JOURNAL OF MEMORY AND LANGUAGE 33, 737–756 (1994)

Memory for Consonants versus Vowels in Heard and Lipread Speech

BEATRICE DE GELDER

Tilburg University, Tilburg, The Netherlands, and Université Libre de Bruxelles, Brussels, Belgium

AND

JEAN VROOMEN

Tilburg University, Tilburg, The Netherlands

Three experiments investigated short-term memory for consonants and vowels in heard and lipread material. In Experiment 1, serial recall of lipread and heard-plus-lipread lists containing vowel-varied or consonant-varied items was examined. Vowel-varied lists had larger recency effects than consonant-varied lists when lipread or heard-plus-lipread. In the following experiments, vowel-varied and consonant-varied lists were used which were heard-only, lipread, or heard-plus-lipread and which could be followed by a suffix presented in one of these modalities. The results clearly established that the recency and suffix effects of lipread speech are sensitive to the vowel/consonant contrast.

Speech is perceived not only by the ear but also by the sight of the speaker's lip movements. Studies undertaken since the late seventies have fully documented the theoretical as well as the practical importance of this second input modality for language (for overviews, see Dodd & Campbell, 1987; Massaro, 1987). As a consequence the notion that there are modality-specific aspects to language processing has evolved considerably. But before lipreading came into focus, the best-known domain of modality aspects of language processing was the contrasts between processing of spoken and written input in memory tasks. Indeed, one area of modality differences that has been studied intensively is recall performance in immediate memory tasks, which uncovered some characteristic phenomena of immediate memory that appeared to obtain only with spoken and not with written input (Corballis, 1966, Crowder & Morton, 1969). When attention turned to lipreading, the obvious question was whether memory for lipreading would resemble memory for written language rather than memory for spoken language input. The studies looking at modality effects in immediate memory have so far shown that lipread speech is processed like heard input and not like that of read input (Campbell, 1987; Campbell & Dodd, 1984; Greene & Crowder, 1984). If so, one is lead to expect that memory for lipread input is sensitive to phonological aspects of the material to be remembered, just like memory for heard input is. The experiments reported here pursue this research and examine this issue. The specific question addressed is whether a consonant/vowel contrast, which has been found repeatedly with auditory material (Crowder, 1971), is also a critical

Both authors contributed equally and the order of names was decided by the toss of a coin. We are much indebted to R. Crowder, D. Balota, and two anonymous referees for comments on an earlier version. Research was partly supported by the Ministry of Education of the Belgian French-speaking Community, Concerted Research Action "Language processing in different modalities: Comparative approaches." The research of the second author has been made possible by a fellowship of the Royal Netherlands Academy of Arts and Sciences. Address reprint requests to Beatrice de Gelder or Jean Vroomen, Tilburg University, Dept. of Psychology, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. e-Mail: degelder@kub.nl or vroomen@kub.nl.

0749-596X/94 $6.00

variable for understanding short-term memory for lipread information.

From Auditory Memory to Linguistic Processes

Immediate memory experiments have been a favorite domain for understanding differences in performance as a function of input modality. In a typical serial recall task subjects are asked to remember a list of items in the order as presented. If the probability of a correct response is plotted as a function of input positions, one usually obtains a U-shaped function in which the early and the last items are remembered better than the items in the middle of the list. The memory advantage of the final item, the so-called recency effect, is greater when the items were heard rather than when they were read silently; hence it is known as the modality effect. A second, closely related memory phenomenon also sensitive to modality of input is the suffix effect. If a redundant item is spoken after the end of the list—for example, an experimenter saying "STOP"—recency for heard lists is reduced, whereas recency for written lists is much less disturbed (Crowder & Morton, 1969).

The accepted explanation of the modality and the suffix effect was to take them as manifestations of an acoustic sensory memory known as precategorical acoustic storage (PAS) (Crowder & Morton, 1969). The modality effect arises as the supplementary contribution of PAS to the recall of the final heard memory item and it is therefore not observed with written input. The effect of a heard suffix is that it removes the final item from its privileged position by overwriting PAS. This notion of PAS has come under strong attack by studies of lipread input showing some similarities between hearing and lipreading. Spoehr and Corin (1978) were among the first who found that the recency effect of heard-plus-lipread digit lists was disrupted as much by a heard-only suffix as by a lipread suffix. Another important similarity between memory for heard and lipread input was studied by Campbell and Dodd (1980) who found that lipread lists had larger recency effects than written lists. Similarly, Greene and Crowder (1984) found that recency effects for lipread lists were larger than to be expected for lists read silently. The disruptive effect of a suffix on recency was also comparable for heard and lipread memory lists. For instance, Campbell and Dodd (1982) showed that lipread recency was impaired by a heard-only suffix. In line with these observations, Gardiner, Gathercole, and Gregg (1983) found that heard-plus-lipread recency is disrupted by a lipread distracter task, but not by a silent reading. Moreover, Gathercole (1987) demonstrated that heard-plus-lipread and lipread suffixes had comparable disruptive effects on heard-plus-lipread or lipread lists, whereas a written suffix did not interfere.

Overall, the symmetries between hearing and lipreading seem to suggest that recency and suffix effects reflect the activity of a more abstract storage system accessed by both heard and lipread material rather than an auditory-derived sensory store like PAS was taken to be (e.g., Campbell, 1987; Campbell & Dodd, 1984; Crowder, 1983; Greene & Crowder, 1984). A strong version of this view states that recency of heard and lipread lists might be based on a single common memory code. For example, Nairne (1988, 1990) asserts that the trace features on which recency depends are the same for heard and lipread speech because both have the same modality-specific trace features ("modality" to be understood here as input route for spoken language processing). In a similar vein, Crowder and Surprenant (in press) argue that recency effects of heard and lipread speech are equivalent because both result from the involvement of the phonetic processor. Similarly, in studies where subjects were asked to silently articulate a visually presented list (mouthing), recency and suffix effects were observed as in auditorily presented lists (Nairne & Walters, 1983). Thus, recency and suffix effects which were once the signature of an acoustic sensory memory are now taken as reflecting linguistic processing. It appears then that similarities across modalities between recency and suffix effects become strong evidence for the notion that abstract phonetic processes are critical.

Attractive as this may be from a linguistic point of view, such an approach might be too strong. For instance, de Gelder and Vroomen (1992) compared three input modalities of memory lists, i.e., heard-only, lipread, and heard-plus-lipread lists presented with different suffix conditions and observed consistent modality-specific effects. The study showed that a lipread suffix had a stronger effect on recency than a heard-only suffix in a lipread list, whereas a heard-only suffix was more damaging than a lipread suffix in heard-only lists. Moreover, Nairne and Crowder (1982) found that the effect of a mouthed suffix was much weaker than that of an auditory suffix. Turner et al. (1987), in a comprehensive study, also concluded that recency and suffix effects for auditory and mouthed lists were not identical.

The goal of this study is to explore further possible symmetries between memory for heard and lipread speech. The central issue is whether memory for lipread input is sensitive to the vowel/consonant contrast. Earlier studies have established that the extent to which recency effects of heard lists are obtained, appears to depend on some specific characteristics of the vocabulary from which the items are taken. Lists which vary in consonants (e.g., /ba, da, ga/) have small recency effects when compared with lists which vary in vowels (e.g., /bi, ba, bo/) (Cole, 1973; Darwin & Baddeley, 1974). This fact—referred to as the vowel/consonant contrast or the vowel advantage—was one of the findings that lent strong support to the PAS theory, as it seemed to present clear evidence that typical acoustic factors (like stable formants) played a role in short-term memory. Besides, the fact that vowel-varied lists showed more recency than consonant-varied lists fitted well with the notion of auditory memory widely accepted in the seventies (Cole, 1973; Crowder, 1971; Darwin & Baddeley, 1974). If lipreading is a proper speech input modality, there is reason to expect that it is influenced by the same linguistic dimensions of spoken language as heard speech. Still, Turner et al. (1987) found that the difference between vowel-varied and consonant-varied lists was not identical for heard and mouthed lists. An investigation of the vowel advantage with lipread speech might extend our understanding of the similarities between the heard and lipread speech modalities. In return, as we shall see, the finding that memory for lipread material would be affected by the same factors as memory for auditory material would suggest quite clearly that the processes and representations in immediate memory might indeed be abstract linguistic in nature.

