monotone-pause-free condition should have been the most difficult but yielded no significant differences whatsoever. Two of the three effects marked in Figure 2, however, arose from the 5-IP passages. To examine this further, one-tailed Monte Carlo Mann-Whitney tests (N=12 in each case) were applied to the data of Figure 1. For each set and each passage length, we evaluated the between-subjects difference between control data and results combined across prosodic manipulations. Of the six resulting tests, that for the 5-IP passages of set I proved significant, (U = 29.5, $p < .01$). Performance was poorer on the prosodically manipulated 5-IP stimuli.

The results in Figures 1 and 2 indicate that interference with prosody produced somewhat poorer recall of the 5-IP passages. The within-subjects and the between-subjects analyses for shorter passages did not yield similarly dependable effects.

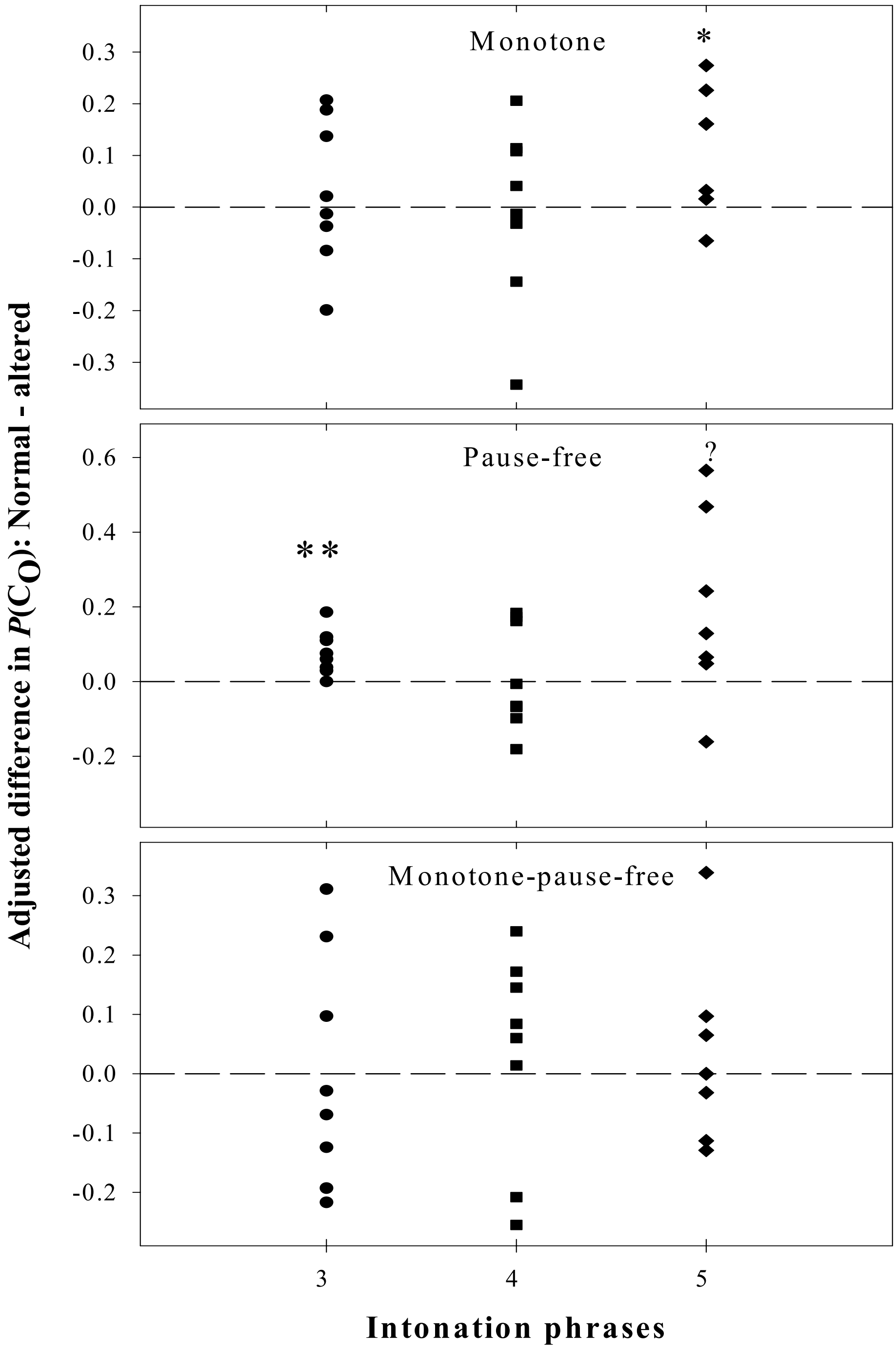

Fig. 2. Differences in $P(C_O)$ between normal and prosodically altered stimuli separated by passage length and type of prosodic manipulation, after adjusting for greater difficulty of set I. Each data set shows results for eight individual subjects. Significant differences identified by one asterisk ($p < .025$) or two asterisks ($p < .01$); question mark indicates marginally significant difference ($p < .05$).

