JOURNAL OF MEMORY AND LANGUAGE 34, 63–88 (1995)

Is the Articulatory Loop Articulatory or Auditory? Reexamining the Effects of Concurrent Articulation on Immediate Serial Recall

PRAHLAD GUPTA AND BRIAN MACWHINNEY

Carnegie Mellon University

Results from the paradigm of *immediate serial recall* form the basis of the influential "articulatory loop" model of auditory–verbal short-term memory (Baddeley, 1986). Central to the development of these ideas have been results obtained in immediate serial recall under the condition of *concurrent articulation*. We reexamine the effects of concurrent articulation and show that findings from immediate serial recall do not uniquely support the articulatory rehearsal hypothesis: the data can be accounted for by assuming a purely auditory rehearsal process. The question of whether the rehearsal process in fact has an "articulatory" component or is purely "auditory" has significance beyond the immediate domain of working memory, and makes contact with a number of important issues concerning phonological processing. We describe a series of experiments aimed at discriminating between the two hypotheses. Our results support an articulatory component in rehearsal, but also indicate that auditory interference plays a significant, but previously unrecognized, role in the concurrent articulation effect.

One easy way of measuring immediate verbal memory is to ask subjects to recall strings of words. In the *immediate serial recall* (ISR) task, the subject is presented with sequences of *unrelated* verbal items (such as digits or words) and is required to recall the sequence in correct order, immediately following its presentation. Presentation of the list may be either auditory or visual. The subject may be required to respond in speech, in writing, or in some other fashion.

This task has played a central role in development of the *working memory* model of Baddeley (1986). In this model, working memory has an auditory–verbal component, which underlies performance of ISR tasks. This component has been termed the "articulatory loop" and, more recently, the "phonological loop."

The articulatory loop system is comprised of two elements (Baddeley, 1990a): (1) a phonological input store within which memory traces will fade if not revived within 1–2 s, and (2) an articulatory control process ("articulatory rehearsal") which serves to maintain memory traces within the phonological store by means of subvocal rehearsal. The control process also provides a way for visually presented items to be fed into the phonological store, provided they are capable of being encoded phonologically and subvocalized. Subvocal rehearsal is a process which operates in real time, with long words taking longer than short.

It appears well established that the rehearsal process cannot be articulatory in the sense of involving the speech musculature; rather, "articulatory" connotes that the process is somehow involved in phonological production *planning* (Baddeley, 1990a), and it is in this sense that we use the term "articulatory" or "output" throughout. Nevertheless, even in the current for-

We thank Julia Evans and Akira Miyake for invaluable input on many points; Alan Baddeley, Rita Sloan Berndt, Jay McClelland, and Henry Roediger, III, for helpful comments on earlier versions of this work; and Sonya Brady for assistance in conducting experiments. Of course, none of them is responsible for any errors or inconsistencies in this paper, nor do they necessarily agree with the opinions expressed here. Correspondence and reprint requests should be addressed to Prahlad Gupta, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213. E-mail: prahlad@cs.cmu.edu.

0749-596X/95 $6.00

mulation of the articulatory loop model, there are major questions that remain, concerning the nature of this rehearsal process that refreshes phonological representations, and this is the question on which we focus in this paper.

The first section of the paper outlines the data and interpretations that constitute the articulatory loop model. In the second section, we show that the accepted interpretations of the data as evidence for an "articulatory" rehearsal process are problematic even with respect to the narrower sense of the term; we show that the data can be accounted for by an alternative "auditory" rehearsal hypothesis. The next three sections of the paper present experimental studies aimed at discriminating between these two accounts. To anticipate, these studies provide new evidence supporting the articulatory rehearsal hypothesis, but also highlight factors that have been ignored in interpreting the data. In the fourth section we turn to a consideration of remaining confounds, which leads into questions about the nature of auditory imagery; we suggest that these confounds highlight the need for supplementing experimental work with neuropsychological and neuroimaging techniques.

The Articulatory Loop

A number of phenomena have motivated and influenced thinking about the articulatory loop. Of these, the following are most pertinent to the present discussion, comprising the key findings that the articulatory loop model was designed to explain. It is the ability to explain these phenomena that is regarded as evidence of the model's success (e.g., Baddeley, 1990a).

1. The Phonological Similarity Effect

In immediate serial recall of lists of words, sequences of similar sounding words are recalled in correct order much less frequently than sequences of dissimilar words of comparable frequency and length. This result holds irrespective of whether the recall stimuli are presented in the auditory or the visual modality (Baddeley, 1986).

2. The Irrelevant Speech Effect

Immediate serial recall of lists of items is disrupted by the presentation of irrelevant spoken material not produced by the subject, despite the fact that the subject is free to ignore this material. The disruptive characteristics of the unattended material appear to be primarily phonological, with nonsense syllables being just as disruptive as meaningful words. Again, this effect obtains regardless of modality of presentation (Baddeley, 1990a).

3. The Word Length Effect

Immediate serial recall performance of word sequences deteriorates as the constituent words in the sequence become longer, whether they are presented auditorily or visually (Baddeley et al., 1975).

4. The Concurrent Articulation Effect

When the subject is required to engage in *concurrent articulation,* i.e., to articulate an irrelevant sound during list presentation,[1] immediate serial recall is markedly impaired, under both visual and auditory presentation of stimuli (Baddeley, 1990a).

5. Interactions

Each of the first three effects has certain further interactions with the fourth. With auditory presentation, the *phonological similarity effect* is still observed under concurrent articulation; that is, when subjects perform ISR under concurrent articulation, they recall sequences of similar sounding words in correct order much less frequently than sequences of dissimilar words, just as in the phonological similarity effect *without* concurrent articulation. Similarly, under concurrent articulation, the *irrelevant speech effect* also seems to be somewhat

[1] We use the more neutral term "concurrent articulation" for what has usually been called "articulatory suppression" in the working memory literature.

preserved, at least under certain conditions (Hanley & Broadbent, 1987); that is, subjects' recall performance under concurrent articulation is further impaired by the presence of irrelevant speech. However, under concurrent articulation, the *word length effect* is abolished if concurrent articulation is required during both list input and recall (Baddeley et al., 1984). That is, when subjects perform ISR under concurrent articulation, it no longer matters whether the words in the list are long or short.

With visual presentation of recall stimuli, concurrent articulation abolishes the effects of phonological similarity, irrelevant speech, and word length (Baddeley, 1990a).

The working memory model account of these various phenomena (e.g., Baddeley, 1990a) is outlined below. A key point to note about this account is that there is an "articulatory rehearsal" process, i.e., an *articulation-based process* that can be invoked to "refresh" *phonological representations* of the recall items.

The *phonological similarity effect* is accounted for by locating phonological errors at the "phonological store", where phonologically similar words will have similar activated representations. As a result, the candidate representations for a particular serial recall position (e.g., for the *third* position to be recalled) are similar, competing, and thus confusable. Maintaining a sequence of such representations in correct order will therefore be more difficult and error-prone than maintaining a sequence of dissimilar representations in correct order.

The *irrelevant speech effect* is accounted for by assuming that auditory stimuli automatically enter the phonological store, where they interfere with the memory traces of recall stimuli. These recall stimuli are assumed to enter the phonological store directly in the case of auditory presentation, and via phonological recoding, if presentation is visual.

The *word length effect* is viewed as a direct consequence of the articulatory nature of rehearsal. Longer words take longer to articulate, and hence longer to rehearse, than shorter words. Consequently, their phonological store traces (which are subject to decay) can be refreshed less frequently than is possible for shorter words, and so fewer long than short words can be maintained in an active state in the phonological store.

Concurrent articulation ties up the articulatory mechanisms, reducing their availability, both for rehearsal and for phonological recoding of visually presented stimuli. As noted previously, concurrent articulation eliminates the *phonological similarity effect* in the case of visual, but not auditory presentation. The account of this is that auditory stimuli can continue to enter the store directly, and so there still is a phonological similarity effect for auditorily presented recall stimuli; however, the phonological recording of visual stimuli is blocked by concurrent articulation, so they no longer enter the phonological store, and so there is no phonological similarity effect for visually presented recall stimuli. Concurrent articulation also eliminates the *irrelevant speech effect* in the case of visual, but not auditory presentation. The account of this is very similar to that for phonological similarity: under concurrent articulation, auditory recall stimuli continue to enter the phonological store, where they encounter interference from the irrelevant stimuli; visual recall stimuli no longer enter the phonological store and hence are no longer subject to interference from irrelevant speech. Finally, concurrent articulation eliminates the *word length effect,* in both visual and auditory presentation. The account here is that, because the word length effect is a result of articulatory rehearsal, anything that blocks rehearsal, such as concurrent articulation, will eliminate the word length effect.

Reexamining Concurrent Articulation

As will be clear from the preceding discussion, the effects of *concurrent articula-*

tion on ISR performance have played a central role in establishing the standard interpretation. However, although this interpretation is in terms of an articulatory refresh, the data do not in fact uniquely support such an interpretation.

Logically, the effects of concurrent articulation could have any or all of the following components: (1) depletion of articulatory resources that would otherwise have been (more) available for rehearsal, (2) depletion (because of having to perform a secondary task) of a *general* processing capacity that is involved in immediate serial recall, and (3) creation of auditory stimuli, which undergo phonological coding and thus interfere with the actual serial recall stimuli—an irrelevant speech effect.

The standard interpretation of ISR data attributes the effects of concurrent articulation to the first of the above factors: the depletion of articulatory resources. The second factor above has also been examined. In a control condition for general processing load, subjects were given the concurrent task of finger-tapping instead of concurrent articulation; this control task had little or no effect on STM performance (Baddeley, 1990a). Assuming that finger-tapping is in fact the appropriate control, this would seem to rule out processing load as the factor underlying concurrent articulation effects. That leaves the third factor above: the auditory interference that might be created by concurrent articulation.

