Research Report

REPETITION DEAFNESS: Repeated Words in Computer-Compressed Speech Are Difficult to Encode and Recall

Michelle D. Miller and Donald G. MacKay
University of California, Los Angeles

Abstract—*This research demonstrates a new cognitive phenomenon known as repetition deafness, a difficulty in immediate recall of repeated words in computer-compressed speech. Sixty-four subjects heard sentences and lists at four speeded rates: 70, 55, 35, and 28 ms/phoneme. Each target word in the materials followed a pretarget word that was either identical (repeated-target condition) or different (unrepeated-target condition), and targets were harder to recall when repeated than unrepeated. Repetition deafness was rate-limited, occurring only with rapid rates of presentation (55 ms/phoneme or less), and decreased in magnitude as structure increased from lists to sentences. Implications for current theories of repetition deficits are discussed.*

According to recent literature (e.g., Kanwisher & Potter, 1989), repetition deafness (RD) is a hypothetical phenomenon that does not occur. What does occur is its visual counterpart, repetition blindness (RB). RB refers to the reduced probability of detecting or recalling a word (or letter) due to prior occurrence of the same word (or letter) in a rapidly presented list or sentence (Bavelier & Potter, 1992; Kanwisher, 1987, 1991; Kanwisher & Potter, 1989, 1990; MacKay, 1969). RB is especially marked with rapid serial visual presentation (RSVP), in which words appear one at a time for a fixed interval at the same location on a computer monitor. However, Kanwisher and Potter (1989) failed to find RD with rapid auditory presentation of sentences and hypothesized that RB occurs at an early or primitive stage of processing that can accurately encode repetition in audition, but not vision. If true in general, this visual specificity hypothesis is important, ruling out theories such as node structure theory (NST; MacKay & Miller, 1992a), which explains RB in terms of connection formation processes that are influenced by presentation rate, structure (lists vs. sentences), and lag (number of words that intervene between repeated words), but are not specific to vision. However, studying repetition deficits may further understanding of how connections are formed—the most fundamental aspect of learning and memory—if modality-free theories such as NST are correct.

We had several reasons for questioning the visual specificity hypothesis. One is that previous work (Wickelgren, 1965, 1966) has demonstrated effects of repetition in immediate recall of spoken digits, and these effects resemble RD. The present study attempted to reconcile this possible conflict between Wickelgren's and Kanwisher and Potter's studies, which differed in materials (lists of digits vs. meaningful sentences), analytic procedures, and important variables (e.g., lag and rate of presentation). Another reason for questioning the visual specificity hypothesis is that RB is unresponsive to changes in visual characteristics of repeated words or letters, such as case (e.g., *a* vs. *A;* Kanwisher, 1987; Bavelier & Potter, 1992), spatial location (Kanwisher & Potter, 1990), and visual format (e.g., *nine* vs. *9;* Bavelier & Potter, 1992). However, RB is responsive to changes in phonology (see Bavelier & Potter, 1992), and because phonology is clearly processed during auditory speech perception, the hypothesized nonoccurrence of RD is puzzling.

STUDY 1: REPETITION DEAFNESS IN WORD LISTS

Study 1 tested the visual specificity hypothesis for rapid auditory presentation of words in lists, using the ingenious procedure developed by Kanwisher (1987). Recall of identical repeated versus unrepeated words was compared in almost identical contexts: Lists containing repeated versus unrepeated targets differed only in a single pretarget word. For example, recall of the target *four* was compared in "one six *nine four* two three" (unrepeated-target version, target and pretarget italicized) and "one six *four four* two three" (repeated-target version); a subject experiencing RD might report hearing only one of the two *four*'s in the latter version. Study 1 also manipulated presentation rate and lag. Although Kanwisher (1987, Experiment 1) systematically manipulated rate and lag using RSVP, her subjects' task was to report what word was repeated in lists that contained only targets that were repeated. The relative retrievability of repeated versus unrepeated targets, although essential for evaluating effects of lag and presentation rate on repetition deficits, has never been examined.

Method

Sixty-four students in introductory psychology classes received partial course credit for participation. All were fluent speakers of English and reported normal hearing.

We constructed repeated- and unrepeated-target versions of 16 lists (see the appendix for details): Pretargets and targets were the same for repeated-target versions and different for unrepeated-target versions (see Table 1). Repeated targets were either repeated immediately (0-lag, e.g., "bush bush") or separated by one short word (1-lag, e.g., "mat plan mat").[1]

1. Note that immediate, or 0-lag, repetition is practicable for rapid auditory presentation, but not RSVP: Because an RSVP word falls without lag on top of its predecessor, 0-lag repetition in RSVP corresponds to a single word presented for twice as long.

