Adaptation is automatic

ARTHUR G. SAMUEL and DONNA KAT
State University of New York, Stony Brook, New York

Two experiments were used to test whether selective adaptation for speech occurs automatically or instead requires attentional resources. A control condition demonstrated the usual large identification shifts caused by repeatedly presenting an adapting sound (/wa/, with listeners identifying members of a /ba/-/wa/ test series). Two types of distractor tasks were used: (1) Subjects did a rapid series of arithmetic problems during the adaptation periods (Experiments 1 and 2), or (2) they made a series of rhyming judgments, requiring phonetic coding (Experiment 2). A control experiment (Experiment 3) demonstrated that these tasks normally impose a heavy attentional cost on phonetic processing. Despite this, for both experimental conditions, the observed adaptation effect was just as large as in the control condition. This result indicates that adaptation is automatic, operating at an early, preattentive level. The implications of these results for current models of speech perception are discussed.

Spoken language can be described as occurring simultaneously on several different levels. For example, the phrase *spoken language* itself can be thought of as a dynamically changing acoustic event, as a series of phonemes or syllables, or as a pair of words. All of these descriptions are accurate. For cognitive psychologists interested in speech perception, a fundamental question is whether such descriptions do in fact refer to representational forms used by human listeners. If we assume that some of them do, the architecture of these different levels becomes a further issue: How, for example, does information at one level affect processing at another level? For at least a decade, there have been two sharply differing answers to the architectural question. The bottom-up/autonomous view is that information is developed successively at the levels suggested above, with no influence of "higher level" representations on lower levels (e.g., Cutler, Mehler, Norris, & Segui, 1987; Massaro, 1989; McQueen, 1991; Norris, 1993). In contrast, top-down/interactive models assert that all (Marslen-Wilson & Welsh, 1978; McClelland, 1991; McClelland & Elman, 1986) or at least some (Connine & Clifton, 1987; Samuel, 1981, 1996) higher level activity can affect the activity at lower levels.

An empirical resolution of this theoretical argument has proved to be impressively elusive. Theorists favoring the autonomous view have generally argued that effects that appear to show top-down perceptual influences can be dismissed as postperceptual, decision-level artifacts. This cannot be said of the top-down lexical effects found in two studies, however, in which the measures used did not de-

pend on explicit identification of the lexically influenced items. In the first of these studies, Elman and McClelland (1988) demonstrated a lexical influence on a phenomenon known as *compensation for coarticulation*. Because such compensation is generally acknowledged to occur at lower levels, the lexical influence qualifies as a top-down effect.

In the second of these studies, Samuel (1997) has reported another case of lexical influence on a lower level process—in this instance, selective adaptation. In a typical adaptation experiment, listeners participate in two tests. First, they identify randomly presented members of a test series varying along a dimension such as place of articulation (e.g., /ba/-/da/), voicing (e.g., /ba/-/pa/), or manner of articulation (e.g., /ba/-/wa/). Then, during an adaptation test, they identify the same test syllables, interspersed in periods of adaptation. These adaptation periods consist of the repeated presentation of an adaptor, which could be a continuum endpoint (e.g., /ba/ or /da/ on /ba/-/da/), a related phoneme (e.g., /pa/ or /ta/ on /ba/-/da/), or even a nonspeech sound with appropriate acoustic properties (e.g., a sawtooth wave on /ba/-/wa/). A large number of studies have shown that adaptation reduces report of the continuum endpoint that shares acoustic-phonetic features with the adaptor (e.g., Eimas & Corbit, 1973; Kat & Samuel, 1984; Samuel, 1986; Sawusch, 1977).

Samuel (1997) coupled the adaptation effect with a second phenomenon, phonemic restoration (Warren, 1970). Phonemic restoration is a perceptual completion effect that occurs when a segment of speech is excised and replaced by another sound (e.g., white noise). Listeners generally report that the speech sounds intact—they perceptually restore the missing speech. Samuel demonstrated that restored phonemes not only sound "real," but are perceptually real enough to produce adaptation. Listeners identified syllables from a /b/-/d/ continuum, before and after adaptation. In the critical conditions, the adaptors consisted of words with either /b/ (e.g., *alpha*et*, *exhi*ition*, etc.) or /d/ (e.g., *arma*illo*, *aca*emic*, etc.) replaced by white noise.