As will be discussed in more detail below, an account of the concurrent articulation effect can in fact be provided in terms of a model that assumes only "mental" or "auditory" refresh of phonological representations, and that attributes the effects of concurrent articulation to auditory interference—essentially, an irrelevant speech effect, created by the subjects' own irrelevant concurrent articulations.

It might be objected that, within the working memory model, there are several converging lines of evidence for the existence of articulatory refresh. Consequently, it might be argued, even if the individual phenomenon of concurrent articulation were amenable to an "auditory interference" account, such an account would not accommodate the various other data that can be accounted for in terms of articulatory rehearsal. Two points are worth making in this regard. *First,* a large part of the data usually adduced in support of articulatory rehearsal derives from ISR with *visual* presentation. However, as we have already seen, the relevant interpretations of these data are crucially dependent on the assumption that visually presented verbal stimuli undergo "phonological recoding" via articulatory mechanisms. This assumption is far from secure. In particular, it has been shown that phonological encoding of visual stimuli can occur even under concurrent articulation (Besner, 1987), and this renders the working memory model account of ISR with visual presentation problematic, even if further assumptions are made, as in Baddeley (1986). *Second,* if only the data from ISR with *auditory* presentation are taken into account, then, as shown below, all the ISR phenomena discussed above can be accounted for within an "auditory rehearsal" model. That is, it seems possible to account for all the data described in the first section of this paper *without* assuming any articulatory involvement in rehearsal processes.

In what follows, we outline both the "articulatory" and "auditory" rehearsal accounts, discussing their explanations of the various effects obtained in ISR tasks with auditory presentation of stimuli. First, however, our conception of short-term storage needs to be clarified. Short-term storage is viewed as a process in which "representations" are entered into a "loop" of activations, maintenance of which requires that representations must be *reentered* into the loop, which requires that they be "refreshed." The details of such processing are not important for present purposes.[2] The points to note here are that (a) evoked representations can (but do not automatically) form input to the short-term memory, (b) the short-term memory can temporarily learn *sequences* of stimuli appearing as input, and (c) the memory of sequences decays rapidly unless the inputs to the system are re-presented, i.e. refreshed. It is interesting to note that such a system appears similar to Hebb's (1949) notion of short-term memory as reverberation in a "closed loop."

The Articulatory Account

According to what we term the *articulatory* account, rehearsal in immediate serial recall involves the refresh of phonological representations by an output phonological process. These representations form the input to an associated short-term memory (STM) system of the kind discussed above. We use the term *refresh* to refer to the reinstatement of a phonological system representation and *rehearsal* to refer to the overall process involving both refresh and maintenance in the STM.

The articulatory account is essentially the articulatory loop model, cast in slightly different terms. *Phonological similarity effects* occur in the STM loop, and arise from the greater confusability of similar items. The *word length effect* reflects the fact that longer words take longer in their output processing; consequently, they are entered into the STM less frequently, and therefore decay more than shorter words would. The *irrelevant speech effect* arises because it interferes with the entry of phonological system representations into the STM loop: although unattended stimuli do not themselves get entered into the STM loop, they reduce the opportunities for recall stimuli to be so entered. *Concurrent articulation* is viewed as reducing the availability of the output processing mechanisms/resources needed to refresh phonological representations.

Note that both irrelevant speech and concurrent articulation reduce the opportunity for refreshed representations to enter the STM loop, but for different reasons, representing different "bottlenecks." With irrelevant speech, the bottleneck is at the point of entry of refreshed representations into the STM loop, whereas under concurrent articulation, the bottleneck is in the refresh process itself. Note also that concurrent articulation could have an irrelevant speech effect as well.

The interactions of concurrent articulation with the other effects are all as previously described for the articulatory loop model. To recapitulate briefly, under concurrent articulation, there should be no change in the phonological similarity effect, as representations are still subject to phonological confusability. Reduction of the word length effect occurs because articulatory processes are tied up in the concurrent articulations; memory items can therefore be refreshed only partially, if at all, thus narrowing or eliminating the difference in refresh rate for long vs short words. Finally, irrelevant speech results in a further deterioration in performance, given that it tends to impose a second bottleneck.

The Auditory Account

According to an *auditory* account, refresh during ISR involves a "mental echo." In this account, the refresh process is auditory, and does not involve articulatory or output processing mechanisms in any way. The "mental echo loop" thus consists of the "replaying" of auditory images of stimuli, resulting in the refreshing of their receptive or input phonological representations.

Phonological similarity effects occur in a local STM loop, just as in the articulatory account. The process of "replaying" a

[2] Recurrent autoassociative networks constitute a class of mechanism with the appropriate properties, and have been discussed by McNaughton & Morris (1987); functionally equivalent mechanisms have been used by Burgess & Hitch (1992).

stimulus takes longer for longer words, and the *word length effect* reflects this fact: the phonological representations of longer words are refreshed via echoing less frequently than those of shorter words, hence the STM associative weights decay more. The account of the *irrelevant speech effect* is identical to the one in the articulatory model: the irrelevant speech effect arises because it imposes a bottleneck on the entry of phonological representations into the STM loop. *Concurrent articulation* in this account is viewed as causing an irrelevant speech effect by reducing the opportunities for recall stimuli to be entered into the STM loop. Under concurrent articulation, there should be no change in the phonological similarity effect. The word length effect is reduced or abolished because the irrelevant speech effect created by concurrent articulation imposes a bottleneck on when refreshed representations can be entered into the STM loop, and this tends to affect long and short words equally. Presentation of actual external irrelevant speech results in further deterioration in performance, because it adds a further bottleneck.

It is worth noting the strength of this auditory account: it explains the same set of data that have been taken to support the articulatory account, and is equally convincing. It is also worth noting that the hypothesized auditory refresh process would be similar to auditory imagery, and that the plausibility of such imagery processes in serial recall is supported by the reported use of *visual* imagery with visually presented recall stimuli (Baddeley et al., 1975).

Articulatory vs Auditory Rehearsal: Implications

Before considering the implications of the articulatory-auditory issue, we will digress briefly, in order to clarify what we do and do not mean when we ask whether rehearsal is articulatory or auditory. As already stated, we conceive of ISR *rehearsal* as a process subserving the maintenance of decaying mental representations of the to-be-recalled items. A key aspect of rehearsal is that it involves the "refresh" of these representations before they have decayed beyond retrieval. The representations could be in articulatory form, or in auditory form, or in some amodal, truly phonological form that is used both in speech perception and in speech production. However, that is not the issue that concerns us here. Rather, the question we raise pertains to the nature of the *refresh* mechanism. It is important here to distinguish between representation and process: even if the representations being refreshed are amodal, the refresh process could still be articulatory. In asking whether it is articulatory, we ask whether this process does or does not involve mechanisms necessary for the motor programming of speech. If it does, it is "articulatory" in that it does draw upon processing mechanisms that by any criterion must be considered to be involved in articulation. If it does not involve any such mechanisms, then it is a more abstract process, one which we have labeled "auditory." It is in this sense that we distinguish between articulatory and auditory. Even in this sense, however, the question of whether rehearsal in immediate serial recall tasks is auditory or articulatory has significance well beyond the bounds of the immediate task domain or even working memory. In particular, it has implications for a number of phenomena related to phonological processing.

First, an important question regarding phonological perception is the extent to which information about the products of "output" phonological processing might be available to "input" phonological processing. There is a long-standing debate over whether speech perception proceeds via articulatory decoding (the "motor theory" of speech perception (Liberman & Mattingly, 1985)). Although the arguments cited in favor of this explicit decoding are not conclusive, they do suggest that articulatory information could form an input to the phonological representation system, especially to its development (for example, there is evidence that sighted children, but not blind children, are significantly more accurate in *producing* visibly articulated than nonvisibly articulated initial segments (Mulford, 1988)). A variety of other kinds of evidence have also been adduced in support of the existence of an *output* → *input* flow. Among these are data on speech monitoring, on repetition priming, on sublexical influences of production on perception, and the dissociation between receptive and productive vocabulary in language learning (Monsell, 1987). If there is an output process mediated refresh of phonological representations in ISR, this supports the possibility of articulatory information being available to receptive phonological processing.

Second, it seems plausible that the phenomenon of "inner speech" or auditory imagery could be related to feedback connections from pre-motoric speech activity to auditory perceptual systems; in fact, thinking on auditory imagery appears to be converging with thinking about the articulatory loop (Smith et al., 1992). The question of whether rehearsal in ISR is articulatory or auditory is thus closely related to the question of whether auditory imagery is purely abstract or whether it relies on articulatory mechanisms. The nature of visual imagery has been an important issue in cognitive psychology, and the articulatory/auditory rehearsal question taps into similar issues with respect to auditory imagery.

Third, recent work by Baddeley and colleagues (Baddeley et al., 1988; Gathercole & Baddeley, 1989; Gathercole & Baddeley, 1990) suggests that the mechanisms involved in immediate serial recall performance may also be involved in vocabulary acquisition, i.e., in the learning of novel phonological forms. The question of whether rehearsal processes in ISR are articulatory or auditory may therefore be important in understanding the phonological processing involved in vocabulary acquisition.

Finally, and at a more general level, there appears in recent years to have been growing interest in examining the relation between perception and action, and a move toward their more integrated treatment (Kolers & Roediger, 1984; Mackay et al., 1987). If rehearsal in immediate serial recall does in fact involve articulatory mechanisms, then, it would be of considerable interest as an example of the use of motor-related information in a purely cognitive task.