5 Discussion

Memory for spoken prose was virtually perfect for passages containing up to 13 words. Performance deteriorated progressively as passage length increased beyond 13 words. Almost one-third of performances were perfect, however, on stimuli with 18 or 19 words. Passages with 25 or 26 words even drew a few perfect performances, but none occurred on the 31-word stimuli.

These results go well beyond the memory span measured with lists of independent items. The larger memory span for prose could result from two influences. First, due to its syntactic and semantic properties, prose may enable more efficient 'chunking' (Miller, 1956) of material in immediate memory. Second, prose may be transferred more efficiently out of immediate memory into a longer-term store.

The majority of imperfect performances on 3-IP, 4-IP, and 5-IP passages arose from omitting words in the middle of the passage. Transfer of the initial words from immediate into a longer-term memory could explain retention of the initial words in a passage. The final words would be held in immediate memory. Omission of words in the middle of a passage would then result.

Word frequency affected performance on the 3-IP, 4-IP, and 5-IP narratives. The difference in difficulty between sets I and II, normal or altered, sprang at least partly from differences in word frequency. The probability of the least frequent word in a passage clearly affected performance on that passage.

5.1 Effects of prosodic manipulations.

We measured the effects of each of three prosodic manipulations on memory for prose: removal of F0 variation, removal of pauses, and removal of both. After adjusting for the greater difficulty of set I, only the 3-IP pause-free passages demonstrated significantly worse performance than did the normal 3-IP passages. The prosodic manipulations had no other effects on memory for the 3-IP and the 4-IP passages. The one positive result on the 3-IP monotone passages may be an artefact of our method of compensating for differences in difficulty between sets I and II.

The 5-IP monotone passages, however, proved significantly harder than the 5-IP normal passages, and the 5-IP pause-free passages were marginally harder than the normal narratives. In line with these results, between-subjects tests showed that the combined 5-IP altered passages of set I were harder to recall than were the normal stimuli. No other similar between-subjects comparisons were significant. Differences between subgroups of listeners exposed to the different types of prosodic manipulation may explain why performance was no different from normal for the 5-IP monotone-pause-free stimuli.

Our generally negative results on the prosodically altered 3-IP and 4-IP passages agree with the findings of Stine and Wingfield (1987). They disagree, however, with previous positive reports on mnemonic effects of intonation (Leonard, 1974; O'Connell, Turner, & Onuska, 1968; Zurif & Mendelsohn, 1972; Paris, Thomas, Gilson, & Kincaid, 2000). All those experiments, however, used relatively short grammatically anomalous sentences, nonsense strings containing English articles and bound morphemes, or utterances with sudden changes of theme. Our straightforward narrative stimuli were very different. The negative findings on pause-free and monotone-pause-free narratives in our experiment also diverge from the results reported by Huttenlocher and Burke (1972), Frankish (1985, 1989), and Martin (1968). All these experimenters found that pauses affect memory for speech. Huttenlocher and Burke and Frankish, however, used lists of random digits, while Martin found that unusually long pauses after principal words decreased recall of grammatical, anomalous, and scrambled utterances.

The differences between our negative results and previous positive findings on the effects of prosody on memory for speech seem to depend on the use of different kinds of stimuli. We employed continuous prose with a relatively simple structure. The previous positive experiments, however, used thematically disjoined or anomalous utterances, lists of unrelated items, nonsense strings, long pauses, or high speech rates. Such conditions are unusual in daily life. They seem to bring out effects of intonation or pauses on memory. When speech contains routine properties and material, however, prosody seems of little importance for memory. The redundancy that other factors afford readily overcomes any effects of interference with prosody. When the going gets difficult, however, prosody may affect memory for speech.

A variety of facts supports this argument. First, increased speech rates facilitate demonstrations of effects of prosody on memory for prose (Stine & Wingfield, 1987; Wingfield, 1975). Second, Paris, Thomas, Gilson, & Kincaid (2000) reported that their variant of citation form prosody interfered with memory for passages of 15-20 words that contained sudden shifts in theme. In contrast, Stine and Wingfield (1987) found little such effect for continuous 16-word prose passages with no thematic disjunctions. Third, cross-sentence splicing experiments (Darwin, 1975; Wingfield, 1975; Wingfield & Klein, 1971) create unusual stimuli. Such experiments bring out a role for prosody in memory for prose. Fourth, prosody facilitates syntactic disambiguation (see Cutler, Dahan, & van Donselaar, 1997), which necessarily involves material that is hard to process. Fifth, only our longest, 5-IP stimuli yielded positive findings that were consistent in both within- and between-subjects analyses.