This research was supported by National Institute of Mental Health Grant RO1MH51663-03, and by Air Force Office of Scientific Research Grant 91-0378. We are pleased to acknowledge their support. We thank Rachael Hoina and Laura Kaell for their help in data collection. Correspondence concerning this article should be addressed to A. G. Samuel, Department of Psychology, State University of New York, Stony Brook, NY 11794-2500 (e-mail: arty@jetsam.psy.sunysb.edu).

These restored phonemes caused significant adaptation shifts. In a control condition, the /b/s or /d/s were replaced by silence, rather than by white noise. Such stimuli do not generally produce perceptual restoration of the deleted consonants, and these stimuli did not produce adaptation effects. This result indicates that the adaptation for the restored case was occurring at relatively low levels (e.g., phonemic), because the control (e.g., *arma_illo*) and experimental (e.g., *arma*illo*) stimuli provided similar high-level (lexical) information.

Samuel's (1997) results provide strong support for the view that activation of higher level representations can affect the activation of lower level ones. Moreover, this top-down effect cannot be attributed to a postperceptual decision stage, because the listeners never had to decide what the words were—the required judgments were on the members of the /b1/-/d1/ test series. As long as the adaptation effects occurred at lower levels (e.g., phonemic) than the information determining the adaptor identity (lexical), the results require a top-down architecture.

In the present study, we examined whether or not adaptation is indeed a low-level, basic process. The control conditions in Samuel's (1997) study indicated that adaptation does not occur at the lexical level. In the present study, we tested whether adaptation requires any cognitive/linguistic resources at all, or whether it is instead a low-level automatic process. Our working hypothesis (based on prior research) was that adaptation is a basic property of early perceptual processing and does not involve higher level processing. In this sense it is preattentive, much like the feature extraction properties assumed in a number of attentional models (e.g., Feature Integration, Treisman & Gelade, 1980; Guided Search, Wolfe, Cave, & Franzel, 1989). If adaptation does prove to be such a low-level process, the success of adaptation by lexically mediated phonemes must be attributed to a true top-down perceptual process.

Of course, understanding the nature and principles behind a well-documented effect such as adaptation is important in its own right, beyond its relevance for the interpretation of Samuel's (1997) results. As such, it should not be surprising that there have been earlier efforts to determine how automatic adaptation is. Mullennix (1986) tested whether adaptation would be reduced when listeners were required to listen for certain target syllables that were presented during the adaptation phases of his experiments. Various combinations of male voice versus female voice of the target syllables were used, as well as of same versus different ear for the adaptor and test syllables. Mullennix found no reduction in adaptor efficacy as a function of distractor task, suggesting a stimulus-based, automatic basis for adaptation. However, this interpretation must be very tentative, because the auditorily presented targets and distractors could themselves produce unknown adaptation effects. In addition, it is not clear that a monitoring task of this sort would continuously require attention. If it did not, listeners could "time-share" between the monitoring and adaptation stimuli.

The one other study of attentional effects in adaptation also was suggestive rather than conclusive. Sussman (1993)

used a /ba/-/da/ test series, with /da/ adaptation, with three relevant attentional conditions. A baseline case (with no distraction) was compared with a visual distraction condition and with an auditory one. In the visual case, the words SEE and SHE were presented on a screen, in a random series, during (auditory) presentation of the adaptor. The task was to count the number of SHE occurrences. This task was apparently very easy; performance was 99% accurate. Not surprisingly, doing this simple task had no effect on the size of the observed adaptation. In the auditory distractor task, the *see* and *she* syllables were presented simultaneously and contralaterally to the (monaural) adaptor, at very low intensity. Listeners were periodically probed on the identity of the most recently presented *see* or *she*. Here, performance was slightly better than chance (59%). No decrease in adaptor efficacy was found. In fact, there was a trend toward *greater* shifts than with no distractor task. This trend could have been caused by the *see* and *she* syllables themselves, because they share place of articulation with the adaptor (/da/). This possibility, and the near-chance performance levels, once again leaves the status of attentional effects on adaptation unclear. The existing literature provides no evidence for an attentional cost to adaptation, but the question remains open.