Discriminating between the Accounts

The obvious question is how to discriminate between the two accounts. It is useful to begin by examining the differing analyses of concurrent articulation more closely. As noted earlier, the effects of concurrent articulation can be analyzed into three components: (1) usage of articulatory resources, (2) creation of auditory interference, and (3) depletion of general resources resulting from having to perform a secondary task.

In explaining the impact of concurrent articulation on ISR, both the articulatory and auditory accounts are consistent with effects due to general resource competition, and auditory interference. Where they differ is in whether they allow for a role for articulatory resource competition. The articulatory account postulates articulatory resource competition, as in this view, ISR draws on articulatory resources. The auditory account disallows articulatory resource competition, as in this view, no articulatory processes are involved in ISR.

This suggests a means of discriminating between the two accounts. According to the auditory account, if we control for components (2) and (3) above of concurrent articulation with a *non-articulatory* task, then ISR performance under concurrent articulation will be equivalent to ISR performance under this non-articulatory control task. The articulatory account, on the other hand, predicts that ISR performance under concurrent articulation will be poorer than in this control task.

The analysis of the processing requirements of concurrent articulation and the control task are shown in Fig. 1a, and Fig. 1b shows the predicted relation between ISR span size with the concurrent control task (Condition I) and ISR span size with concurrent articulation (Condition II), under each of the two accounts. The auditory account predicts no difference between conditions, whereas the articulatory account predicts worse performance in Condition II (concurrent articulation).

Experiment 1

Given the experimental logic outlined above, the question is determining the right controls. According to Baddeley (1990b), *finger-tapping* is an appropriate control for the general resource demands of concurrent articulation, and has often been used as such (e.g., Baddeley et al., 1981, 1991;

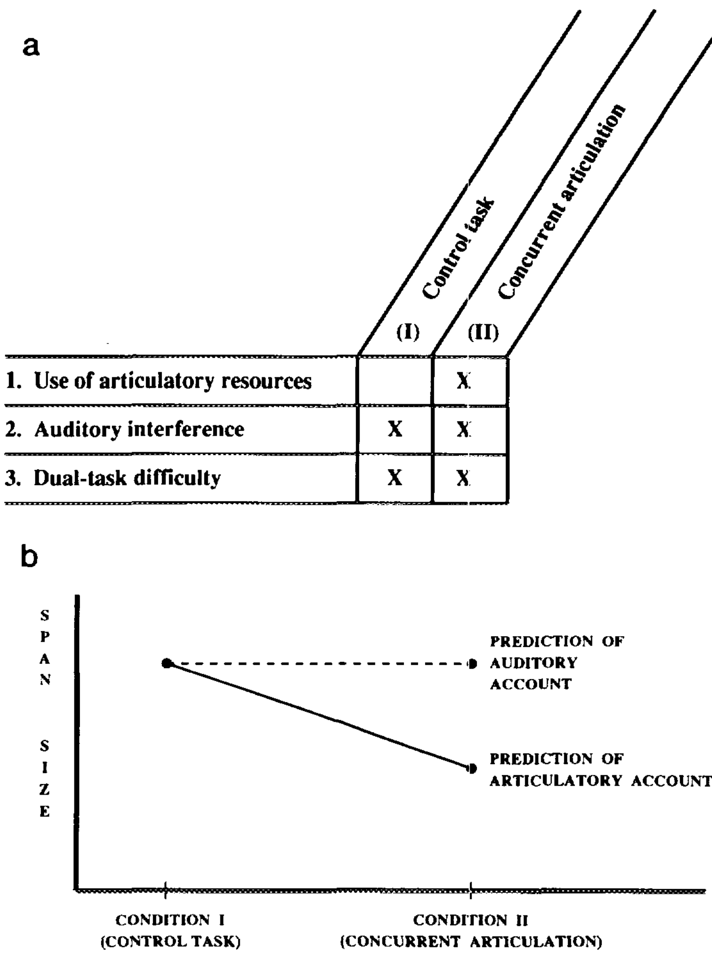

FIG. 1. Experimental logic. (a) Analysis of requirements of a concurrent articulation task (Condition II) and of a task that controls for auditory interference and dual-task difficulty (Condition I). (b) Predicted span sizes. The *articulatory* account predicts that span size will be lower in ISR under concurrent articulation (Condition II), than in ISR with the control task (Condition I). The *auditory* account predicts that there will be no difference.

Papagno et al., 1991).[3] An appropriate control for the auditory interference created by concurrent articulation should be exposure to repeated utterances of the same speech sound that is articulated during concurrent articulation. Combining these two controls with ISR creates a task that should discriminate between the two hypotheses. In this task, subjects would have to perform ISR while (i) concurrently tapping a finger and (ii) hearing repeated utterances of a speech sound. If general resource demands and auditory interference have been controlled for, then this task differs from performance of ISR under concurrent articulation only in its non-usage of articulatory resources.[4] To test these predictions, we devised an experiment based on the logic just outlined.

Method

Subjects. Thirty-six undergraduate students at Carnegie Mellon University participated in this experiment for course credit.

Design. There were two conditions in the experiment. In both conditions, the subjects' primary task was to recall sequences of digits that were presented to them auditorily. In each condition, subjects were also required to perform a secondary task during presentation of the recall stimuli. Each subject performed the primary task under each of the secondary task conditions, leading to a within-subjects repeated measures design. Each subject was randomly assigned to one of the two possible treatment orders.

Materials. One token of each of the digits *1* through *9* spoken by a female native speaker of American English was recorded as digitized sound on a Macintosh computer. Random sequences of these tokens were generated, varying in length from 4 digits to 11 digits.

Procedure. Each digit sequence was presented auditorily under computer control, at the rate of one digit per second. One trial consisted of presentation of one sequence of a particular length. For example, one trial at list length four consisted of auditory presentation of a sequence of four digits such as *3, 8, 2, 5.*

There were eight trials at each list length. Presentation of the lists began with sequences of four digits. If a subject recalled in correct serial order five or more of the eight sequences (trials) at a particular list length, the next higher list length was introduced. At the beginning of trials of each list length, the subject was told what the length of sequences would be. The longest list length for which a subject correctly recalled five or more sequences was taken as the measure of that subject's digit span in that condition.

Subjects were seated facing the computer screen, on which a cross appeared after presentation of each sequence. Subjects were instructed to repeat the sequence orally, in order, as soon as the cross appeared. The experimenter wrote down the subjects' responses, without providing any feedback about their accuracy. Initiation of each trial was controlled by the experimenter: before each trial, the experimenter asked the subject "Ready?" and when the subject nodded or said "Yes," the experimenter started the trial.

[3] Although we are not aware of any research actually validating this assumed equality of general resource demands, we nevertheless felt that finger-tapping was a reasonable control, at least as a first approximation.

[4] According to the motor theory of speech perception (Liberman & Mattingly, 1985), interpreting speech requires the use of articulatory mechanisms. On this theory, listening to repeated speech utterances in the control task *would* involve some usage of articulatory resources, as pointed out by one of the reviewers of this article. However, it seems reasonable to assume that any such usage of articulatory mechanisms would be less than that required for actual concurrent articulation, and with this assumption, the "articulatory" and "auditory" accounts still make differing predictions. The articulatory account still predicts superior performance under the control interference task, because of its lesser usage of articulatory resources. In the auditory account, usage of articulatory resources is irrelevant to ISR, and so the account still predicts equivalent performance in the two interference conditions.

In addition to the primary task, subjects were required to perform each of two secondary tasks. In the secondary task in the *Tapping + Noise* condition, subjects were required to concurrently tap the index finger of their dominant hand, throughout presentation of each sequence of digits, until the point when spoken recall began. At the same time, a recording was played of the word *the* being repeated every 500 ms, which subjects were instructed to ignore. However, they were instructed to tap their finger at the same rate as the repetitions of *the* (although they were not required to maintain synchrony with the tape). The recording was turned on by the experimenter at the start of each trial, and was turned off as soon as the cross appeared on the computer screen, indicating that the subject should respond.

In the secondary task in the *Concurrent Articulation* condition, subjects were required to concurrently and repeatedly articulate the word *the*, throughout presentation of each sequence of digits, until the point when spoken recall began. At the beginning of the condition, subjects were played a recording of the word *the* being repeated every 500 ms, and practiced articulating *the* in synchrony with the recording. They were asked to maintain that same rate of articulation during the actual experimental procedure (during which the tape was turned off by the experimenter).

Results

As noted above, subjects' digit span in each condition was taken to be the greatest list length at which they performed correctly on five or more of the eight trials. Mean digit span in the Tapping + Noise condition was 6.97. Mean digit span in the Concurrent Articulation condition was 5.89. The difference was significant ($F(1,35) = 48.09$, $MS_e = .44$, $p < .0001$).

Discussion

This result is consistent with the predictions of the articulatory account, but not the auditory account. That is, the results support the notion that depletion of articulatory resources plays a role in the effect of concurrent articulation on immediate serial recall performance. This, in turn, supports the idea that rehearsal during normal ISR does involve articulatory phonological processing resources.

Of course, this interpretation of our results is crucially dependent on two assumptions. First, we have assumed that the finger tapping task in the Tapping + Noise condition of our experiment is in fact an appropriate control for the general processing load imposed by Concurrent Articulation in the other interference condition. Second, we have assumed that exposure to repeated utterances of *the* under Tapping + Noise is a good control for any auditory interference effects of Concurrent Articulation of *the* in the other condition. In fact, however, the following arguments could be made against the appropriateness of these controls:

1. Any *non-articulatory* resources drawn on by ISR are not as depleted by finger-tapping as by concurrent articulation.
2. It is easier to ignore a tape playing a speech sound such as *the* than to ignore one's own concurrent articulation of the sound.
3. Concurrent articulation of *the* generates internal auditory noise (due to bone conduction and vibration) which is not present when listening to a tape playing *the*.