Input Modality versus Modularity of Speech Processes

What predictions can be derived from the short history of heard and lipread speech memory research? Although initially the finding of a vowel advantage in memory added to the evidence in favor of PAS, subsequent research on lipread speech brought linguistic dimension of immediate memory to the foreground. The older perspective offers little reason to expect that the vowel/consonant contrast will show up in lipread lists. Such a negative finding would be in line with the argument claiming that larger recency effects for vowels are found because they persist longer in acoustic memory than in consonants (cf. Cole, 1973; Greene and Crowder, 1984; Pisoni, 1973, 1975). If an acoustic property were the critical difference between vowels and consonants, there is no reason to expect that a similar contrast should be observed in lipread lists since they have no acoustic properties at all. The recent perspective of a speech module, however, moves one a step away from explanations exclusively based on acoustic properties. A module provides a niche for lipreading data and the emerging similarities in immediate memory phenomena for heard and lipread speech. Modality-specific aspects (heard versus lipread) would belong to early, peripheral, sensory-based stages, whereas module-based or modular processes are abstract and ignore modality differences.

These two alternatives are not entirely clear though. One critical issue is that of locus of persistent modality differences. In fact, in the context of modality effects, the notion of an abstract speech module is somewhat ambiguous. Processes and representations might be termed "abstract" when the focus is on the contrast between peripheral sensory aspects and module based processes. However, if one grants that hearing and lipreading are two speech modalities, there is still room for modality differences within a module-based view. As a matter of fact, logically speaking the notion of a specialized speech module carries no implications for specificity of input modality. Thus, modality on input to the speech module and modularity of the representations and processes are orthogonal issues (Cutler, 1989; de Gelder & Vroomen, 1989). Although the present study does not focus on it, the issue of modality specificity of short-term memory systems will arise as the suffix paradigm will be used in Experiments 2 and 3.

The present experiments investigate whether memory for lipread lists would exhibit a vowel/consonant contrast and how such a result fits the just mentioned theoretical alternatives. In Experiment 1, vowel-varied or consonant-varied lists were heard-plus-lipread or lipread. The subsequent experiments pursue that issue by comparing the two input modalities separately and in combination, thereby examining the effect of different kinds of suffixes. One expects that heard vowel-varied lists should show a strong recency effect, whereas consonant-varied lists (heard and lipread) do not or only in a limited way. When lipread vowel-varied list have a strong recency effect, the advantage of vowels would not derive from peripheral auditory processes like stable formant configurations or other parameters of acoustic nature, but it would reflect processing activities within a phonetic processor or speech module.

Experiment 1

The experiment was designed to compare recency effects of heard-plus-lipread and lipread lists which varied in consonants or vowels. We expected that heard-plus-lipread lists varying in consonants would have small recency effects when compared with vowel-varied lists (Cole, 1973, Crowder, 1971). If this consonant/vowel difference is based on processes within an a-modal speech processor, analogous results are expected for lipread lists. If, on the other hand, an acoustic property underlies the consonant/vowel contrast, a large vowel/consonant contrast should be observed in heard-plus-lipread lists, though not in lipread lists.

Method

Subjects. Thirty-two students (Male = 12, Female = 20) were tested. All were native Dutch speakers with no reported hearing or seeing deficit.

Material. A female speaker reciting the digit lists was recorded on Sony U-matic video. The video picture showed the whole of the speaker's face under even lighting conditions. In the heard-plus-lipread condition, heard as well as lipread information was present. For the lipread presentation, the sound track was deleted from the master forcing the subjects to rely entirely on lipreading. Lists of seven naturally spoken items were presented at a rate of three items per 2 s. They were drawn pseudorandomly with replacement from visually distinct consonant-varied (i.e., /ba, da, va/) or vowel-varied vocabularies (i.e., /bi, ba, bo/) (see Jackson, Montgomery, & Binnie, 1976; Woodward & Barber, 1960). The lists were lipread or heard-plus-lipread which resulted in four conditions. Each condition consisted of 10 lists. Within a list, no item appeared more than twice in a row. Lists were preceded by a 200-ms warning tone.

Procedure. Each of the four conditions was presented in one block of 10 trials. Because of potential unequal practice effects, the presentation mode of the lists was not counterbalanced and the heard-plus-lipread lists were presented first, followed by the lipread lists. For half of the subjects the consonant-varied items were presented first followed by the vowel-varied items, for the other half the order was reversed. The experiment was preceded by a practice session in lipreading where each memory item had to be lipread three times in isolation. Practice continued with eight memory lists in which the four experimental conditions occurred twice. Instructions specified the presentation mode of the lists, its length, the vocabulary from which the items were taken, and the number of trials. Subjects were asked to write down their responses from left to right on a prepared sheet within 12 s. No backtracking was allowed. Subjects were told that responses would only be scored correct if they were in the correct position and encouraged not to leave blanks but to guess if unsure. They were closely supervised during the experiment to ensure that instructions were followed. Subjects were tested in groups of four to eight. The response sheets were copied into a computer and all scoring and tests were done by program.

Results

The mean proportion of correct recall as a function of the serial position for the vowel-varied and consonant-varied items for each presentation mode is shown in Fig. 1. Recency of vowel-varied lists was larger than that of consonant-varied lists; heard-plus-lipread lists had larger recency effects than lipread lists; and these two factors in-

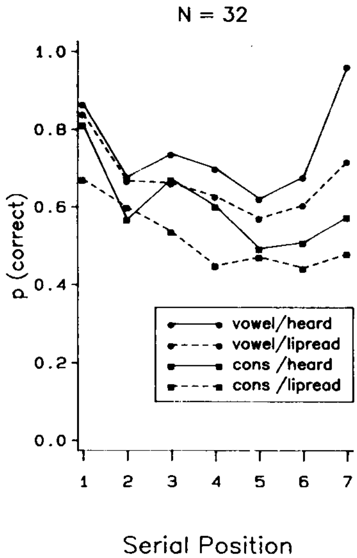

FIG. 1. Experiment 1: Proportion of correct responses as a function of serial position for the lipread and heard-plus-lipread lists.

teracted: recency was largest when vowel-varied lists were heard-plus-lipread.

These generalizations were supported by a 2 (List mode) × 2 (Vocabulary) × 7 (Serial position) analysis of variance (ANOVA) on the proportion of correct responses. There was a significant main effect of list mode [$F(1,31) = 26.69, p < .001, MS_e = .05$] because heard-plus-lipread lists were recalled better than lipread lists. The effect of vocabulary was significant [$F(1,31) = 57.15, p < .001, MS_e = .08$] because of the superior recall of the vowel-varied lists. There was a main effect of serial position [$F(6,186) = 29.08, p < .001, MS_e = .03$] and there were significant interactions among presentation mode and serial position [$F(6,186) = 6.07, p < .001, MS_e = .02$] and among vocabulary and serial position [$F(6,186) = 8.48, p < .001, MS_e = .02$]. The second-order interaction among presentation mode, vocabulary, and serial position was also significant [$F(6,186) = 3.24, p < .005, MS_e = .02$].

To analyze recency effects, recency was calculated as the difference in number of items recalled between the last serial position and the average of positions 5 and 6.

This difference score was chosen because it partials out the effect that mispreceptions might occur more often in lipread lists than in heard-plus-lipread lists. Several other measures of recency have appeared in the literature, and we wanted to make sure that our results were not depending on our particular measure of recency. We therefore also calculated recency as the difference between the final item and the average of all other serial positions. Because this measure led to identical conclusions, only the analyses of the first measure will be reported. These mean recency scores are presented in Table 1.

A 2 (Vocabulary) × 2 (List mode) ANOVA on the recency scores indicated that lipread lists had smaller recency effects than heard-plus-lipread lists [$F(1,31) = 14.12, p < .001, MS_e = .03$], and that consonant-varied lists had smaller recency effects than vowel-varied lists [$F(1,31) = 25.13, p < .001, MS_e = .03$]. Of interest was a significant interaction among list mode and vocabulary [$F(1,31) = 4.35, p < .05, MS_e = .03$]. Post hoc tests (Fisher's LSD, $\alpha = .01$, critical difference = .08) based on the error term of the omnibus interaction confirmed that the recency effect of heard-plus-lipread lists which varied in vowels was significantly larger than all other recency effects, and that lipread vowel-varied lists had larger recency effects than lipread consonant-varied lists.

Discussion

The question addressed in this experiment was whether recall of lipread and heard-plus-lipread memory lists was affected by the consonant/vowel distinction. The results showed that lipread as well as heard-plus-lipread lists had larger recency effects when the items varied in vowels rather than in consonants. This finding clearly suggests that sensitivity for the consonant/vowel distinction might not be a matter of acoustic differences between vowels and consonants. We discuss this result and relate it to the complementary observation which is the interaction that appears to exist between the modality of the list and the vocabulary of the items.