Address correspondence to Donald G. MacKay, Psychology Department, University of California, Los Angeles, Los Angeles, CA 90024.

Table 1. *Example materials for Study 1 (lists) and Study 2 (sentences)*

Kind of stimulus	Example
	Lists (Study 1)
Repeated-target list	burden *mat* plan *mat* thunder hand
Unrepeated-target list	burden *age* plan *mat* thunder hand
	Sentences (Study 2)
Repeated-target sentence	They wanted to play *sports* but *sports* were not allowed
Unrepeated-target sentence	They wanted to play *ball* but *sports* were not allowed
Ungrammatical filler	When we went to the it was very crowded
Normal filler	When we went to the store it was very crowded

Note. Targets and pretargets are italicized.

Subjects were informed that they would hear word strings played at various speeded rates over headphones and that they were to repeat each string aloud as soon as it ended. Subjects were warned that one or more words in a string would sometimes be repeated and that they were to report each word in order, and as many times as they heard it. Lists were presented at four rates (28, 35, 55, and 70 ms/phoneme[2]) that varied orthogonally with list type (see the appendix for details).

Results and Discussion

Figure 1 (left panel) plots correct recall of repeated and unrepeated targets as a function of presentation rate. The effect of rate was significant ($F[3, 189] = 68.26$, $p < .001$), and across all rates, subjects correctly recalled unrepeated targets significantly more often than repeated targets (41% vs. 29%; $F[1, 63] = 10.28$, $p < .05$). This inhibitory effect of repetition reopens issues previously considered closed, for example, the question of whether earlier results (Wickelgren, 1965, 1966) are relevant to current procedures and theories of repetition deficits.

Lag had no effect on RD (i.e., correct report for unrepeated minus repeated targets), with no significant differences at any rate or combination of rates (highest $t[1, 63] = 0.86$, $p = .393$). However, RD interacted with rate ($F[3, 189] = 6.93$, $p < .001$) because of a floor effect at 28 ms/phoneme and absence of RD at 70 ms/phoneme (see Fig. 1). This latter finding suggests that RD, like RB, is a rate-limited phenomenon that does not occur at rates slower than about 70 ms/phoneme. Absence of RD at our slowest rate also suggests that RD is not attributable to response bias or other subject strategies (which should be more rather than less effective with increased time to apply them).

2. Our ms/phoneme rates used Anderson's (1991) definition of phonemes and corresponded to 90, 113, 177, and 225 ms/word, or 11.11, 8.85, 5.65, and 4.44 words/s.

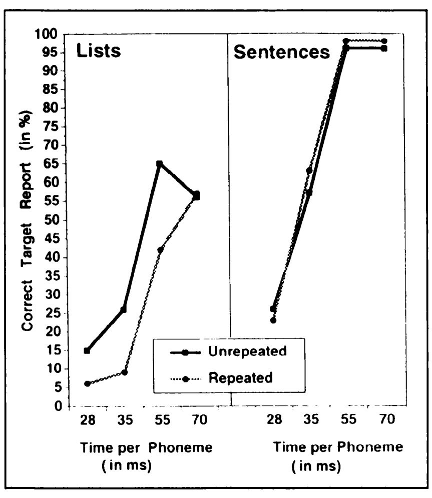

Fig. 1. Percentage correct target report for lists (left panel; Study 1) and sentences (right panel; Study 2) as a function of repetition condition and presentation rate.

STUDY 2: REPETITION DEAFNESS IN SENTENCES

Because RD is clearly sensitive to temporal factors, and because Kanwisher and Potter (1989) looked for RD only in sentences at a single rate (133 ms/word), with relatively long lags (one to three words) separating repeated words, it is possible that RD may occur in sentences at more rapid rates or with shorter lags. This possibility seemed all the more likely because, relative to reading, auditory sentence processing benefits from more extensive prior practice (MacKay, 1981; 1982; 1987, p. 72) and faster rates of processing (Marslen-Wilson, 1989). Study 2 therefore attempted to replicate Kanwisher and Potter's (1989) results for sentences using shorter lags and a range of rapid rates to determine whether RD would occur for some rate or some lag.