These possibilities led us to try and tighten up these controls in a second experiment, which we describe next.

EXPERIMENT 2

As noted above, Experiment 1 supported the view that the phonological loop has an articulatory component, rather than being purely auditory. However, such interpretation of the results depended crucially on the appropriateness of our controls for (1) the usage of "general" (non-articulatory) resources in concurrent articulation and (2) the auditory interference engendered by concurrent articulation of *the*. We have already discussed possible confounds.

In trying to tighten up our controls, we began by refining our analysis of the component processes in concurrent articulation. As compared with the earlier analysis, we now made a distinction between two kinds of auditory interference: *external* auditory interference and *internal* bone conduction/vibration, so that the analysis of concurrent articulation now involved the following four factors: (1) Use of articulatory resources, (2) external auditory interference, (3) internal bone conduction, and (4) dual-task difficulty. Clearly, the appropriate control task would have to control for all the latter three of these factors.

The control task we devised was one in which subjects were required to click their teeth throughout ISR stimulus presentation. Subjects were asked to perform this tooth clicking in synchrony with a tape playing the same speech stimulus that they themselves produced in the concurrent articulation condition. The rationale was that tooth clicking should control for the nonarticulatory resource demands of concurrent articulation more closely than does finger-tapping, being in the same motor modality as concurrent articulation. Additionally, requiring subjects to maintain synchrony with the external speech stimulus would prevent them from ignoring it, thus addressing the concern that it may be easier to ignore an externally presented speech sound than it is to ignore one's own concurrent articulation of that sound. Finally, the tooth clicking would generate internal bone conduction.

We further reasoned that if auditory interference could be eliminated from the concurrent articulation task (CA), then any remaining effect on recall performance would be attributable only to usage of articulatory resources and/or general dual-task difficulty. Requiring subjects to perform concurrent articulation *silently* would be such a task. Subjects would be required to concurrently articulate a word, engage in all the articulatory movements necessary to produce it, but without generating any sound, and without whispering. Such *silent concurrent articulation* (SCA) would eliminate auditory interference (both external, and internal bone conduction noise) from the concurrent articulation task. If recall performance under SCA did not differ significantly from performance under CA, then auditory interference could be ruled out as a factor. A significant difference would indicate, on the other hand, that auditory interference does play a role in the effect of concurrent articulation on ISR.

Finally, we reasoned that the concurrent articulation stimulus *the* used in Experiment 1 is an articulatorily simple as well as highly overlearned word, and its repeated production could perhaps be automated relatively easily, thus minimizing the use of articulatory resources and articulatory planning. We therefore decided to use the work *blank* instead, which is still semantically neutral, but is less familiar, longer, and more complex than *the*, so that concurrent articulation should be less easily automatized.

This led us to the three experimental conditions whose logic is outlined in Fig. 2. The analysis of component processes is as shown. Note that the aim of the control task was, as in Experiment 1, to isolate the effects of articulatory resource usage. Accordingly, the articulatory and auditory accounts make the same differing predictions as before for ISR performance under Concurrent Articulation vs. ISR performance under the control interference task (Clicking + Noise). Specifically, the articulatory account predicts a significant difference between recall performance in the Concurrent Articulation and Clicking + Noise conditions, whereas the auditory account predicts no significant difference between these two conditions. In addition, the Silent Concurrent Articulation condition would enable us to assess directly whether or not

	CA ("blank") (I)	Tooth click + hearing "blank" (II)	SCA ("blank") (III)
1. Use of articulatory resources	X		X
2. External auditory interference	X	X	
3. Internal bone conduction	X	X	
4. Dual-task difficulty	X	X	X

FIG. 2. Logic of Experiment 2. Analysis of requirements of a concurrent articulation task (CA), a control task (Clicking + Noise), and a silent concurrent articulation task (SCA).

there is an effect of auditory interference in concurrent articulation.

Method

Materials and design were as in Experiment 1, except that there were now four experimental conditions instead of two. These were a baseline condition, in which subjects performed immediate serial recall without any secondary task; and three conditions in which subjects also performed a secondary task concurrently with presentation of recall stimuli. As in Experiment 1, this was a within-subjects repeated measures design. Each subject performed the ISR task at baseline and under all three secondary task conditions, with presentation order of the secondary tasks being counterbalanced among subjects.

Subjects. Twenty-four undergraduate students at Carnegie Mellon University participated in the experiment for course credit and/or payment. None of these subjects had participated in the first study.

Procedure. For each subject, we first assessed digit span in an ISR task without any accompanying secondary task. In this baseline condition as well as in other conditions to be described, presentation of digits for recall was exactly as in Experiment 1. Baseline digit span was treated as being the highest list length at which the subject correctly recalled at least five of the eight trials.

Subsequently, each subject was asked to perform serial recall on lists of numbers whose length was his/her baseline span size, while concurrently performing one of the secondary tasks shown in Fig. 2 (in all cases, the secondary task was performed throughout presentation of the recall stimuli, but not during spoken recall). The dependent measure in each condition was therefore not span size as in Experiment 1, but the *number of trials correctly recalled* in that condition. The advantage of this procedure is that span does not have to be reestablished for each subject in each secondary task condition, which would involve testing at several list lengths. Rather, each subject performs each secondary task condition at a single list length corresponding to their baseline span size, thus greatly reducing testing time. Even though different subjects are tested at different list lengths, the recall task can be regarded as being equated in difficulty across subjects and conditions, because each is tested at his/her span, in each condition (Waters et al., 1992). In each experimental condition, the subject's primary task was to perform ISR for eight presentations (trials) of span-length digit sequences.

In the secondary task in the *Concurrent Articulation* condition, subjects additionally were required to concurrently articulate the word *blank* throughout presentation of each sequence of digits, until the point when spoken recall began. At the beginning of the condition, subjects were played a recording of the word *blank* being repeated every 500 ms and practiced articulating *blank* in synchrony with the recording (they were also given one trial of practice in doing this concurrently with ISR). They were asked to maintain the same rate of articulation during the actual experimental procedure. At the beginning of each trial, the tape was played so that subjects could begin concurrent articulation at the correct rate. The tape was then turned off and presentation of the recall stimuli was initiated.

In the secondary task in the *Clicking + Noise* condition, subjects were required to click their teeth concurrently with performance of ISR, and in synchrony with a tape playing utterances of the word *blank*, repeated every 500 ms. Subjects were asked to click their teeth by repeatedly opening and shutting their jaws so that their upper and lower back teeth came in contact, making a clicking noise. At the beginning of the condition, the experimenter demonstrated what was required, and subjects were given practice in performing this tooth clicking in synchrony with the tape, and also one practice trial of doing this concurrently with ISR. On each actual trial, subjects began the secondary task prior to presentation of the recall stimuli: the tape started playing before stimulus presentation.

In the secondary task in the *Silent Concurrent Articulation* condition, subjects were required to *silently* articulate the word *blank* concurrently with ISR. This condition was exactly like the Concurrent Articulation condition, except that the concurrent articulation was silent. Subjects were instructed to do everything that they would in actually saying the word aloud, and merely refrain from producing the sounds (and from whispering). They were explicitly instructed to make all the articulations involved in saying the word, and to think of themselves as actually saying the word. They were given practice in performing such repeated silent articulation, and also one practice trial of doing this concurrently with ISR.

As subjects were not generating any sound, an obvious potential problem is being sure that they had understood and were actually following instructions. To avoid this problem, they were asked at the end of practice to confirm that they understood what was required, and that they had been following the instructions (and not merely moving their lips, for instance). They were reminded of these instructions twice more during the task, and asked to confirm that they were following them. During practice, and at random intervals during the actual trials, the experimenter also confirmed by visual inspection that instructions were being followed. For these reasons, we are confident that subjects did in fact perform silent concurrent articulation as required. Note that the experimenter did not monitor subjects continuously during actual trials, to avoid the possibility of discomfiting them and thereby affecting their performance. For the same reason, the random monitoring was performed inconspicuously, and without staring directly at the subject.

Results

Results are shown in Fig. 3. There was a significant effect of type of secondary task on ISR performance ($F(3,23) = 23.97$, $MS_e = 1.85$, $p < .0001$). All six pairwise comparisons of treatment means were made using the Tukey procedure with a family confidence coefficient of .95. Performance in

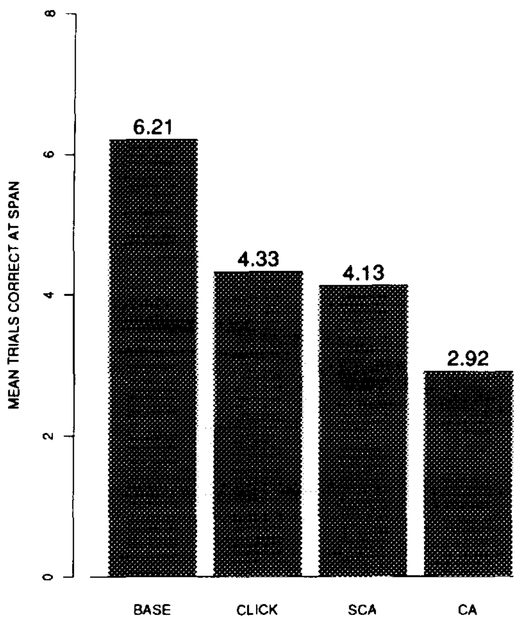

FIG. 3. Results of Experiment 2. Mean number of trials correct in ISR under various conditions. "BASE" is baseline ISR performance (with no secondary task). "CLICK" is ISR under the control interference condition (Clicking + Noise). "SCA" is ISR under silent concurrent articulation. "CA" is ISR under overt concurrent articulation.

all secondary task conditions was significantly lower than baseline. Performance under Concurrent Articulation was significantly lower than under both of the other secondary task conditions (Clicking + Noise, and Silent Concurrent Articulation). However, the latter two conditions did not differ significantly from each other.