In the present study, we sought to answer this question more definitively. In two experiments, we used distractor tasks during adaptation that shared two important properties: First, the distractor tasks were all visually presented, preempting any possible unintended adaptation by the distractors. Second, the tasks were carefully calibrated to be in an optimal range of difficulty, far from both the ceiling and floor levels that can complicate interpretation. In a third experiment, we measured the attentional demands of the distractor tasks explicitly and found that they did indeed consume substantial attentional resources. If adaptation effects are unaffected by such tasks, we can more confidently conclude that adaptation is an automatic, low-level, preattentive process.

In Experiment 1, the distracting task was a sequence of arithmetic problems presented on a screen. The task was to decide whether the answer given was correct (e.g., " $9 + 7 = 16$ " is correct, but " $9 + 7 = 15$ " is not). In Experiment 1, we also used a set of stimuli and conditions optimized for separately looking for central versus peripheral adaptation effects. If attention plays any role in adaptation, one would expect it to be found for the central component. In Experiment 2, we used a more demanding variation of the arithmetic task and a visually presented rhyming task. The latter provided a test for any involvement of domain-specific (i.e., linguistic) resources during adaptation. Collectively, the conditions of the first two experiments constituted a sensitive probe of the extent to which adaptation processes require cognitive or linguistic resources.

EXPERIMENT 1

In Experiment 1, we examined whether doing a rapid series of arithmetic problems during the periods of adaptation would cause a reduction in the observed adaptation

effect. The arithmetic problems were presented at a rate of one every 2 sec, a task that pilot work showed to be extremely attentionally demanding. If adaptation requires cognitive resources, the arithmetic task should reduce the observed shifts. Both adaptors and test syllables were presented monaurally. By varying the ear of adaptor presentation (ipsilateral vs. contralateral to the test syllables), any central components of adaptation could be examined separately from peripheral ones. The contralateral test isolated central processes from more peripheral ones. Experiment 1 provided a very sensitive test of this issue because the particular test items and adaptor have been shown to yield effects at both central and peripheral levels (Samuel, 1988).

Method

Subjects

Eleven native American English speakers, with normal hearing, participated. They received course credit or were paid for their participation. One additional subject failed to identify the syllables consistently and was not included in the analyses.

Stimuli

The test series. An eight-item synthetic /ba/-/wa/ continuum was used as the test series. The eight tokens were generated on the Klatt (1980) software synthesizer. The stimuli varied in the duration and rate of change of the first three formant transitions: The /ba/ endpoint had short, rapid transitions (20 msec); the /wa/ endpoint had longer, more gradual changes (55 msec). The transition duration changed in 5-msec increments across the continuum. The stimuli were identical to the voiced /ba/-/wa/ syllables used by Samuel (1988); a full description of the stimulus parameters is available there.

The adaptor. The endpoint /wa/ served as the adaptor. This adaptor was chosen because it produced large adaptation effects in previous research (Samuel, 1988). Moreover, the previous research indicated that approximately half of the large adaptation effect was attributable to central, rather than peripheral, processes. If cognitive resources are a factor in adaptation, it is likely that their influence would be most apparent under adaptation conditions mediated by central processes. The use of /wa/ as an adaptor, under the conditions described below, permitted a sensitive test of any possible attentional influence.

Procedure and Apparatus

Each individual participated in two sessions held at least 24 h apart. Each session included a baseline identification test, followed by an adaptation test (with /wa/ as the adaptor). Half of the subjects engaged in the distractor task during the adaptation periods of their first session; half did so during their second session. Within each of these two groups, half of the subjects wore their headphones in the normal orientation, and half wore them reversed (left channel to right ear).

Each baseline test included 42 randomizations of the eight-item /ba/-/wa/ continuum. The first six randomizations were practice and were not scored. Each adaptation test began with 90 presentations (60 sec) of /wa/, followed by 36 adaptation blocks. Each block included 30 sec of adaptation (45 iterations), followed by one randomization of the eight tokens for identification. Both speed and accuracy of responding were stressed. Subjects responded by pushing buttons labeled "B" and "W," which were the middle two buttons on a four-button response panel.

Adaptation was monaural—the adaptor was always output over the same D/A channel. Therefore, because of the headphone orientation counterbalancing, half of the listeners received the adaptor over the left headphone, and half received it over the right one.