The results obtained with heard-plus-lipread lists confirm earlier findings on the advantage of vowel-varied lists with auditory presentation (Crowder, 1971). More importantly, our present results with lipread lists show that the phenomenon of a memory advantage of vowel-varied lists extends to lipread material. As the peripheral or sensory properties of lipread speech are very different from those of heard speech, this result calls for a characterization of the vowel/consonant contrast in other than only peripheral auditory terms. The notion of sonority (e.g., Selkirk, 1984) puts the consonant/vowel contrast in a linguistic perspective. Vowels and liquids (like /l/ and /r/) are usually grouped high in the sonority hierarchy, nasals (like /m/ and /n/) are intermediate, and obstruents (like /p/ or /f/) are low in sonority. Telling from the present results, memory for lipread syllables appears to respect a sonority dimension. So our results confirm for lipread presentation what has recently been observed for auditory stimuli by Surprenant and Speer (1990). Their finding that recency effects vary with the linguistic dimension of sonority appears to carry over from the domain of heard to that of lipread speech. Besides, appealing to a sonority scale might explain that a small recency effect was still found in the consonants-varied lists since one of our consonant-varied items was the labiodental /va/ which has a moderate sonority value. The present result and its generaliza-

TABLE 1
MEAN RECENCY SCORES FOR EACH LIST MODE AND VOCABULARY

Vocabulary	List mode: Lipread	List mode: Heard-plus-lipread	
Consonants	.02	.07	.05
Vowels	.13	.31	.22
	.08	.19	

tion in terms of sonority differences between list items appear as straightforward support to the strong hypothesis that recency effects are the signature of the involvement of an abstract phonetic processor.

The data do not support an acoustic explanation of the vowel advantage. That perspective is then also unsatisfactory to explain the second finding of this experiment, i.e. that there is superior recency when there is an acoustic component, like in heard-plus-lipread lists. What might explain this recency advantage of heard-plus-lipread over lipread presentations? One possibility we cannot rule out given the present stimuli is that the recency advantage of the heard-plus-lipread lists over lipread lists is due to the fact that the former is a bimodal presentation and the latter unimodal. This possibility will be studied in the following experiment where heard-only and lipread lists are compared with heard-plus-lipread lists.

What other explanations for the superior recency of heard-plus-lipread lists are available? A possibility is that more traces are available at recall when the items are heard-plus-lipread instead of lipread because the heard component is richer in phonetic information. In line with this suggestion are the results of Campbell, Garwood, and Rosen (1988) who found that lipread suffix was less effective than a heard-plus-lipread suffix. They concluded that lipreading gives rise to a phonetic percept that is underspecified with respect to at least some component of that trace. For instance, lipread information does not convey voicing. The idea is thus that hearing-plus-lipreading leaves a phonetically "richer" trace than lipreading does and that the recency advantage is based on this phonetic trace. The more fully specified that trace, the more recency is observed on the last item. It looks particularly appropriate then to examine this notion of relatively less specified traces by making use of the suffix phenomenon. When an extra utterance (the suffix) is appended at the end of a list, the advantage of the final memory item is removed because the suffix overwrites traces on which recency was based (Crowder & Morton, 1969). A prediction suggested by the above notion would be that a suffix will be more effective when it is presented in the modality that one believes to leave a richer trace. If hearing leaves more phonetic traces than lipreading does, one would expect that recency effects in heard-only as well as heard-plus-lipread lists are larger than in lipread lists. Following this line of thought, one might also find that heard-plus-lipread suffixes and heard-only suffixes are more effective than lipread suffixes. The following experiments addressed that issue.

EXPERIMENTS 2A AND 2B

Experiment 2 consists of two parts. Experiment 2a examined visually distinct vowel-varied items (i.e./pi, pa, po/) that were heard-only, lipread, or heard-plus-lipread. The lists were followed by a pure tone or by a suffix presented in one of these modalities. In Experiment 2b, consonant-varied items (/pa, ta, fa/) were used. Even if no recency and, therefore, no suffix effect were expected in this latter case, comparing the two sets of data might allow a further analysis of the vowel advantage for lipread speech.

Method

Subjects. The group of Experiment 2a consisted of 64 students (Male = 28, Female = 36). Thirty-eight students (Male = 12, Female = 26) participated in Experiment 2b. All subjects were native Dutch speakers with no reported hearing or seeing deficit. No subject participated in the previous experiment.

Materials. A recording was made of a female speaker reciting the lists on Sony U-matic video. The video picture showed the whole of the speaker's face under even lighting conditions. In the heard-plus-lipread condition subjects could hear the speaker as well as watch her lips move. For lipread lists, the sound track was deleted forcing the subjects to rely on lipreading. For the heard-only presentation, the items were spoken into an Uher reel-reel tape recorder. This soundtrack was inserted into the videotape showing the speaker sitting quietly. Note that in the heard-only condition the face remained visible. In this way a lipread or a heard-plus-lipread suffix could be presented without a sudden fade-in of the speaker.

Lists of seven naturally spoken items were presented at a rate of three items per two sec. For Experiment 2a the items were drawn pseudorandomly with replacement from the visually distinct vocabulary /pi, pa, po/ (Jackson, Montgomery, & Binnie, 1976). For Experiment 2b the items were /pa, ta, fa/. There were three presentation modes: heard-only, lipread, and heard-plus-lipread. There were four suffix conditions: a no-suffix condition in which a 200-ms tone of 1000 Hz was presented after list presentation, and three speech suffixes presented in either the heard-only, lipread, or heard-plus-lipread mode. Thus, there were 12 conditions. Each condition consisted of six lists for a total of 72 experimental lists. Within a list, no item appeared more than twice in a row. Across lists, in each condition each item occurred twice on each serial position. Suffixes were presented in rhythm with item presentation. The word "STOP" was chosen as suffix because a pilot study had shown that it was visually distinct from the memory items. Lists were preceded by a 200-ms warning tone. To ensure that subjects were watching the screen during presentation of a suffix, the face of the speaker faded out right after its presentation and it was immediately followed by a colored circle appearing for 1 s in the middle of the screen. A green circle required subjects to start responding by recalling the list, a red circle (i.e., a catch trial) signalled subjects to write down crosses on the response sheet.

Procedure. Each presentation mode of the lists consisted of two blocks of 12 trials each. In order to counterbalance practice effects, the modality of the list-blocks was heard-plus-lipread, heard-only, lipread, lipread, heard-only, and heard-plus lipread. In each block, the four suffix conditions were presented three times in random order with no more than two consecutive trials in the same suffix condition. Each block was preceded by a warm-up list. There was one catch trial in each block. The experiment was preceded by a practice session in lipreading where each memory item and the suffix word "STOP" was presented silently three times in isolation. Practice continued with 27 memory lists in which the 12 experimental conditions occurred twice, including three catch trials. Instructions specified the modality of the lists, its length, and the number of trials. Subjects were told to "ignore" the suffix and to look at the video screen until a green or a red circle appeared. Then they were asked to write down their responses from left to right on a prepared sheet. They were given 12 s for this. No backtracking was allowed. Subjects were told that responses would only be scored correct if they were in the correct position. They were encouraged not to leave blanks but to guess if they were unsure. They were closely supervised during the experiment to ensure that they followed instructions. Subjects were tested in groups of 16 to 20. The response sheets were copied into a computer and all scoring and tests were done by program.

Results

Results are given separately for Experiments 2a and 2b, and then a comparison is made with Experiment 1. For Experiment 2a, the mean proportion of correct responses as a function of the serial position is presented in Fig. 2, separately for each modality of the list and suffix.