Discussion

In interpreting these results, it may be helpful to refer back to the analysis of the three experimental conditions given in Fig. 2.

First, the significant difference between the overt concurrent articulation condition and the tooth clicking control condition indicates that the usage of articulatory resources in concurrent articulation plays a role in its impact on ISR performance. This is consistent with the articulatory rehearsal hypothesis, but not with the auditory hypothesis. Second, the significant difference between the overt and silent concurrent articulation conditions indicates that auditory interference also does play a role in the effects of (overt) concurrent articulation on ISR performance. That is, a significant but hitherto unrecognized component of the concurrent articulation effect is in fact an irrelevant speech effect. Thus, the concurrent articulation effect appears to have two components: the usage of articulatory resources, and an irrelevant speech effect. Third, the lack of any significant difference between the silent concurrent articulation and tooth clicking control conditions suggests that the relative impacts of these two components may be approximately equal (see Fig. 2).

The results of Experiment 2 thus indicate a significant auditory interference effect in concurrent articulation. However, as can be seen from Fig. 2, our manipulations did not assess the relative impacts of "internal bone conduction" and "external" auditory interference. Additionally, although the concurrent articulation stimulus used in Experiment 2 was more complex than that in Experiment 1, it could still be argued that performing repeated concurrent articulation of a word such as *blank* might quickly become automated, especially in view of its monosyllabicity. The attempt to address these considerations led to our next experiment.

EXPERIMENT 3

As noted above, one aim of our third experiment was to separate the effects of "internal bone conduction" and "external" auditory interference. The secondary task we devised was lip-syncing to a speech sound that was being played externally. The logic is illustrated in Fig. 4. Lip-syncing to a word (which is heard externally) differs from overt articulation of that word in its lack of internal bone conduction. Silent articulation lacks both the internal and external auditory interference components of overt articulation. Experiment 2

	(I) CA ("kwelstry")	(II) Lip-sync to "kwelstry"	(III) SCA ("kwelstry")
1. Use of articulatory resources	X	X	X
2. External auditory interference	X	X	
3. Internal bone conduction	X		
4. Dual-task difficulty	X	X	X

FIG. 4. Logic of Experiment 3. Analysis of requirements of concurrent articulation (CA), lip-syncing, and silent concurrent articulation (SCA).

indicated that silent concurrent articulation is significantly less difficult than overt concurrent articulation, and we assumed that this effect would hold up. The question in Experiment 3 was whether lip-syncing would differ significantly from either silent or overt concurrent articulation. A significant difference from overt but not from silent concurrent articulation would indicate that internal bone conduction but not external auditory interference plays a significant role in the concurrent articulation effect. A significant difference from silent but not from overt articulation would indicate that external auditory interference, but not internal bone conduction, plays a significant role in the concurrent articulation effect. A significant difference from both silent and overt concurrent articulation would indicate that *both* external auditory interference and internal bone conduction play a significant role in the concurrent articulation effect.

A second aim of the experiment was to further reduce the possibility of automatized performance of concurrent articulation. We reasoned that an articulatorily complex, bisyllabic, nonword stimulus would not permit use of an existing overlearned, automated motor program, and would maximize the articulatory planning needed for the various kinds of articulation task. To this end, we used the nonsense word *kwelstry* as the concurrent articulation response.

Method

Materials and design were exactly as in Experiment 2, with a baseline condition and three secondary task conditions. As in Experiment 2, the experimental design was a within-subjects repeated measures design. Each subject performed the ISR task under all four conditions, with presentation order of the three secondary task conditions being counterbalanced among subjects.

Subjects. Thirty-six students and research staff members at Carnegie Mellon University participated in the experiment for payment and/or course credit. None of the subjects had participated in either of the other studies.

Procedure. Baseline digit span was assessed for each subject exactly as in Experiment 2. Subsequently, each subject was asked to perform serial recall on sequences of digits whose length was *one less than* his/her baseline span size, while concurrently performing one of the secondary tasks shown in Fig. 4. This was because we had increased the difficulty of the concurrent articulation stimulus (*kwelstry*, as compared with *the* and *blank* in Experiments 1 and 2), and we wanted to ensure that we did not reduce performance to floor. As in Experiment 2, the dependent measure was the number of trials correctly recalled under each condition.

In the secondary task in the *Concurrent Articulation* condition, subjects were required to overtly articulate the nonword *kwelstry* concurrently with performance of ISR. The procedure was exactly as in Experiment 2, except for the list length (one less than span size) and the articulation stimulus. The *Silent Concurrent Articulation* condition also differed from Experiment 2 only in these two respects.

In the *Lip-Syncing* condition, subjects were required to articulate *kwelstry* silently and concurrently with ISR, and in synchrony with a tape playing repeated utterances of *kwelstry*. Subjects were given the same instructions as for silent concurrent articulation, and in addition were instructed to maintain synchrony with the tape.

Results

Results are shown in Fig. 5. We made planned pairwise comparisons of ISR performance, as follows. (i) Performance was significantly higher under the Silent Concurrent Articulation condition than under Lip-Syncing ($F(1,35) = 6.79$, $MS_e = 1.60$, $p < .05$). (ii) Performance was significantly higher under the Lip-Syncing condition than under overt Concurrent Articulation ($F(1,35) = 9.74$, $MS_e = 1.55$, $p < .005$). (iii) Performance was significantly better under the Silent Concurrent Articulation condition than under overt Concurrent Articulation ($F(1,35) = 31.83$, $MS_e = 1.62$, $p < .0001$).

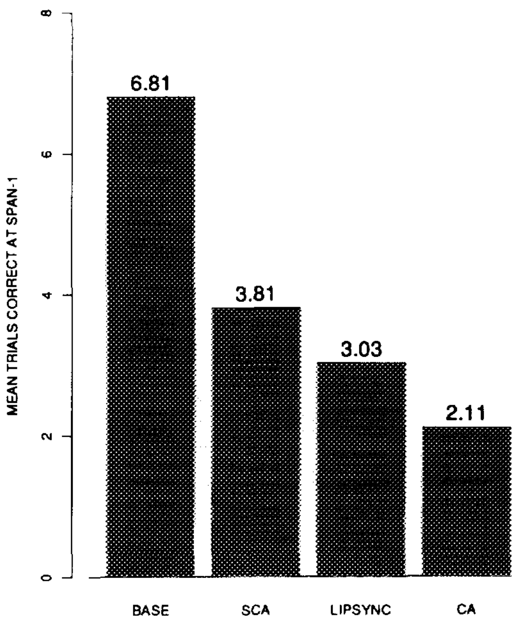

FIG. 5. Results of Experiment 3. Mean number of trials correct in ISR under various conditions. "BASE" is baseline ISR performance (with no secondary task). "SCA" is ISR under silent concurrent articulation. "LIPSYNC" is ISR under lip-syncing to auditory interference. "CA" is ISR under overt concurrent articulation.

Discussion

In interpreting these results, it may be helpful to recall the analysis of the three experimental conditions given in Fig. 4. First of all, the significant difference between the overt and silent concurrent articulation conditions replicates the findings of Experiment 2, thus confirming the role of auditory interference in the concurrent articulation effect.

Second, there is a significant difference between the lip-syncing and overt concurrent articulation conditions. This indicates that the effects of auditory interference in concurrent articulation have a significant component due to *internal bone conduction*. Third, there is also a significant difference between the lip-syncing and silent concurrent articulation conditions, which indicates that *external auditory interference* has a significant impact on ISR performance under (silent) concurrent articulation. Taken together, these results show that *both* external auditory interference and internal bone conduction play a significant role in the Concurrent Articulation effect.

The external auditory interference in the Lip-Syncing condition can be regarded as irrelevant speech. Our results therefore indicate that there is an irrelevant speech effect even under Silent Concurrent Articulation. This is consistent with previous findings regarding the effects of irrelevant speech on recall under (overt) concurrent articulation. For example, Hanley and Broadbent (1987) obtained results indicating that the detrimental effect of irrelevant speech on ISR is preserved under overt concurrent articulation. That is, recall performance is worse when subjects perform concurrent articulation *and* are exposed to irrelevant speech than when they only perform concurrent articulation. Baddeley et al. (1991) employed irrelevant speech stimuli that were "sandwiched" between the recall stimuli, i.e., an irrelevant speech stimulus was played in between every two recall stimuli, while subjects also performed concurrent articulation. It was found that recall performance was worse under these conditions than when subjects only performed concurrent articulation.

Clearly, the present results are consistent with these findings; however, they also serve to extend the earlier results. This is because the present study differs from these earlier studies in one important respect. In Experiment 3 discussed above, subjects' concurrent articulations (whether overt or silent) in all three conditions were *synchronized* with the external irrelevant speech (whether self-generated or on tape). There was no such synchronization in either of the two previous studies. In the work by Hanley and Broadbent (1987), subjects' concurrent articulations were at the rate of three times per second, whereas the irrelevant speech was synchronized with the presentation of recall stimuli, at the rate of once every second. Similarly, because the Baddeley et al. (1991) study employed sandwiched irrelevant speech stimuli, there is no reason to expect that this irrelevant speech maintained synchrony with subjects' concurrent articulation.