Method

Subjects and method were identical in Studies 1 and 2 except that stimuli were sentences. Indeed, both experiments were run in the same session, separated by a 4-min break, with Study 2 either preceding or following Study 1, in counterbalanced order. However, we present the two studies separately here to clarify exposition of their differing results. For example, effects of repetition differed significantly across the two experiments ($F[1, 63] = 15.09$, $p < .001$, with type of material as a within-subjects factor and presentation order as a between-subjects factor).

Subjects heard and immediately recalled 32 strings, 16 experimental sentences randomly interspersed among 16 fillers (see Table 1). Use of words from lists in Study 1 was avoided in constructing the sentences. The experimental sentences averaged nine words long and were grammatically acceptable; repeated- and unrepeated-target versions of a sentence were similar in meaning and syntax. Eight fillers were ungrammatical, and eight were normal sentences with novel syntax and no repeated words. Fillers ensured that subjects would not focus on repetitions and could sometimes expect ungrammatical sentences, as might occur because of RD. Lags, rates, and compression procedures were the same as in Study 1.

Instructions were the same as in Study 1 except that subjects were not told that sentences would contain repeated words. Subjects were asked for verbatim recall without paraphrasing or "fixing the sentences up, even though some might sound strange or ungrammatical." Sessions began with four representative practice sentences, and counterbalancing was the same as in Study 1.

Results

Scoring procedures were the same as in Study 1. Figure 1 (right panel) shows mean percentage of correct report for repeated and unrepeated targets as a function of presentation rate. The main effect of rate was significant ($F[3, 189] = 236.83, p < .05$), but neither the overall effect of repetition ($F[1, 63] = 2.02, p > .16$) nor the Repetition × Rate interaction ($F[3, 189] = 2.00, p > .116$) was significant. Because of a ceiling effect, no evidence of RD was possible at the slowest presentation rates (55 and 70 ms/phoneme), but the pattern was similar for faster rates: Subjects correctly reported repeated targets at least as often as unrepeated targets averaged across rates and lags (71% vs. 69%). However, a post hoc test indicated a significant effect of lag on RD at 35 ms/phoneme, with greater RD for 0- than 1-lag targets (Walsh test, $z = -3.20, p < .002$). The same test indicated that for 0-lag targets presented at 35 ms/phoneme, unrepeated targets were recalled significantly more often than repeated targets ($z = -4.28$, $p < .001$), and this RD effect for 0-lag targets was at least as large (29%) as the largest RD effect for lists in Study 1 (23% at 55 ms/phoneme).

GENERAL DISCUSSION

Results for Study 2 suggest that RD may occur for sentences with immediately repeated targets presented at a relatively rapid rate (35 ms/phoneme). However, RD for sentences is clearly nonexistent at slower rates and longer lags, replicating and extending Kanwisher and Potter's (1989) observation for 1- to 3-lag sentences presented at 133 ms/word. An adequate theory must therefore explain why RD is in general weaker for sentences (Study 2) than lists (Study 1).[3] The NST (MacKay, 1987, 1990; MacKay & Miller, 1992a) provides one such explanation. Consider first the main basis in NST for repetition deficits in sentences. Under NST, a single node or set of nodes represents a lexical concept in long-term memory, and encoding proceeds in part by forming connections from already existing lexical nodes to phrase-level nodes (MacKay & Miller, 1992a). For example, consider the nodes illustrated in Figure 2 for comprehending the sentence "They wanted to play sports, but sports were not allowed." The single lexical node for the repeated concept (*sports*) must quickly connect with two phrase nodes, for the verb phrase "to play sports" and the proposition "sports were not allowed" (see Fig. 2). Lexical nodes for unrepeated concepts (e.g., *to, were,* and *allowed*) become connected with only a single phrase node, and these one-to-one connections can be formed quickly and in parallel, whereas the one-to-many connections from repeated concepts to phrase nodes require more time because they must be formed in sequence: A single node can generate connections to only one other node at a time (MacKay, 1990). Given the time pressure of rapid presentation, the first connection from a repeated concept may be formed, but not the second, resulting in RD, a failure to encode and retrieve the second instance of a concept, together with its phonology. This explains why slower presentation rates eliminate RD: When sufficient time separates pretarget from target, repeated and unrepeated words become equally easy to encode and retrieve.

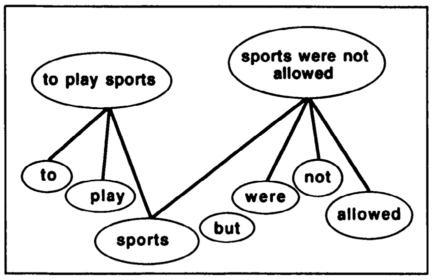

Fig. 2. Selected nodes for encoding the auditory sentence "They wanted to play sports, but sports were not allowed." Note that a single lexical node represents the repeated word "sports." Acoustic and phonological nodes are omitted from the figure.