A 3 (List mode) × 4 (Suffix) × 7 (Serial position) ANOVA was performed on the proportion of correct responses. There was a main effect of list mode [$F(2,126) = 62.04, p < .001, MS_e = .09$]. Post hoc tests

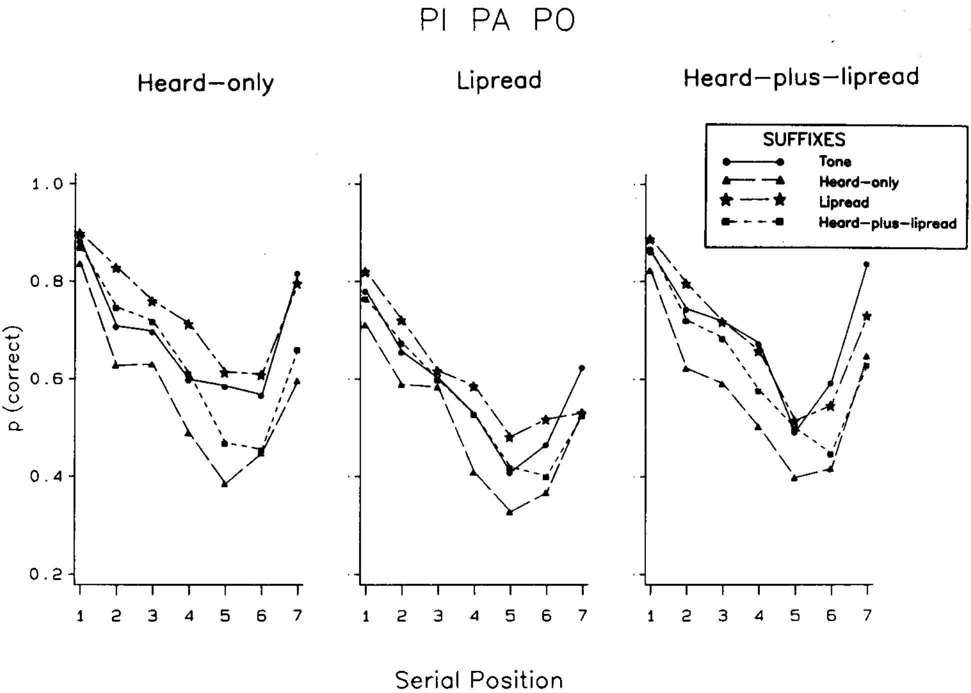

FIG. 2. Experiment 2a: Proportion of correct responses as a function of heard-only, lipread, and heard-plus-lipread lists for vowel-varied CV items.

(Fisher's LSD, α = .01) indicated that this was mainly due to the lower recall for the lipread lists. The main effects of suffix [$F(3,189) = 65.25$, $p < .001$, $MS_e = .07$], and serial position were significant [$F(6,378) = 112.26$, $p < .001$, $MS_e = .11$]. There were significant interactions among list mode and serial position [$F(12,756) = 3.16$, $p < .001$, $MS_e = .03$], among list mode and suffix [$F(6,378) = 2.94$, $p < 01$, $MS_e = .05$], and among suffix and serial position [$F(18,1134) = 7.41$, $p < .001$, $MS_e = .03$]. The second-order interaction among list mode, suffix, and serial position was also significant [$F(36,2268) = 1.75$, $p < .005$, $MS_e = .03$]. Inspection of the figure suggests that the second-order interaction is located in the recency part of the curves.

Recency was calculated as in experiment 1 (see Table 2). A 3 (List mode) × 4 (Suffix) ANOVA was performed on the recency scores. There was a main effect of list mode [$F(2,126) = 15.22$, $p < .001$, $MS_e = .04$] because lipread lists had less recency than heard-only and heard-plus-lipread lists (Fisher's LSD, α = .01). The effect of a suffix was significant [$F(3,189) = 9.19$, $p < .001$, $MS_e = .04$] as was the interaction between the modality of the list and the suffix [$F(6,378) = 2.36$, $p < .05$, $MS_e = .05$].

Post hoc tests (Fisher's LSD) were performed which were based on the error term of the interaction. For each list mode, the

TABLE 2
MEAN RECENCY SCORES FOR EACH SUFFIX AND LIST MODE OF /pi, pa, po/ LISTS

	List mode			
Suffix	Heard-only	Lipread	Heard-plus-lipread	
Tone	.24	.19	.30	.24
Heard-only	.18	.19	.24	.20
Lipread	.18	.03	.20	.14
Heard-plus-lipread	.20	.12	.16	.16
	.20	.13	.23	

recency effect in the suffix conditions was compared with the appropriate no-suffix condition. The tests indicated that a lipread suffix decreased recency in lipread lists ($\alpha = .01$, critical difference = .098), and heard-plus-lipread and lipread suffixes ($\alpha = .05$, critical difference = .073) were effective in heard-plus-lipread list. Comparing recency effects among the no-suffix conditions indicated that recency of lipread lists was smaller than that of heard-plus-lipread lists ($\alpha = .01$).

We now turn to Experiment 2b in which consonant-varied items (/pa, ta, fa/) were used. The results are plotted in Fig. 3. As can be seen, recency effect were small if compared with Experiment 2a.

On the proportion of correct responses, a 3 (List mode) × 4 (Suffix) × 7 (Serial position) ANOVA was run with list mode, suffix, and serial position as within-subjects variables. There was a main effect of presentation mode of the list [$F(2,74) = 13.06$, $p < .001$, $MS_e = .06$]. Post hoc tests (Fisher's LSD, $\alpha = .01$) indicated that recall of the lipread list was lower. The effect of a suffix was significant [$F(3,111) = 18.67$, $p < .001$, $MS_e = .06$], as was serial position [$F(6,222) = 98.15$, $p < .001$, $MS_e = .05$]. There was a significant interaction among suffix and serial position [$F(18,666) = 5.50$, $p < .001$, $MS_e = .03$], and a significant second-order interaction among modality of the lists, suffix, and serial position [$F(36,1332) = 1.59$, $p < .02$, $MS_e = .03$]. All other interactions were not significant (all $p > .10$).

Recency was calculated as in the previous experiments in order to investigate whether the second-order interaction was located in the theoretically interesting recency part of the serial position curves. Mean recency scores are presented in Table 3. A 3 (List mode) × 4 (Suffix) ANOVA on the recency scores indicated that there was a main effect of suffix [$F(3,111) = 4.77$, $p < .005$, $MS_e = .06$], and of modality of the list [$F(2,74) = 3.51$, $p < .05$, $MS_e =$

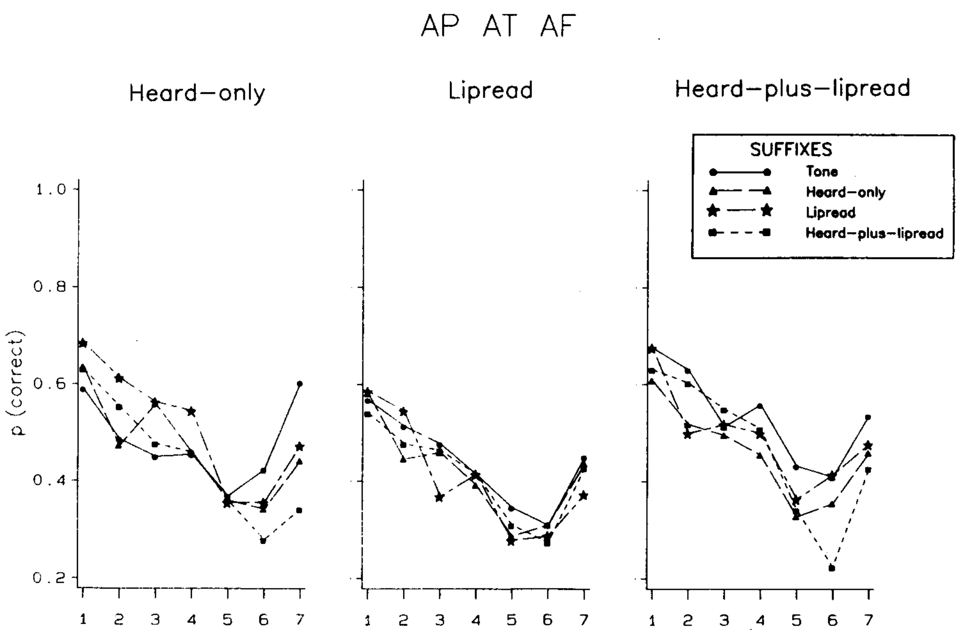

FIG. 3. Experiment 2b: Proportion of correct responses as a function of heard-only, lipread, and heard-plus-lipread lists for consonant-varied CV items.

TABLE 3
MEAN RECENCY SCORES FOR EACH SUFFIX AND LIST MODE OF /pa, ta, fa/ LISTS

	List mode			
Suffix	Heard-only	Lipread	Heard-plus-lipread	
Tone	.11	.00	.09	.06
Heard-only	.08	.06	.10	.08
Lipread	.07	−.03	−.09	−.02
Heard-plus-lipread	.01	−.05	.00	−.02
	.07	−.01	.02	

.06], but the interaction between modality of the list and suffix was not significant [$F(6,222) = 1.59, p = .150, MS_e = .05$].

A separate ANOVA was performed which compared the recency scores from the no-suffix conditions of Experiments 2a and 2b. The vocabulary of which the items were taken (vowel-varied or consonant-varied items) was treated as a between-subjects variable, the modality of the list was within-subjects. The ANOVA showed that there was a significant main effect of vocabulary [$F(1,100) = 28.85, p < .001$], because vowel-varied items had larger recency effect than consonant-varied items. The modality of the list was also significant [$F(2,200) = 6.63, p < .002$]. Post-hoc tests (Fisher's LSD, $\alpha = .01$, critical difference = .072) showed that, across the two experiments, lipread lists had smaller recency effect than heard-only and heard-plus-lipread list. The interaction between vocabulary and modality of the list was not significant [$F(2,200) < 1$]. Thus, experiments 2a and 2b replicate and extend the major finding of experiment 1: vowel-varied items have more recency than consonant-varied items, whether heard or lipread. Moreover, heard-plus-lipread and heard-only lists have larger recency effects than lipread lists.