The present findings therefore show that external auditory interference will further impact ISR performance under (silent) concurrent articulation, *even if synchronized with the concurrent articulations*. This has implications for the "articulatory" and "auditory" accounts in the second section. As discussed there, bottlenecks will be imposed on the entry of recall stimuli into a short-term memory system by two kinds of interference effects: (i) the interruption of rehearsal and (ii) the presence of interfering auditory speech stimuli. In terms of this conceptualization, the results of the Hanley and Broadbent (1987) and Baddeley et al. (1991) studies indicate that this bottleneck effect is additive if both types of interference are present *asynchronously*. Results from the present Experiment 3 indicate that the bottleneck effect is also additive if both types of interference are *synchronously* present.

We also derived serial position curves for the various conditions in Experiment 3, for the 12 subjects with a baseline span of nine digits (this span size constituted the single largest group of subjects). It seemed an interesting possibility that our interference tasks might have systematic interactions with primacy, recency, and suffix effects.[5] For example, auditory interference might tend to reduce recency effects, and articulatory interference might particularly impact primacy; such a finding would create an interesting connection between research on suffix-modality effects and research on the working memory model, which has had relatively little to say about recency or suffix effects. We found, however, that the se-

[5] The *primacy* and *recency* effects refer to the advantage in recall for, respectively, the first few and last few items on a list. The *suffix* effect refers to the abolition of the recency effect if an irrelevant speech stimulus is presented after the last item of an auditory list.

rial position curves at baseline and under the various interference conditions show classical primacy and recency effects, and there does not appear to be any systematic impact of type of interference condition. This is consistent with serial position data from Hanley and Broadbent (1987), which show normal primacy and recency effects under (overt) concurrent articulation, under irrelevant speech, and under both concurrent articulation and irrelevant speech. We therefore concluded that serial position data do not add much to our experimental analyses.

General Discussion

Experimental Results

In our re-analysis of concurrent articulation, we pointed out that performing ISR under concurrent articulation imposes general difficulty as well as auditory interference. It could be argued that these two factors are sufficient to account for the concurrent articulation effect in immediate serial recall, without positing interference with an articulatory resource as a factor. We outlined how both the usual *articulatory* rehearsal account and an alternative *auditory* rehearsal account can accommodate the key phenomena of ISR.

In Experiment 1, we adopted the customary control for general difficulty of processing in concurrent articulation tasks (finger tapping). In addition, we attempted to control for the auditory interference that is generated in a concurrent articulation task, which does not appear to have been examined previously as a component of the concurrent articulation effect. In the control for concurrent articulation, subjects (i) were exposed to repetitions of the same speech sound that they themselves repeated under concurrent articulation and (ii) were also required to tap their finger at the same rate as the repeated speech sounds. Subjects' ISR performance was significantly better in this control condition than under concurrent articulation. At first sight therefore, the results of Experiment 1 were consistent with the predictions of the articulatory account but not the auditory account.

A number of questions can be raised, however, about the adequacy of finger-tapping as a control for the general difficulty of concurrent articulation, and also about the adequacy of exposure to speech as a control for the auditory interference generated by concurrent articulation. In Experiment 2, we attempted to tighten up these controls. In the new control task, subjects were exposed to speech sounds as in Experiment 1, but were now required to also click their teeth, not only at the same rate as the speech sound, but also in synchrony with them. This control task was intended to approximate the general difficulty of concurrent articulation much more closely than finger tapping. Additionally, in this control task subjects were generating internal (bone conduction) sound, as in concurrent articulation, and were prevented from possibly ignoring the external speech by being required to maintain synchrony with it. Overall, we felt this was a much more stringent control task than that in Experiment 1. Additionally, in Experiment 2 we also examined the effects of silent concurrent articulation on ISR, in order to be able to factor out the auditory interference component of concurrent articulation entirely. The task analysis of conditions in Experiments 2 and 3 is shown in Fig. 6 and may be helpful to consult as needed (the reader should ignore the shaded outer row and column, for the moment).

Subjects' performance in the control task in Experiment 2 was significantly better than under concurrent articulation. This finding is explicable under the articulatory account (depletion of articulatory resources available for rehearsal), but not under the auditory account (usage of articulatory resources should be irrelevant to rehearsal). To the extent that we had controlled for everything except the use of articulatory re-

	EXPERIMENT 2 (I) CA ("blank")	EXPERIMENT 2 (II) Tooth click + hearing "blank"	EXPERIMENT 2 (III) SCA ("blank")	EXPERIMENT 3 (I) CA ("kwelstry")	EXPERIMENT 3 (II) Lip-sync to "kwelstry"	EXPERIMENT 3 (III) SCA ("kwelstry")	Tooth + hearing & thinking "blank"
1. Use of articulatory resources	X		X	X	X	X	??
2. External auditory interference	X	X		X	X		X
3. Internal bone conduction	X	X		X			X
4. Dual-task difficulty	X	X	X	X	X	X	X
Speech inside the head	X		X	X	X	X	X

FIG. 6. Analysis of conditions in Experiments 2 and 3 combined (excluding the shaded row and column). See Discussion in text.

sources under concurrent articulation, this supported the articulatory rehearsal hypothesis. Performance under silent concurrent articulation was also significantly better than under overt concurrent articulation, indicating that auditory interference plays a significant but previously unrecognizable role in the concurrent articulation effect. ISR performance in both the control task and under silent concurrent articulation were significantly lower than baseline, but did not differ significantly from each other. This suggests that the auditory interference in overt concurrent articulation impacts ISR to approximately the same degree as the use of articulatory resources in overt or silent concurrent articulation (see Fig. 6).

In Experiment 3, we further examined the auditory interference effects of concurrent articulation. Silent concurrent articulation eliminates both the external auditory interference and internal bone conduction noise of overt concurrent articulation. To determine the relative contributions of these kinds of auditory interference to the concurrent articulation effects, we devised a new lip-syncing condition. Subjects in Experiment 3 thus performed ISR under three concurrent task conditions: overt articulation, lip-syncing, and silent articulation.

ISR performance under overt concurrent articulation was significantly lower than under silent concurrent articulation, replicating the finding in Experiment 2. The new results were that ISR performance in the lip-syncing condition was significantly better than under overt concurrent articulation, and significantly worse than performance under silent concurrent articulation.

As shown in Fig. 6, the task analysis of lip-syncing is that it involves the use of articulatory resources, external auditory interference, and general processing load (dual-task difficulty). Compared with overt concurrent articulation, lip-syncing lacks the internal bone conduction noise component. The significantly better ISR performance under lip-syncing than under overt concurrent articulation therefore indicates that the internal bone conduction noise engendered by overt concurrent articulation has a significant effect on ISR performance. Compared with silent concurrent articulation, lip-syncing additionally involves external auditory interference. The significantly worse ISR performance under lip-syncing than under silent concurrent articulation therefore indicates that the external auditory interference engendered by overt concurrent articulation also has a significant effect on ISR performance. This latter finding is consistent with previous investigation of the effects of irrelevant speech on ISR under concurrent articulation (Hanley & Broadbent, 1987; Baddeley et al., 1991), and extends those earlier results to the case where the irrelevant speech is *synchronous* with concurrent articulation.

The detrimental impact of concurrent articulation on ISR performance has usually been explained solely in terms of a reduction of articulatory resources available for rehearsal (see the earlier discussion of the articulatory and auditory accounts). The results of Experiments 2 and 3 indicate, however, that a significant part of this detrimental effect is in fact due to two auditory interference components of concurrent articulation: the internal bone conduction noise that is generated by concurrent articulation, as well as the external auditory interference that is also generated.

Overall, the results of the present experiments both support and qualify the "articulatory" hypotheses. In view of the results of Experiment 2, the argument that rehearsal under ISR has a component that is articulatory in nature is now considerably stronger, and the purely auditory rehearsal account less tenable. However, Experiments 2 and 3 also highlight the role of auditory interference in creating the concurrent articulation effect, showing that this effect has been overinterpreted as reflecting the use of articulatory mechanisms in ISR.

Rehearsal and Auditory Imagery

The experiments as so far described appear to support the articulatory hypothesis, although they also indicate that internal auditory interference (bone conduction noise) and external auditory interference both play a significant role in the concurrent articulation effect. However, there is a factor we have not taken into account in our task analyses, because there seemed no way of separating its effects from those of articulatory resource usage. During rehearsal, irrespective of whether such rehearsal is by articulatory or by auditory processes, there is "speech inside the head." According to the articulatory account, this speech inside the head is the concomitant of articulatory rehearsal. According to the auditory account, this "speech inside the head" *is* rehearsal, and there is no accompanying articulatory involvement. But speech inside the head is also, of course, generated by concurrent articulation. So far, we have distinguished "internal bone conduction" and "external" auditory interference effects in concurrent articulation. We now need to add "speech inside the head" to the interference effects caused by concurrent articulation.

With this further refinement of the componential analysis of concurrent articulation, our analysis of the various experimental conditions changes. These changes can be seen by now including the shaded row in Fig. 6 as part of the analysis. As can be seen, the tooth-clicking control of Experiment 2 in fact does not control for the "inner speech" effect of concurrent articulation. The difference between ISR performance under concurrent articulation and under the tooth clicking control condition can therefore be explained by a purely auditory account: it can be argued that the differential impact of concurrent articulation (compared with the control) is due, not to the differential usage of articulatory resources, but to the inner speech engendered by concurrent articulation (but not by the control), which interferes with the inner speech that is part of an auditory rehearsal process.