Thus, according to NST, factors that facilitate the rapid linking of words to phrases or, in the case of lists, chunks should diminish the likelihood of repetition deficits. One of these factors is syntax. Sentences (unlike lists) contain syntactic cues that indicate in redundant, nested fashion how words link together into phrases. For example, consider the words "to play sports" in Figure 2. The infinitival determiner "to" signals that a verb follows and calls for links with a verb-phrase node, to which the object "sports" also becomes linked. By enabling rapid formation of word-to-phrase links, syntax reduces the probability of RD in sentences relative to lists.

Another factor that can differentially reduce RD for sentences relative to lists is prosody, acoustic cues to phrase structure such as timing, stress, pitch, and intonation (Levelt, 1990, pp. 365–412). Like syntax, prosodic cues occur in sentences but not lists and, under NST, reduce RD by making it easier for listeners to determine how words combine together into phrases, thereby increasing the time available for forming the one-to-many links required to encode and retrieve repeated words.

Prior practice, variability, and interference are a final set of factors that can differentially reduce RD for sentences relative to lists. Unlike the sometimes highly practiced connections for familiar phrases such as *to play sports*, connections for chunks in lists are novel, unpracticed, and variable; their precise nature can vary from list to list, subject to subject, and trial to trial. Some subjects may chunk lists by linking each word with superordinate nodes representing

3. The list–sentence difference is difficult to explain under a memory load hypothesis, whereby repetition deficits are inversely related to short-term memory capacity (Park & Kanwisher, 1991). The greater RD for lists than sentences does correlate inversely with one index of memory capacity, namely, overall level of correct recall, which was greater for sentences than lists (70% vs. 34%). However, if level of correct recall is equated for lists and sentences (e.g., for subjects perceiving lists at 35 ms/phoneme in Study 1 and sentences at 28 ms/phoneme in Study 2; see Fig. 1), lists but not sentences exhibit RD, suggesting that memory capacity is irrelevant to the greater RD for lists.

ordinal positions in the list. Other subjects may encode lists into beginning, middle, and end chunks, the size of which will vary and cannot normally be determined until the end of the list (Wickelgren, 1979). Moreover, consistent application of either of these chunking strategies will result in cross-list interference (Wickelgren, 1979) that can further slow the chunking process and increase the likelihood of RD in lists relative to sentences.

Acknowledgments—This report was presented as a poster to the 32nd Annual Meeting of the Psychonomic Society, San Francisco, November 1991, and as a colloquium at UCLA and UC Berkeley in October 1992. The authors gratefully acknowledge support from National Institute on Aging Grant 1 R01 AG 09755 to Donald G. MacKay and a National Science Foundation Graduate Fellowship to Michelle D. Miller. We also thank Deborah Burke for loaning various pieces of equipment essential to this research and for helpful comments on an earlier version of this article. Data for a random half of the subjects reported here have been used for comparison with 32 older subjects in a subsequent report on effects of aging on repetition deafness (MacKay & Miller, 1992b).

REFERENCES

Anderson, J.R. (1991). *Cognitive psychology and its implications* (2nd ed.). San Francisco: Freeman.

Bavelier, D., & Potter, M.C. (1992). Visual and phonological codes in repetition blindness. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 134–147.

Francis, W.N., & Kuçera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.

Kanwisher, N.G. (1987). Repetition blindness: Type recognition without token individuation. *Cognition, 27*, 117–143.

Kanwisher, N.G. (1991). Repetition blindness and illusory conjunctions: Errors in binding visual types with visual tokens. *Journal of Experimental Psychology: Human Perception and Performance, 17*, 404–421.

Kanwisher, N.G., & Potter, M.C. (1989). Repetition blindness: The effects of stimulus modality and spatial displacement. *Memory & Cognition, 17*, 117–124.

Kanwisher, N.G., & Potter, M.C. (1990). Repetition blindness: Levels of processing. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 30–47.

Levelt, W.J.M. (1990). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

MacKay, D.G. (1969). The repeated letter effect in the misspellings of dysgraphics and normals. *Perception & Psychophysics, 5*, 102–106.

MacKay, D.G. (1981). The problem of rehearsal or mental practice. *Journal of Motor Behavior, 13*, 274–285.