Discussion

Experiments 2a and 2b tried to replicate the finding that lipread lists are sensitive to the vowel/consonant contrast (Experiment 1). A second goal was to understand why, against this background of similarity, hearing-plus-lipreading still shows a recency advantage over lipreading alone. The comparison of hearing-plus-lipreading with hearing-only explores the causes of this difference. Moreover, the suffix effects should allow us to explore the idea of phonetic underspecification of the visual speech representation.

The results show a replication, strengthening, and extension of the previous finding of the sensitivity of lipread lists to the consonant/vowel distinction. We first look at the results from the no-suffix conditions. If one compares Tables 2 and 3, it is clear that vowel-varied lists have larger recency effects than consonant-varied lists irrespective of whether they are lipread or heard-plus-lipread. Thus, recency effects are larger if the critical information of the list items is a vowel rather than a consonant, irrespective of the presentation modality. Hence, an acoustic explanation cannot be the reason for differences in recency between vowels and consonants. Beside this major result, we again observe that heard-plus-lipread and heard-only lists have larger recency effects than lipread lists. But we can exclude the possibility that the advantage of the heard-plus-lipread lists is due to a bimodal input, since heard-only lists are remembered as well as heard-plus-lipread lists. This confirms previous results where it was also found that a double input condition did not boost memory performance (de Gelder & Vroomen, 1992).

With this new result of a vowel advantage for lipread speech, we can readdress the issue of the superior recency of heard-plus-lipread lists. Ideally, the finding of selective superior recency might contribute to our understanding of the generalised vowel advantage. The modality advantage of hearing the list items seems to point to an effect that appears to be specific for processing speech in the heard modality. A possible explanation we mentioned already would have it that the larger recency effect in heard-only and heard-plus-lipread lists is derived from an extra contribution that hearing might make to recency by offering a richer phonetic trace. We can now appreciate this suggestion by looking at differences in the suffix effects. If the heard trace is phonetically richer, one should find that a heard-plus-lipread and heard-only suffix is at least as powerful or more so than a lipread suffix and, therefore, has an equally strong if not stronger effect on recency of a lipread list. Likewise, a lipread suffix should be less effective than the other suffixes because it has less potential to overwrite the richer auditory trace.

This prediction from the phonetic trace approach was, however, only partially confirmed. The data show that the strongest suffix effects are obtained when the modality of list and suffix match (see also de Gelder & Vroomen, 1992). For instance, a lipread suffix is more powerful with lipread lists than a heard-only suffix. Thus, the suggestion that a heard representation is phonetically richer than a lipread representation appears by itself insufficient to explain the data. The present results again point to the importance of modality-specific aspects. A way of taking up this challenge of modality differences is to see whether they can be conceptualized within in broader framework than that suggested by the phonetic trace approach which limits modality differences to a quantitative contrast only, namely that of poor versus rich traces. Instead, one might contrast a quantitative notion of modality specificity, where a rich trace is a matter of more information, with a qualitative notion of modality specificity. In the latter perspective, there is more to differences in traces originating from different modalities of input than the amount of information conveyed by the respective modalities.

The next experiment was designed in order to assess whether one might generalize the previous findings to memory items with a different linguistic structure. The order of the consonant and vowel segments in Experiments 2a and 2b was the same since a consonant was always followed by a vowel in a typical CV syllable. In Experiment 2b this meant the use of rhyming items, which allowed us to observe that lipreading is sensitive to the vowel/consonant contrast. But at the same time the logic of the vowel/consonant contrast introduces a possibly confounding factor, i.e. the use of rhyming items only in Experiment 2b, which complicates the comparison between the Experiments 1, 2a, and 2b. In order to extend the generality of the effects observed so far, another experiment was designed where the syllable structure or the order of critical segments was inverted. Instead of /pi, pa, po/, we used /ip, ap, op/ as memory items and the /pa, ta, fa/ were replaced with lists consisting of /ap, at, af/ items. This comparison maintains the syllable structure symmetry and rules out the effect of rhyme.

Experiments 3a and 3b

Consonant and vowel-varied lists are again contrasted, but the difference with Experiments 2a and 2b is that the syllable structure is VC instead of CV. As a consequence, the items in Experiment 3b do not rhyme, and a more straightforward comparison of the vowel effect in the lipread modality is possible.

Method

Subjects. For Experiment 3a the group consisted of 59 students (Male = 22, Female = 37), and for Experiment 3b the group consisted of 37 students (Male = 11, Female = 26) who participated. No subject took part in any of the previous experiments.

Materials. In Experiments 3a and b, exactly the same stimuli and procedures were used as in Experiments 2a and b, except that the /pi, pa, po/-items were replaced with /ip, ap, op/, and /pa, ta, fa/ items were replaced with /ap, at, af/.

Results

Results are first presented for the /ip, ap, op/ and /ap, at, af/ sets separately and are then followed by the appropriate comparisons. Mean recall accuracy for each serial position and for each condition was calculated, and these data were used for analyses. Figure 4 presents for the /ip, ap, op/ items the mean proportion of correct responses as a function of the serial position, separately for each presentation mode of the list and suffix. As can be seen, the results closely resemble those of Experiment 2a. Lipread lists have small recency effects when compared with the other modalities, and the effect of a suffix depends on the match between the modality of the list and the suffix.

A 3 (List mode) × 4 (Suffix) × 7 (Serial Position) ANOVA was performed on the proportion of correct responses. There was a main effect of list mode [$F(2,116) = 42.75, p < .001, MS_e = .10$], because of the lower recall of the lipread lists (Fisher's LSD, $\alpha = p < .001, MS_e = .10$. The effect of a suffix was significant [$F(3,174) = 68.01, p < .001, MS_e = .07$], as was serial position [$F(6,348) = 85.78, p < .001, MS_e = .08$]. There were significant interactions among list mode and serial position [$F(12,696) = 3.14, p < .001, MS_e = .03$], and among suffix and serial position [$F(18,1044) = 7.41, p < .001, MS_e = .03$]. The second-order interaction among list mode, suffix, and serial position was also significant [$F(36,2088) = 2.30, p < .001, MS_e = .03$].

Recency was calculated as in the previous experiments. Mean recency scores are presented in Table 4.

A 3 (List mode) × 4 (Suffix) ANOVA on the recency scores indicated that there was a significant effect of list mode [$F(2,116) = 7.04, p < .001, MS_e = .05$], mainly because lipread lists had less recency than in the

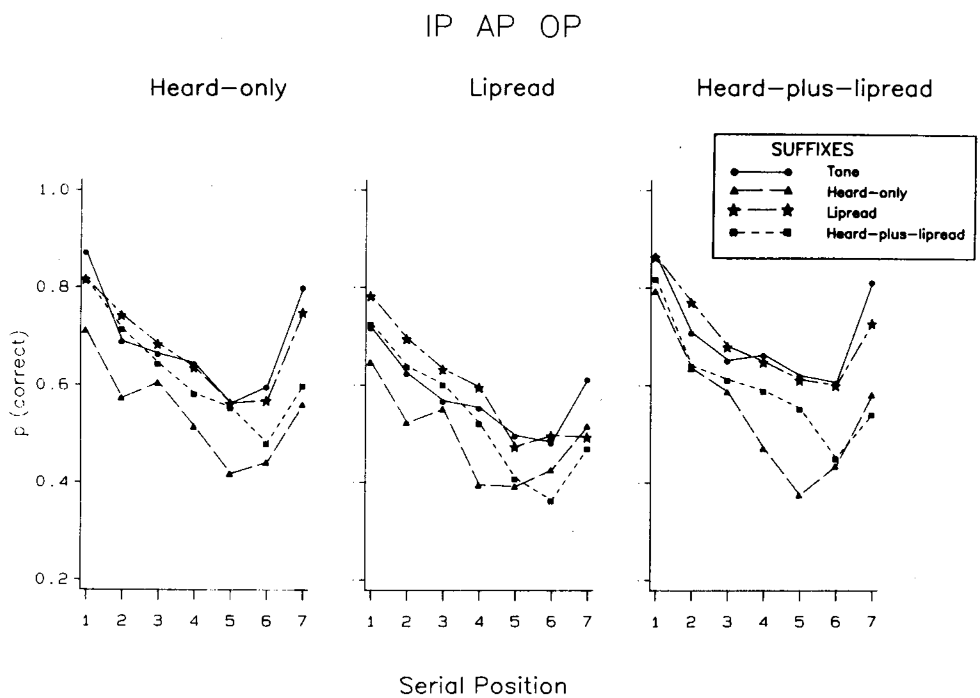

FIG. 4. Experiment 3a: Proportion of correct responses as a function of heard-only, lipread, and heard-plus-lipread list for vowel-varied VC items.