The data in Figure 3 give further support to our argument. Total counts of words in correct order were obtained across subjects for each normal passage. We also did this for each passage combining over the three altered conditions. This yielded $P(C_{OG})$, the grand proportion of ordered words correct for a given passage under either normal or pooled altered conditions. In Figure 3, $\log[P(C_{OG})]$ for prosodically altered passages is plotted on the vertical axis against $\log[P(C_{OG})]$ for normal passages. Points that fall around the diagonal line indicate no essential effect of altered prosody. Four points clearly depart from the diagonal. All have relatively low values of $\log[P(C_{OG})]$ for normal passages, compared to the six points clustered at the upper right of the figure. The four normal narratives that yielded these data therefore tended to be relatively difficult even prior to any prosodic changes. The prosodic alterations made them even harder to remember. Three of the four passages contained five intonation phrases.

The most puzzling result in Figure 3 is the poor performance on a 3-IP passage (filled circle). Nothing seems to differentiate it from the other 3-IP passages. In addition, performance was equally good on the normal and altered versions of one 5-IP passage. It seems quite comparable to the other 5-IP passages that yielded reduced performance when altered. Factors other than minimum $\log(p)$ and passage length seem to underlie these two discrepancies.

In summary, prosody may facilitate memory for speech only as processing becomes increasingly difficult. Previous research has shown that difficult conditions include high speech rates, ambiguous or even abnormal syntax, thematic discontinuity, and unstructured sequences of items. Our findings now add two additional factors to this list: increasing utterance length and lower word frequency. As long as processing demands are sufficiently low and redundancy is sufficiently high, however, prosody seems unnecessary for remembering spoken material.

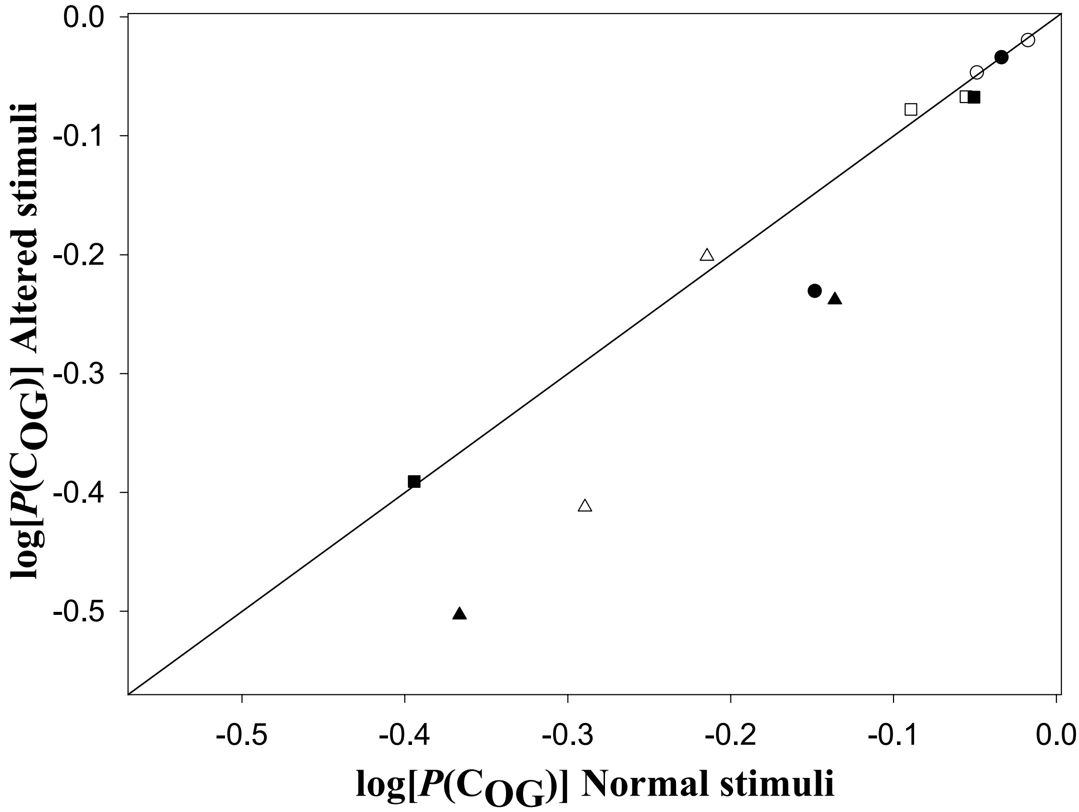

Fig. 3. Log[$P(C_O)$] for prosodically altered passages plotted against log[$P(C_O)$] for normal passages. Filled symbols: set I; open symbols: set II. Circles, squares, and triangles: 3-IP, 4-IP, and 5–IP stimuli, respectively.