It might seem that the appropriate control would be to have subjects perform the tooth-clicking task, as in Experiment 2, but this time also concurrently "thinking" the articulation stimulus (e.g., *blank*). Such a control is shown in the last (shaded) column of Fig. 6. Would such a control finally distinguish between the auditory and articulatory hypotheses?

The answer is no, because we are back at precisely the question we began with. In its original form, the question was whether rehearsal is articulatory or auditory. In its new form, the question is, what is going on in "thinking" a verbal stimulus? According to the articulatory account, "thinking" *blank* involves articulatory planning of *blank;* under this account, it is not possible to set up a task involving repeatedly thinking a word but not planning its articulation. ISR performance under the two tasks would therefore be predicted not to differ significantly. The auditory account also makes the same prediction, but for a different reason. That is, according to the articulatory account, both concurrent articulation and the proposed speech-inside-the-head control task have an articulatory component, and they will therefore interfere equally with articulatory rehearsal. According to the auditory account, both concurrent articulation and the control task involve (non-articulatory) speech-inside-the-head, and so they will interfere equally with the purely auditory rehearsal process; the additional involvement of articulatory processes in concurrent articulation will be irrelevant to this non-articulatory rehearsal process. The fundamental question we have returned to is that of the nature of auditory imagery: "auditory" or "articulatory?" There has, in fact, been a convergence of thinking about auditory imagery and the "articulatory loop" (Smith et al., 1992), and the present discussion highlights this point.

Auditory imagery has been defined as "the introspective persistence of an auditory experience, including one constructed from components drawn from long-term memory, in the absence of direct sensory instigation of that experience" (Intons-Peterson, 1992, p. 46). Researchers have talked of an "inner ear" and an "inner voice," and these have been identified respectively with the articulatory loop's "phonological store" and "articulatory rehearsal" (Campbell, 1992; Smith et al., 1992). An important issue in this area has been the roles of the inner ear and the inner voice in various phenomena of auditory imagery, a question of obvious relevance to the present discussion. Several points are worth noting.

First, some of the manipulations designed to assess the roles of the inner ear or inner voice have been the same as in the working memory literature: concurrent articulation, and irrelevant speech.[6] For example, in one experiment, subjects were asked to judge whether each of a printed list of words (such as *larks, dogs, halves, cats*) would be pronounced with a final "s" or "z" sound. Performance was examined in a baseline condition, under concurrent articulation, with concurrent auditory input, and with both concurrent articulation and concurrent auditory input. Concurrent articulation appreciably reduced performance, whereas concurrent auditory input had no impact on performance. From this, it was concluded that the inner voice but not the inner ear is involved in this particular kind of auditory imagery (Smith et al., 1992, pp. 108–109).

Second, the indications (based on the kind of logic just exemplified) are that the inner voice is needed for a variety of imagery tasks. However, different aspects of the inner voice may be more or less relevant to different imagery tasks. For example, in a voicing judgment task similar to the one described above, concurrent *humming* was sufficient to disrupt performance (Smith et al., 1992, p. 106).

Third, there are types of auditory imagery (for example, imagining the timbre of a specific musical instrument) that seem to rely on the inner ear and not on the inner voice (Smith et al., 1992, pp. 104, 108). It

[6] Other working-memory related research in auditory imagery includes the investigation of rehearsal in different imagined voices (e.g., Geiselman & Bjork, 1980).

seems intuitively likely that the role of the inner voice would be more prominent for imagery of speech-like sounds.

Finally, it should be clear that the confound discussed above with respect to the experiments we have reported in this paper applies equally to auditory imagery. Concurrent articulation also generates "speech in the head" and so could have its impact on rehearsal (or imagery) via interference with rehearsal (or imagery) that is completely auditory (or inner-ear) based, and that does not in fact involve articulatory processes in rehearsal (or imagery).

The Need for Neurophysiological Evidence

The situation we have described above has some parallels with the debate that continued for some years over whether visual imagery is propositional or visual (Pylyshyn, 1973; Kosslyn et al., 1979; Pylyshyn, 1981). In both cases, the issue is whether or not a fundamental physiologically definable system (the visual perceptual system or the articulatory system) is involved in the process. With respect to visual imagery, it is worth noting that at least one prominent researcher concluded that this was an unresolvable question for cognitive psychology (Anderson, 1978). It is further worth noting that some of the most influential data bearing on the visual imagery debate came via the techniques of the *neurosciences* and not from the standard methodology of experimental psychology (e.g., Farah, 1988). In particular, evidence that cortical areas involved in visual processing are also involved in visual imagery has had an important bearing on this issue. This suggests that it would be appropriate to ask what the neural substrates of articulation are, and whether there is evidence implicating these substrates in rehearsal.

We believe the present question is an excellent example of a psychological issue that can benefit from neurophysiological data, because it is possible to specify, with some confidence, at least some of the neural substrates that should be involved in ISR, if rehearsal is indeed articulatory. In particular, Broca's area (Brodmann Area 44) lies immediately anterior to the premotor cortex and has a long history of being viewed as the secondary motor area specialized for speech motor control/programming (Barlow & Farley, 1989; Geschwind, 1979; Gracco & Abbs, 1987; Liepmann, 1915; Mayeux & Kandel, 1985; Mohr, 1976). There also seems to be agreement that Broca's area is *not* responsible for several functions that have at various times been attributed to it: the functions of grammar and syntactic processing cannot be localized to Broca's area (Kean, 1985; Bates et al., 1988), and the syndrome of Broca's aphasia does not arise from damage to Broca's area alone (Kertesz et al., 1979; Mateer, 1989; Mohr, 1976; Naeser & Hayward, 1978). It is therefore possible to identify Broca's area as being implicated in articulatory programming, planning and control. Involvement of Broca's area thus constitutes a clear diagnostic for articulatory planning activity. Therefore, if verbal working memory and, in particular, rehearsal processes, truly involve articulatory planning, they should rely in some way on processing in Broca's area.

Unfortunately, the neurophysiological literature pertinent to this question is limited, and far from conclusive. Two recent studies (Waters et al., 1992; Martin et al., under review) have examined ISR performance in patients identified as having speech planning deficits, i.e., as having apraxia of speech. These studies do not provide direct evidence about the effect of Broca's area lesions on ISR performance; rather, they provide at best indirect evidence, by examining ISR performance in patients who exhibit *the kinds of deficits that would be expected* following Broca's area lesions.[7] In the earlier study (Waters et

[7] In these studies, the lesions are either not reported (Waters et al., 1992: patients are merely described as having suffered left-hemisphere cerebro-vascular accident, with five of the six patients being classified as

al., 1992), a group of apraxic patients were found to exhibit a pattern of ISR performance similar to that observed in normals under concurrent articulation, suggesting that their rehearsal processes were impaired, and supporting the idea of articulatory involvement in rehearsal. In contrast, in another recent study (Martin et al., under review), a single patient diagnosed as having a selective deficit of articulatory planning was found to exhibit the pattern of word-length effects shown by normals under unimpeded ISR, suggesting that the patient's rehearsal processes were unimpaired despite deficits of articulatory planning. This would seem to challenge the notion that rehearsal relies on articulatory mechanisms.

However, neither of these studies can be viewed as providing strong evidence one way or another. In the Waters et al. (1992) study, on the one hand, apraxia of speech was not the patients' only deficit (the patients were nonfluent aphasics), and so, as Martin et al. (under review) have noted, it is difficult to ascribe the observed pattern of ISR performance solely to the articulatory programming deficit. In the Martin et al. (under review) study, on the other hand, it is not at all clear whether the patient was in fact performing rehearsal. He was only tested on lists of between four and six digits in length, and at these list lengths it is quite feasible to account for word-length effects without positing rehearsal. Moreover, the normal controls' overall performance was even worse than that of the patient, which is difficult to explain if they were indeed rehearsing. It may well be that neither the patient nor the controls were rehearsing. In view of this possibility, the results of this study cannot be viewed as definitive, either.

In bringing neuropsychological data to bear on the present issue, one obvious difficulty arises from the rarity of focal lesions to Broca's or other premotor areas (and/or the rarity of "pure" speech apraxias) in adults. A second problem arises from the difficulty in determining whether a patient is or is not rehearsing, given that premorbid digit span is unknown, and given that the list lengths that can feasibly be tested with such patients are within the range of normal performance without rehearsal. For these reasons, it may be difficult to obtain conclusive evidence regarding the present issue from neuropsychological studies.

Functional neuroimaging has the potential to avoid some of these problems through more direct examination of processing in intact neural structures in normal subject populations, and is therefore the most effective methodology for examining the involvement of Broca's area in rehearsal. It is therefore interesting that brain imaging studies bearing on the question of auditory vs. articulatory rehearsal have in fact been reported recently by Paulesu et al. (1993), who found activation of Broca's area in both an immediate serial recall task and a rhyme judgement task; they interpreted this as indicative of articulatory rehearsal (or the "inner voice").[8]

At first sight, the data of Paulesu et al. (1993) would appear to settle the question of whether rehearsal and the inner voice are articulatory or auditory. In fact, however, interpretation is confounded by the fact that stimuli were presented visually. Visual stimuli have been argued to undergo phonological encoding via articulatory processes, as discussed in the second section. As we pointed out in that section, this hypothesis about the phonological encoding of visual stimuli is not indisputable, and for this reason we have excluded ISR with visual presentation from our analyses. Nev-

Broca's aphasics), or else are non-focal (Martin et al., under review).