TABLE 4
MEAN RECENCY SCORES FOR EACH SUFFIX AND LIST MODE OF /ip, ap, op/ LISTS

	List mode			
Suffix	Heard-only	Lipread	Heard-plus-lipread	
Tone	.13	.04	.13	.10
Heard-only	.02	.03	.03	.02
Lipread	.08	−.12	.03	.00
Heard-plus-lipread	−.03	−.07	−.07	−.05
	.05	−.03	.03	

other modalities. The effect of a suffix was significant [$F(3,174) = 9.36, p < .001, MS_e = .04$], as was the interaction between list mode and suffix [$F(6,348) = 3.23, p < .005, MS_e = .04$]. Post hoc tests based on the error term of the interaction revealed that the heard-only suffix decreased recency in heard-only lists (Fisher's LSD, $\alpha = .05$, critical difference = .072), but there was no effect on recency in lipread or heard-plus-lipread lists. The lipread suffix decreased recency in lipread lists (Fisher's LSD, $\alpha = .01$, critical difference = .096) and in heard-plus-lipread lists ($\alpha = .05$), but there was no effect of a lipread suffix in heard-only lists. A heard-plus-lipread suffix decreased recency in heard-only and heard-plus-lipread lists ($\alpha = .01$), but not in lipread lists. A comparison among the no-suffix conditions showed that recency in lipread lists was less than in heard-only ($\alpha = .01$) and heard-plus-lipread lists ($\alpha = .05$). Experiment 3a thus again replicated that lipread lists have less recency than in the other modalities, and that the effect of a suffix depends on the match between the modality of the list and the suffix.

We now turn to the Experiment 3b. The results, as plotted in Fig. 5, show that the recency effects of /ap, at, af/ items are somewhat larger than those of the rhyming /pa, ta, fa/ items of Experiment 2b. Hence, the consonant/vowel contrast is attenuated.

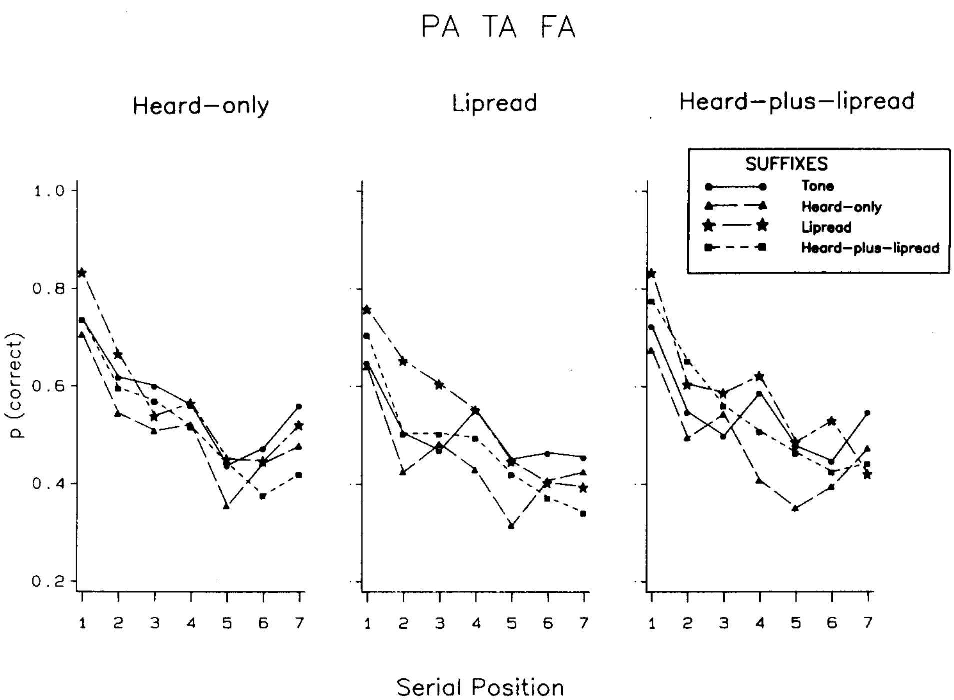

FIG. 5. Experiment 3b: Proportion of correct responses as a function of heard-only, lipread, and heard-plus-lipread list for consonant-varied VC items.

This seems to be the case for all modalities in which the lists are presented.

A 3 (List mode) × 4 (Suffix Condition) × 7 (Serial Position) ANOVA was performed on the proportion of correct responses. There was a main effect of list mode [$F(2,72) = 20.05$, $p < .001$, $MS_e = .07$]. Post hoc tests (Fisher's LSD, $\alpha = .01$) confirmed that this was due to the lower recall of the lipread lists. The effect of a suffix was significant [$F(3,108) = 4.17$, $p < .01$, $MS_e = .08$], as was serial position [$F(6,216) = 45.64$, $p < .001$, $MS_e = .10$]. There were significant interactions among modality of list and suffix [$F(6,216) = 2.15$, $p < .05$, $MS_e = .06$], among suffix and serial position [$F(18,648) = 3.21$, $p < .001$, $MS_e = .03$] and a second-order interaction among modality of list, suffix, and serial position [$F(36,1296) = 2.03$, $p < .001$, $MS_e = .03$].

The second-order interaction was investigated by a 3 (List mode) × 4 (Suffix) ANOVA on recency scores (see Table 5). There were no main effects of suffix [$F(3,108) = 1.23$, $p = .30$, $MS_e = .05$] or modality of the list [$F(2,72) < 1$], and the interaction was only marginally significant [$F(6,216) = 2.02$, $p = .064$, $MS_e = .05$].

Two separate analyses were performed comparing the no-suffix conditions of the present experiment with those of Experiment 3a (/ip, ap, op/ items) and 2b (/pa, ta, fa/ items). The first analysis is concerned with the vowel/consonant contrast, the second looks for the effect of rhyme across different input modalities. With respect to the first, an ANOVA was run on the recency effects of the no-suffix conditions of Experiments 3a (/ip, ap, op/) and 3b (/ap, at, af/). The vocabulary of the items (vowel-varied or consonant-varied items) was between-subjects, and modality of the list was treated as within-subjects variable. There was no main effect of vocabulary [$F(1,94) = 1.30$, $p = 2.58$, $MS_e = .06$], but the effect of modality of the list was significant [$F(2,188) = 4.32$, $p < .02$, $MS_e = .05$]. Post hoc test (Fisher's LSD, $\alpha = .01$, critical difference = .079) indicated that the recency effect of lipread lists was smaller than that of heard-only lists. The interaction between vocabulary and modality of the list did not reach significance [$F(2,1188) = 1.02$, $p = .36$, $MS_e = .05$]. It thus appeared that the vowel/consonant contrast in Experiments 3a and 3b was smaller than that of the previous two experiments. Inspection of the Tables 3 and 5 suggests that the factor rhyme of the consonant-varied items contributed to this difference: the non-rhyming /ap, at, af/ items had larger recency effects than the rhyming /pa, ta, fa/ items. To investigate this statistically, and to see whether the effect of rhyme depends on the modality of the list, an ANOVA was performed on the recency scores of the no-suffix conditions of Experiments 2b (/pa, ta, fa/ items) and 3b (/ap, at, af/) with modality of the list as within-subjects factor. As already suggested, non-rhyming items (/ap, at, af/) had larger recency effects than the rhyming items (/pa, ta, fa/) [$F(1.73) = 6.90$, $p < .01$, $MS_e = .06$]. There was again a main effect of presentation mode of the list [$F(2,146) = 3.38$, $p < .05$, $MS_e = .05$] because lipread lists had smaller recency effects than lists presented in the other modalities. The interaction between vocabulary and modality of the list was not significant [$F(2,146) < 1$]. Thus, the effect of rhyme on recency was equally strong across the modalities in which lists were presented.