References

Berry, C. 1994. *Your voice and how to use it*. London: Virgin Books: 125-126.

Boersma, P. and Weenink, D. 1996. PRAAT: A system for doing phonetics by computer. Report of the Institute of Phonetic Sciences, University of Amsterdam 132.

Carpentier, F. and Moulines, E. 1990. Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Communication*, *9*: 453-457.

Cutler, A., Dahan, D., and van Donselaar, W. 1997. Prosody in the comprehension of spoken language: A critical review. *Language and Speech 40*: 141-201.

Darwin, C. J. 1975. On the dynamic use of prosody in speech perception. In A. Cohen and S. G. Nooteboom Eds., *Structure and process in speech perception*. Berlin: Springer-Verlag: 178-193.

Frankish, C. 1985. Modality-specific grouping effects in short-term memory. *Journal of Memory and Language 24*: 200-209.

Frankish, C. 1989. Perceptual organization and precategorical acoustic storage. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *15*: 469-479.

Huttenlocher, J. and Burke, D. 1976. Why does memory span increase with age? *Cognitive Psychology*, *8*: 1-31.

Institute of Electrical and Electronics Engineers. 1969. *IEEE recommended practice for speech quality measurements* IEEE No. 297. New York: Author.

Kilgarriff, A. 1997. Putting frequencies in the dictionary. *International Journal of Lexicography 10*: 135-55.

Ladd, D. R. 1996. *Intonational phonology*. Cambridge: Cambridge University Press.

Leonard, L. B. 1974. The role of intonation in the recall of various linguistic stimuli. *Language and Speech*, *16*: 327-335.

Mandel, S. 1993. *Effective presentation skills* Revised ed.. Menlo Park, CA: Crisp Publications.

Martin, J. G. 1968. Temporal word spacing and the perception of ordinary, anomalous, and scrambled strings. *Journal of Verbal Learning and Verbal Behavior, 7*: 154-157.

Miller, G. A. 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review 63*: 81-97.

O'Connell, D. C., Turner, E. A. and Onuska, L. A. 1968. Intonation, grammatical structure, and contextual association in immediate recall. *Journal of Verbal Learning and Verbal Behavior, 7*, 110-116.

Paris, C. R., Thomas, M. H., Gilson, R. D. and Kincaid, J. P. 2000. Linguistic cues and memory for synthetic and natural speech. *Human Factors*, *42*, 421-431.

Robinson, G. M. 1977. Rhythmic organization in speech processing. *Journal of Experimental Psychology: Human Perception and Performance 3*: 83-91.

Saito, S. 1998. Effects of articulatory suppression on immediate serial recall of temporally grouped and intonated lists. *Psychologia*, *41*: 95-101.

Speer, S. R., Crowder, R. G. andThomas, L. M. 1993. Prosodic structure and sentence recognition. *Journal of Memory and Language*, *32*, 336-358.

Stine, E. L., & Wingfield, A. 1987. Process and strategy in memory for speech among younger and older adults. *Psychology and Aging*, *2*: 272-279.

Syntrillium Software Corporation, 1996. *Cool96*.

Wechsler, D. 1981. *Wechsler adult intelligence scales revised.* San Antonio TX: Psychological Corporation and Harcourt Brace Jovanovich.

Wingfield, A. 1975a. Acoustic redundancy and the perception of time-compressed speech. *Journal of Speech and Hearing Research 18*: 96-104.

Wingfield, A. 1975b. The intonation-syntax interaction: prosodic features in the perceptual processing of sentences. In A. Cohen and S. G. Nooteboom Eds., *Structure and process in speech perception.* Berlin: Springer-Verlag: 148-156.

Wingfield, A. and Klein, J. F. 1971. Syntactic structure and acoustic pattern in speech recognition. *Perception and Psychophysics 9*: 23-25.

Wingfield, A., Lahar, C. J., and Stine, E. A. L. 1989. Age and decision strategies in running memory for speech: Effects of prosody and linguistic structure. *Journal of Gerontology: PSYCHOLOGICAL SCIENCES 44*: 106-113.

Zurif, E. B. and Mendelsohn, M. 1972. Hemispheric specialization for the perception of speech sounds: The influence of intonation and structure. *Perception and Psychophysics 72*: 329-332.