MacKay, D.G. (1982). The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behavior. *Psychological Review, 89*, 483–506.

MacKay, D.G. (1987). *The organization of perception and action: A theory for language and other cognitive skills*. New York: Springer-Verlag.

MacKay, D.G. (1990). Perception, action, and awareness: A three body problem. In W. Prinz & O. Neumann (Eds.), *Relationships between perception and action* (pp. 269–303). Berlin: Springer-Verlag.

MacKay, D.G., & Miller, M. (1992a, November). *Repetition deafness and blindness: Rapidly processed repeated concepts are difficult to encode*. Paper presented at the 33rd Annual Meeting of the Psychonomic Society, St. Louis.

MacKay, D.G., & Miller, M. (1992b, April). *The repetition deafness phenomenon in young and older adults*. Paper presented at the 4th Biennial Cognitive Aging Conference, Atlanta.

Marslen-Wilson, W. (1989). Access and integration: Projecting sound onto meaning. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 3–24). Cambridge, MA: MIT Press.

Park, J., & Kanwisher, N. (1991, June). *(Non)determinants of repetition blindness*. Paper presented at the Western Attention Conference, University of California, Davis.

Wickelgren, W.A. (1965). Short-term memory for repeated and non-repeated items. *Quarterly Journal of Experimental Psychology, 17*, 14–25.

Wickelgren, W.A. (1966). Associative intrusions in short-term recall. *Journal of Experimental Psychology, 72*, 853–858.

Wickelgren, W.A. (1979). *Cognitive psychology*. Englewood Cliffs, NJ: Prentice-Hall.

(RECEIVED 1/25/93; REVISION ACCEPTED 5/26/93)

APPENDIX: METHODOLOGICAL DETAILS, STUDY 1

Materials

Lists contained six, seven, or eight unrelated content words that were equated for average frequency using Francis and Kuçera (1982). Lists containing 0-lag and 1-lag repetitions were identical in length ($M = 7$ words). Pretarget and target words were closely matched in frequency and average length (2–4 phonemes) and never occurred first or last in a list. Repetition of words across lists was avoided.

Compression Procedures

To achieve the four presentation rates, a speaker of standard American English (M.M.) repeatedly adjusted her output until list duration corresponded to 140 ± 3 ms/phoneme (a slow rate) and 110 ± 3 ms/phoneme (a brisk rate), measured digitally using a MacRecorder installed in a Macintosh SE computer. The software (SoundEdit 2.0.1) sampled the input 22,000 times per second and then compressed the digitized record by scanning the samples, identifying recurrent or nearly recurrent samples, deleting these redundant samples until 50% of the total samples remained, and abutting these residual samples. This procedure accelerated the original recordings by a factor of 2, but gave fairly high intelligibility and left pitch unchanged. The two initial recordings of the materials were compressed twice in this manner, once for the 70 and 55 ms/phoneme rates, and again for the 35 and 28 ms/phoneme rates.

Procedure and Design

After computer compression, lists were transferred to eight audiotapes, with each subject randomly assigned to hear one of the eight tapes. Each tape contained 16 lists (8 repeated-target versions, with 2 presented at each of the four rates, plus unrepeated versions of the remaining 8 lists, likewise with 2 presented at each of the four rates), with versions and rates counterbalanced across subjects. A 250-ms tone alerted subjects to an upcoming list, which arrived 100 ms later, and subjects responded during an 8,000-ms silent period that followed each list.

Order of rates and repetition conditions varied randomly across the eight tapes, except that average serial position for repeated- and unrepeated-target lists was kept constant. Tapes were played to subjects via headphones linked to a four-track tape deck. Subject output was tape-recorded, but the experimenter (M.M.) also transcribed responses "on line" onto scoring sheets that were unique for each tape. Each session began with a practice tape that contained two repeated- and two unrepeated-target lists, one at each of the four rates in "random" order, except that the fastest rate never came first.

Analyses

Each trial was scored for correct target recall (i.e., inclusion of the target word in a subject's response). To distinguish target from pretarget in recall of 0-lag lists, both target and pretarget had to be included in order to count as correct target recall, for both repeated- and unrepeated-target versions. For 1-lag lists, serial position of a target relative to other recalled words almost always disambiguated whether target or pretarget had been recalled. In rare instances when ambiguity remained, repeated targets were scored as correct even though only one was recalled, a highly conservative procedure that works against observing RD (i.e., decreased report of repeated relative to unrepeated targets). Planned analyses were univariate analyses of variance, and post hoc tests were t tests unless indicated otherwise.