TABLE 5
MEAN RECENCY SCORES FOR EACH SUFFIX AND LIST MODE OF /ap, at, af/ LISTS

	List mode			
Suffix	Heard-only	Lipread	Heard-plus-lipread	
Tone	.21	.12	.11	.15
Heard-only	.09	.14	.12	.11
Lipread	.12	.09	.09	.10
Heard-plus-lipread	.02	.14	.14	.10
	.11	.12	.12	

Discussion

The goal of Experiments 3a and 3b was to replicate the vowel advantage with lipread speech and to extend it to items with a different linguistic structure so as to eliminate the possibly confounding factor of inter-item rhyme. As concerns the replication aspect, we note that the present data add new evidence in support of the preliminary conclusions of Experiments 1, 2a, and 2b. The basic finding is that lipread input is sensitive to the consonant/vowel distinction. Above that, we also observed that lipread lists are, like heard-only and heard-plus-lipread input, sensitive to rhyme. Another important aspect which has been replicated several times in the previous experiments is that lipread lists have, in general, smaller recency effects than heard-only or heard-plus-lipread lists. Moreover, we observed that suffix effects in vowel-varied lists depend on the match between modality of the list and suffix: suffix effects are larger if the modality of the list and the suffix match. Finally, these findings can be extended to items with a VC structure, instead of being specific for the more common CV syllable.

General Discussion

The goal of these experiments was to investigate whether the consonant/vowel distinction can be observed in memory performance when the lists to be recalled were lipread, and thus lacking the acoustic component, hitherto credited as the basis of this distinction. The major finding of this study is that there is an important symmetry in memory for consonants versus vowels across the difference in input modality. The vowel advantage is observed across all modalities using various combinations of consonants and vowels. Lists of syllables varying in vowel are better remembered than lists consisting of syllables varying in consonant, whether they are heard, lipread, or heard-plus-lipread. In a similar vein, we observed that items which rhymed had smaller recency effects than corresponding non-rhyming items, but the effect of rhyme was equally strong for each input modality. These results suggest that processing within the speech module ignores peripheral sensory differences and favors aspects of linguistic structure that transcend the differences between input modalities. Taken together, the experiments point to the importance of the consonant/vowel contrast for lipread speech memory and thus fit in well with proposals made in recent studies with heard-only presentation that linguistic, and not auditory factors are critical for performance in immediate memory tasks. However, we also find that there is still an advantage for the heard presentation. The critical issue then is how to combine in one model of processing and immediate memory, symmetries as well as differences between speech input modalities.

Needless to say, relating the advantage of heard vowel presentation to the original notion of an extra acoustic contribution (e.g., Crowder, 1971) is no longer very appealing, since a retreat to this position would ignore the major finding of these experiments, which is the similarity across modalities. More critically still, an acoustic explanation of the vowel advantage would place the origin of the recency effect outside the speech domain. By now this is not a very desirable approach because of the evidence reviewed above together with the present data which suggest that the recency effect is in the phonetic processor.

Do the present data then converge toward supporting the central role of an abstract speech processor in immediate memory? The cross-modal similarities observed here are certainly compatible with the notion that the recency and the suffix effects reflect the role by the speech module. If one adopts that strong perspective on the modularity of speech processing, these results of heard and lipread similarities are certainly predictable. The discovery of a vowel advantage of lipread vowels is compatible with the notion of a phonetic processor extracting abstract information in accordance with phonetic principles. Along these same lines, one is entitled to expect that the sonority dimension will be respected equally with auditory and visual input. In what follows, we discuss how this approach creates problems of its own.

A possible challenge for the explanatory power of an abstract linguistic module comes from a different angle. When items are visually presented and silently mouthed, recency and suffix effects are obtained like with heard presentation (e.g., Nairne & Walters, 1983). One way of accounting for these data is to suppose that not only heard and lipread, but also gestural information is stored in an a-modal abstract speech memory which, in turn, contributes to the recency effect. But, like with heard and lipread presentation, modality-specific differences have been found between heard and mouthed stimuli. In a study by Turner et al. (1987), recency and suffix effects were obtained with auditorily presented items varying in vowels. Like in the present study, these effects were smaller when the items rhymed and varied in consonants. However, Turner et al. (1987) also found that these typical effects were much weaker when the items were visually presented and mouthed. Turner et al. (1987) concluded that recency and suffix effects of heard and mouthed lists are not mediated by the same mechanisms. One can agree with this conclusion, but the question is whether the distinction in memory performance between hearing and mouthing poses a genuine problem for the notion of an abstract speech processor. This depends on one's views about the relation between perception and production. If one admits that abstract representations are gestures, one is led to believe in an intimate link between perception and production. The Turner et al. (1987) results might challenge one or another kind of motor theory of speech perception, but if there is room for different representation formats for perception and production, then mouthing data need not conflict with the present results.

However, at least two aspects of the presents findings are somewhat more difficult to fit in with the notion of an abstract speech processor. One concerns the vowel/consonant contrast itself. Liberman and collaborators (1967) have argued that there are good grounds to postulate a special speech processor which is required more for the perception of some speech sounds than for others. In particular, stop consonants need more of the services of the speech processor than do vowels. However, as recency is supposed to reflect the activity of this speech processor, one might expect that consonants have larger recency effects than vowels. Yet, we observe just the opposite. A second difficulty is that an abstract phonetic processor explanation tends to pay little attention to modality differences in speech processing. If recency reflects these abstract representations, there is not much reason to expect differences due to original modality of presentation. But they are observed here since recency for heard input is larger than for lipread input, and, moreover, suffix effects are modality specific.

An alternative approach already mentioned is based on the notion that the phonetic trace of a lipread speech event is underspecified, whereas for auditory speech it is more fully specified or richer (e.g., Campbell, Garwood, & Rosen, 1988). The notion can be related to the model Nairne (1990) proposes for the explanation of recency effects. Recency results from the processing traces left behind by the final memory item and it reflects the extent to which traces allow to distinguish between the items of the memory set. For instance, the feature place of articulation allows to distinguish between /pa/ and /ta/, but the feature stop-consonantal does not. One might explain the distinction between vowels and consonants by assuming that the vowels /a, i, o/ leave more distinct traces than the consonants /p, t, f/, possibly because vowels are perceputally more distinct from each other. Consequently, consonants are more often confused and recency decreases. This approach can explain the results of the first experiment and the no-suffix condition results from the other ones. That is, vowel-varied lists show more recency than consonant-varied lists because vowels leave more distinct traces than consonants. Moreover, hearing leads to more recency than lipreading because the former leaves more phonetic traces, and, following the additive factor logic, these two factors interact like in Experiment 1, because the difference between hearing versus lipreading and between vowels versus consonants relates to the same mechanism. But obviously, what one would need if this line of thought is to be pursued is an independent measure of the encoded features of heard and lipread vowels and consonants and, moreover, a similarity analysis of these features.

The relation between the previously sketched notion of poor versus rich memory traces and the obtained suffix effects is, however, more complicated. It seems to be a reliable finding that lipread suffixes are less disruptive in heard memory lists (Campbell, Garwood, & Rosen, 1988; de Gelder and Vroomen 1992). An interpretation for this finding could be that the interfering effect of a suffix depends on the amount of phonetic overlap between the final item and the suffix. Since a lipread suffix is phonetically poorer than a heard suffix, the former is less damaging than the latter. However, the same reasoning cannot be applied when a heard-only suffix followed a lipread list. In both experiments, the heard-only suffix did not interfere with lipread recency. Moreover, it was not the case that there was a kind of floor effect in recency of lipread lists, as lipread suffixes did interfere. Similar findings with digit lists were obtained by de Gelder and Vroomen (1992) who also found that a heard-only suffix did not interfere with lipread lists. Moreover, Greene and Crowder (1984, Experiment 3) found a small but significant effect showing that lipread suffixes were more effective than heard-plus-lipread suffixes in lipread list. These findings suggest that the amount of *phonetic* overlap between the final item and the suffix is not the only factor controlling suffix effects. In other words, even by adopting the phonetic trace proposal, one is still forced to consider input modality.

It has been known for some time that the overall similarity between list items and suffix plays a role in the size of the suffix effects (Morton, Crowder & Prussin, 1971; Watkins & Watkins, 1980). The general finding is that suffix effects will be larger the more the suffix resembles the memory list. For instance, a mismatch in voice or a change in location makes a suffix less powerful (Morton, Crowder & Prussin, 1971). The differential suffix effects as obtained in the present study might be an instance of this general principle. On this account, heard suffixes are, in a physical sense, more similar to heard memory lists than to lipread lists and vice versa. This notion can explain the critical finding that lipread suffixes in lipread lists are more disruptive than are heard suffixes, even though the latter consist of richer phonetic traces. Thus, as far as suffix effects are concerned, speech module based explanations need to leave room for such low level effects. Low level effects and modality effects are, however, not incompatible and may well jointly contribute to the magnitude of the suffix effects.