[8] Subjects perform more poorly on a rhyme judgement task if concurrent articulation is required of them (Besner, 1987), from which it has been concluded that the "inner voice" is involved in articulating the to-be-judged stimuli.

ertheless, this hypothesis does raise questions about what the observed Broca's area activation represents: it could in fact be evidence of articulatory recoding of visual stimuli in both tasks, rather than of articulatory rehearsal or the inner voice.

Although this particular study may not have resolved the auditory/articulatory issue (nor was is specifically designed to do so), it indicates a promising line of research that could eventually provide more definitive data. We do not wish to suggest neuropsychological/neuroimaging data as the panacea for all psychological inquiry. However, for specific issues such as the present one, cognitive psychology may benefit greatly from drawing on the methodologies of the neurosciences.

Summary and Conclusions

In this paper, we focused on the effects of concurrent articulation on immediate serial recall performance. These effects have been central to development of the "articulatory loop" model of auditory-verbal short term memory. We reexamined these effects, showing that they are consistent with an alternative "auditory" account that posits a nonarticulatory rehearsal process not in any way dependent on articulatory mechanisms. We showed that this auditory hypothesis can also account for other key phenomena observed in immediate serial recall.

The nature of the rehearsal process (whether articulatory or auditory) has implications for several phenomena of phonological processing. We described three experiments that were designed to discriminate between the auditory and articulatory hypotheses. We concluded (1) that our results support an articulatory component in rehearsal, but also (2) that a hitherto unrecognized irrelevant speech effect is a significant part of the concurrent articulation effect.

We then discussed still other factors that could be playing a role in concurrent articulation, and showed how the articulatory/auditory issue at this point converges with questions about the nature of auditory imagery. We suggested that such issues may best be resolved by augmenting the results of psychological experimentation with data from neuroimaging studies. We described a recent neuroimaging study that, although inconclusive, represents a first step in this direction.

We conclude that our analysis in this paper points up some inadequacies in the widely influential working memory model. The experiments we designed go some way toward resolving these interpretational inadequacies, and discriminating between the auditory and articulatory hypotheses. However, ultimate resolution of these questions may be greatly facilitated by data from brain imaging studies. Precise specification of the articulatory loop model pushes against the limits of conventional psychological inquiry.

References

Anderson, J. R. (1978). Arguments concerning representations for mental imagery. *Psychological Review*, 85, 249–277.

Baddeley, A. D. (1986). *Working memory*. New York: Oxford University Press.

Baddeley, A. D. (1990a). The development of the concept of working memory: Implications and contributions of neuropsychology. In G. Vallar & T. Shallice (Eds.), *Neuropsychological impairments of short-term memory*. Cambridge, England: Cambridge University Press.

Baddeley, A. D. (1990b). *Human memory*. Boston: Allyn and Bacon.

Baddeley, A. D., Eldridge, M., & Lewis, V. J. (1981). The role of subvocalization in reading. *Quarterly Journal of Experimental Psychology*, 33, 439–454.

Baddeley, A. D., Lewis, V., & Vallar, G. (1984). Exploring the articulatory loop. *Quarterly Journal of Experimental Psychology*, 36A, 233–252.

Baddeley, A. D., Papagno, C., & Norris, D. (1991). Phonological memory and serial order: A sandwich for TODAM. In W. E. Hockley & S. Lewandowsky (Eds.), *Relating theory and data: Essays on human memory in honor of Bennet B. Murdock*. Hillsdale, NJ: Erlbaum.

Baddeley, A. D., Papagno, C., & Vallar, G. (1988). When long-term learning depends on short-term storage. *Journal of Memory and Language*, 27, 586–595.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, **14**, 575–589.

Barlow, S. M., & Farley, G. R. (1989). Neurophysiology of speech. In D. P. Kuehn, M. L. Lemme, & J. M. Baumgartner (Eds.), *Neural Bases of Speech, Hearing and Language*. Boston, MA: Little, Brown.

Bates, E., Thal, D., & Janowsky, J. (1988). Early language development and its neural correlates. In I. Rapin & S. Segalowitz (Eds.), *Handbook of neuropsychology, Volume 6, child neurology*. Amsterdam: Elsevier.

Besner, D. (1987). Phonology, lexical access in reading, and articulatory suppression: A critical review. *Quarterly Journal of Experimental Psychology*, **39A**, 467–478.

Burgess, N., & Hitch, G. J. (1992). Toward a network model of the articulatory loop. *Journal of Memory and Language*, **31**, 429–460.

Campbell, R. (1992). Speech in the head? Rhyme skill, reading, and immediate memory in the deaf. In D. Reisberg (Ed.), *Auditory imagery*. Hillsdale, NJ: Erlbaum.

Farah, M. J. (1988). Is visual imagery really visual? Overlooked evidence from neuropsychology. *Psychological Review*, **95**, 307–317.

Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language*, **28**, 200–213.

Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language-disordered children: Is there a causal connection? *Journal of Memory and Language*, **29**, 336–360.

Geiselman, R. E., & Bjork, R. A. (1980). Primary vs. secondary rehearsal in imagined voices: Differential effects on recognition. *Cognitive Psychology*, **12**, 188–205.

Geschwind, N. (1979). Specializations of the human brain. *Scientific American*, **241**(3), 190–199.

Gracco, V. L., & Abbs, J. H. (1987). Programming and execution processes of speech movement control: Potential neural correlates. In E. Keller & M. Gopnik (Eds.), *Motor and sensory processes of language*. Hillsdale, NJ: Erlbaum.

Hanley, R. J., & Broadbent, C. (1987). The effect of unattended speech on serial recall following auditory presentation. *British Journal of Psychology*, **78**, 287–297.

Hebb, D. O. (1949). *Organization of behavior*. New York: Wiley.

Intons-Peterson, M. J. (1992). Components of auditory imagery. In D. Reisberg (Ed.), *Auditory imagery*. Hillsdale, NJ: Erlbaum.

Kean, M. L. (Ed.) (1985). *Agrammatism*. New York: Academic Press.

Kertesz, A., Harlock, W., & Coates, R. (1979). Computer tomographic localization, lesion size, and prognosis in aphasia and nonverbal impairment. *Brain and Language*, **8**, 34.

Kolers, P. A., & Roediger, H. L., III. (1984). Procedures of mind. *Journal of Verbal Learning and Verbal Behavior*, **23**, 425–449.

Kosslyn, S. M., Pinker, S., Smith, G., & Schwartz, S. P. (1979). On the demystification of mental imagery. *Behavioral and Brain Sciences*, **2**(3), 535–581.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, **21**, 1–36.

Liepmann, H. (1915). Diseases of the brain. In C. W. Barr (Ed.), *Curschmann's textbook on nervous diseases* (Vol. 1). Philadelphia: Blakiston.

Mackay, D. G., Allport, A., Prinz, W., & Scheerer, E. (1987). Relationships and modules within language perception and production: An introduction. In A. Allport, D. G. Mackay, W. Prinz, & E. Scheerer (Eds.), *Language perception and production: Relationships between listening, speaking, reading and writing*. New York: Academic Press.

Martin, R. C., Blossom-Stach, C., Yaffee, L. S., & Wetzel, W. F. (under review). Consequences of a motor programming deficit for rehearsal and written sentence comprehension.

Mateer, C. A. (1989). Neural correlates of language function. In D. P. Kuehn, M. L. Lemme, & J. M. Baumgartner (Eds.), *Neural bases of speech, hearing and language*. Boston, MA: Little, Brown.

Mayeux, R., & Kandel, E. R. (1985). Natural language, disorders of language, and other localizable disorders of cognitive functioning. In E. R. Kandel & J. H. Schwartz (Eds.), *Principles of neural science* (2nd ed.). New York: Elsevier.

McNaughton, B. L., & Morris, R. G. M. (1987). Hippocampal synaptic enhancement and information storage within a distributed memory system. *Trends in Neurosciences*, **10**(10), 408–415.

Mohr, J. P. (1976). Broca's area and Broca's aphasia. In H. Whitaker & H. Whitaker (Eds.), *Studies in neurolinguistics* (Vol. 1). New York: Academic Press.

Monsell, S. (1987). On the relation between lexical input and output pathways for speech. In A. Allport, D. G. Mackay, W. Prinz, & E. Scheerer (Eds.), *Language perception and production: Relationships between listening, speaking, reading and writing*. New York: Academic Press.

Mulford, R. (1988). First words of the blind child. In M. D. Smith & J. L. Locke (Eds.), *The emergent lexicon*. New York: Academic Press.

NAESER, M. A., & HAYWARD, R. W. (1978). Lesion localization in aphasia with cranial-computed tomography and Boston diagnostic aphasia exam. *Neurology*, **28**, 545.

PAPAGNO, C., VALENTINE, T., & BADDELEY, A. D. (1991). Phonological short-term memory and foreign-language learning. *Journal of Memory and Language*, **30**, 331–347.

PAULESU, E., FRITH, C. D., & FRACKOWIAK, R. S. J. (1993). The neural correlates of the verbal component of working memory. *Nature*, **362**, 342–345.

PYLYSHYN, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin*, **80**, 1–24.

PYLYSHYN, Z. W. (1981). The imagery debate; Analogue media vs. tacit knowledge. *Psychological Review*, **87**, 16–45.

SMITH, J. D., REISBERG, D., & WILSON, M. (1992). Subvocalization and auditory imagery: Interactions between the inner ear and inner voice. In D. Reisberg (Ed.), *Auditory imagery*. Hillsdale, NJ: Erlbaum.

WATERS, G. S., ROCHON, E., & CAPLAN, D. (1992). The role of high-level speech planning in rehearsal: Evidence from patients with apraxia of speech. *Journal of Memory and Language*, **31**, 54-73.

(Received October 7, 1993)

(Revision received May 26, 1994)