Finally, a more speculative alternative focuses plainly on modality differences by making room for modality specificity inside the speech module. The distinctive aspect of this suggestion is that modality specificity is part of the specific linguistic computations on the signal. It is thus different from what Penney (1989) calls a sensory-based acoustic code which she distinguishes from a phonological code. The notion of an early modality-specific stage in speech processing explains commalities between auditory and lipread presentation, as well as different recency effects for heard and lipread material. The commonalities, most critically here, the consonant/vowel distinction, follow from the assumption that in the two cases the extraction of information conforms to linguistic representations, i.e. minimally, representations defined over linguistic variables and operational in the domain of linguistic input. Only, instead of assuming that extraction is from the earliest stages on driven by abstract linguistic knowledge, we propose that initially it takes place and is driven, in each modality, by a modality-specific subprocessor for speech. Subsequently, abstract processing might take over independently or jointly with retention in phonological memory. The latter might be based on phonological recording as proposed in, for example, Baddeley's (1990) notion of a phonological loop in phonological memory. One can still image in that case that the modality-specific phonetic representations retain their distinctive properties. Against this background, the observed differences can equally well come into their rights. Differential suffix effects can be accounted for if we assume that immediate memory representations have modality-specific and modality-independent components. A suffix will be more disruptive if it overwrites more modality-dependent and modality-independent traces of the final item. This explains the finding that the suffix effect depends on the match between the input modality of the list and the suffix. Such a modality-specific component does not clash with the idea that recency is based on traces of the phonetic processor (Crowder & Surprenant, in press) because it is, in itself, part of the processor.

It is worth noting that there is no logical or theoretical objection against a proposal that modality-specific subprocesses occur within the abstract speech processor. The notion of an abstract processor is ambiguous, since it marks the contrast between sensory and phonetic processes as well as that between modality aspects. Besides, it is an intriguing question whether the abstract representations of the theory builder are in any way real for a cognitive system that only "knows" concrete, modality-specific phonetics. At the empirical level, the phonetic modality-specific view evoked here is compatible with evidence from other studies and paradigms. This notion of heard and lipread speech perception is compatible with a model of late integration, for instance, Massaro's (1987) proposal that the heard and the visual speech signal are combined after they have been evaluated independently of each other. As a matter of fact, the existence of the so-called McGurk illusion (McGurk & MacDonald, 1976) testifies to a significant degree the autonomy of the input modalities even within the speech module. Further suggestions in support of separate subprocessors comes from data suggesting a dissociation between heard and lipread input making subjects insensitive to the McGurk illusion (de Gelder, Vroomen, & van der Heijden, 1990; de Gelder, de Sconen, & Gepner, 1992).

REFERENCES

BADDELEY, A. (1990). *Human memory: theory and practice*. Hillsdale, NJ: Erlbaum.

CAMPBELL, R. (1987). Common processes in immediate memory. In D. A. Allport, D. G. Mackay, W. Prinz, & E. Scheerer (Eds.), *Language perception and production: Common mechanisms in listening, speaking, reading, and writing*. New York: Academic Press.

CAMPBELL, R., & DODD, B. (1980). Hearing by eye. *Quarterly Journal of Experimental Psychology*, 32, 85–99.

CAMPBELL, R., & DODD, B. (1982). Some suffix effects on lipread lists. *Canadian Journal of Psychology*, 36, 509–515.

CAMPBELL, R., & DODD, B. (1984). Aspects of hearing by eye. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X*. Hillsdale, NJ: Erlbaum.

CAMPBELL, R., GARWOOD, J., & ROSEN S. (1988). Adding sound to lipread lists: The effects of serial recall of adding an auditory pulse train and a pure tone to silently lipread lists. *Memory and Cognition*, 16, 210–219.

COLE, R. A. (1973). Different memory functions for consonants and vowels. *Cognitive Psychology*, 4, 39–54.

CORBALLIS, M. C. (1966). Rehearsal and decay in immediate recall of visually and aurally presented items. *Canadian Journal of Psychology,* **20,** 450–461.

Crowder, R. G. (1971). The sound of vowels and consonants in immediate memory. *Journal of Verbal Learning and Verbal Behaviour,* **10,** 587–596.

Crowder, R. G. (1983). The purity of auditory memory. *Philosophical Transactions of the Royal Society of London, B,* **302,** 251–265.

Crowder, R. G., & Morton, J. (1969). Precategorical acoustic storage (PAS). *Perception and Psychophysics,* **5,** 363–373.

Crowder, R. G., & Surprenant, A. M. (in press). On the linguistic module in auditory memory. In B. de Gelder & J. Morais (Eds.), *Speech and reading.* Hillsdale, NJ: Erlbaum.

Darwin, C., & Baddeley, A. (1974). Acoustic memory and the perception of speech. *Cognitive Psychology* **6,** 41–60.

de Gelder, B., & Vroomen, J. (1989). Models in the mind, modules on the lips. *Behavioral and Brain Sciences,* **124,** 762.

de Gelder, B., & Vroomen, J. (1992). Abstract versus modality-specific memory representations. *Memory and Cognition,* **20,** 533–538.

de Gelder, B., & Vroomen, J. (in press). Memory deficits for heard and lipread speech in young and adult dyslexics. In B. de Gelder & J. Morais (Eds.), *Language and literacy.* Hillsdale, NJ: Erlbaum.

de Gelder, B., de Sconen, S., & Gepner, B. (1992, January). *Face processing in autistics: A generalised deficit?* Paper presented at the Xth Cognitive Neuropsychology Workshop, Bressanone.

Dodd, B., & Campbell, R. (1987). *Hearing by eye: the psychology of lip-reading.* London: Erlbaum.

Gardiner, J. M., Gathercole, S. E., & Gregg, V. H. (1983). Further evidence of interference between lipreading and auditory recency. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **9,** 328–333.

Gathercole, S. E. (1987). Lipreading: Implications for theories of short-term memory. In B. Dodd & R. Campbell (Eds.), *Hearing by eye: The psychology of lipreading.* Hillsdale, NJ: Erlbaum.

Greene, R. L., & Crowder, R. G. (1984). Modality and suffix effects in the absence of auditory stimulation. *Journal of Verbal Learning & Verbal Behaviour,* **23,** 371–382.

Jackson, P. L., Montgomery, A. A., & Binnie, C. A. (1976). Perceptual dimensions underlying vowel lipreading performance. *Journal of Speech and Hearing Research,* **19,** 796–812.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review,* **74,** 431–461.

Massaro, D. W. (1987). *Speech perception by ear and eye.* Hillsdale, NJ: Erlbaum.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature* **264,** 746–748.

Morton, J., Crowder, R. G., & Prussin, H. A. (1971). Experiments with the stimulus suffix effect. *Journal of Experimental Psychology: Monographs,* **91,** 169–190.

Nairne, J. S. (1988). A framework for interpreting recency effects in immediate serial recall. *Memory and Cognition,* **16,** 343–352.

Nairne, J. S. (1990). A feature model of immediate memory. *Memory and Cognition,* **18,** 251–269.

Nairne, J. S., & Crowder, R. G. (1982). On the locus of the stimulus suffix effect. *Memory and Cognition,* **10,** 350–357.

Nairne, J. S., & Walters, V. L. (1983). Silent mouthing produces modality- and suffix-like effects. *Journal of Verbal Learning and Verbal Behavior,* **22,** 475–483.

Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory and Cognition,* **17,** 398–422.

Pisoni, D. B. (1973). Auditory and phonetic codes in the discrimination of consonants and vowels. *Perception and Psychophysics,* **13,** 253–260.

Pisoni, D. B. (1975). Auditory short-term memory and visual perception. *Memory and Cognition,* **3,** 7–18.

Selkirk, E. O. (1984). *Phonology and syntax: The relation between sound and structure.* Cambridge: MIT Press.

Spoehr, K. T., & Corin, W. J. (1978). The stimulus suffix effect as a memory coding phenomenon. *Memory and Cognition,* **6,** 583–589.

Surprenant, A. M. & Speer, S. R. (1990, November). *The influence of segmental sonority on immediate memory for syllables.* Paper presented at the 31th meeting of the Psychonomic Society, New Orleans.

Turner, M. L., LaPointe, L. B., Cantor, J., Reeves, C. H., Griffeth, R. H., & Engle, R. W. (1987). Recency and suffix effects found with auditory presentation and with mouthed visual presentation: They're not the same thing. *Journal of Memory and Language,* **26,** 138–164.

Watkins, O. C., & Watkins, M. J. (1980). Echoic memory and voice quality: recency recall is not enhanced by varying presentation voice. *Memory and Cognition,* **8,** 26–30.

Woodward, M. F., & Barber, C. G. (1960). Phoneme perception in lipreading. *Journal of Speech and Hearing Research,* **3,** 212–222.

(Received February 5, 1993)

(Revision received October 10, 1993)