Test syllables were also presented monaurally, during both the identification tests and the adaptation tests. The procedure was to

present one randomization of the eight test items over one channel, and to present the next randomization over the other channel. Note that for the adaptation tests, this procedure resulted in an alternation of ear of presentation on an adaptation block by adaptation block basis: One block of eight test items would be heard in the same ear as was the adapting sound; the next would be in the opposite ear. Thus, during each session, 18 baseline and 18 adapted observations were collected for each token, in each ear.

As noted above, during one of each listener's two sessions, the adaptation periods were accompanied by the arithmetic distractor task. The task consisted of a series of addition problems presented on an ADDs Viewpoint 3A+ terminal screen. Each problem was of the form $X + Y = \text{ANS}$, with X and Y each being a random integer from 0 to 15. The ANS value was determined probabilistically: In each case, there was a .5 probability of ANS's being the correct sum, there was a .25 probability of its being 1 or 2 more than the correct sum, and there was a .25 probability of its being 1 or 2 less than the correct sum. The subjects were instructed to evaluate each equation and to keep a running count of the number of *incorrect* equations. Each problem was on the screen for 2 sec. At the end of each adaptation sequence, the phrase "ERRORS = ____?" was presented, with the blank containing the actual number of incorrect equations during the block ($p = .5$), one more than the actual number ($p = .25$), or one less ($p = .25$). The subjects responded by pushing one of the two outside buttons on their four-button response boards; these outside buttons were labeled "YES" and "NO." After all the subjects had entered a response, the eight-item identification block was presented. The distractor task was quite demanding and, subjectively, seemed to require constant attention.

Results and Discussion

The distractor task yielded performance in the desired range: 74% correct responses. This level of accuracy was far from both the ceiling (100%) and floor (50%), indicating that the task was challenging enough to require substantial attention, but not hard enough to cause listeners to give up.

The responses to each of the eight stimuli were averaged and analyzed separately for each ear of presentation for each listener. Ear of presentation on both the baseline and adaptation test was defined as either *ipsilateral* to the adaptor or *contralateral* to it. The resulting identification/adaptation functions are shown in Figure 1, broken down by whether or not the arithmetic task was done during the adaptation phases.

The average percentage of "B" responses for the middle four members of the /ba/-/wa/ continuum was used as the primary summary statistic. This measure has proved to be very sensitive, because it focuses on the region of most robust and reliable adaptation (see Samuel & Newport, 1979, for additional advantages). These scores were entered into an analysis of variance with the following three factors: distractor task (distractor vs. no distractor), ear (ipsilateral vs. contralateral), and adaptation (baseline identification vs. adaptation). As Figure 1 makes clear, there was a robust effect of adaptation [baseline, 65.5% "W"; adapt, 25.8% "W", which is a 39.7% decrease in "W" report; $F(1,10) = 107.74, p < .001$]. There was also a strong effect of ear of presentation [$F(1,10) = 19.10, p < .001$]. This main effect was really just due to differences after adaptation, of course (since "ipsilateral" and "contralateral" are not relevant on the baseline), resulting in a reliable ear $\times$ adaptation interaction [$F(1,10) = 19.31, p < .001$]. These

Figure 1. Identification of the /ba/-/wa/ test syllables, before and after adaptation in Experiment 1. The two left panels are from the contralateral testing conditions; the right panels show the ipsilateral cases. Data in the top two panels come from the experimental (arithmetic) condition. The no-distractor control case is shown at the bottom.

effects replicate Samuel's (1988) results with these stimuli, with large adaptation effects overall, and substantial effects even with test stimuli presented contralaterally to the adaptor. These contralateral effects averaged 31.6%, indicating a large central adaptation effect.

The central question of Experiment 1 involved the third factor, distractor task. No main effect was predicted, because the main effect included both baseline and adaptation data, and no main effect was found [$F(1,10) < 1$, n.s.]. The statistic of interest was the interaction of distractor task with adaptation: Did the arithmetic task reduce the difference between the adaptation condition and its baseline? As Figure 1 shows, the distractor task did not have this effect. The 41.8% shift without distraction was indistinguishable from the 37.4% effect found with arithmetic [$F(1,10) < 1$, n.s.]. It was suggested above that the attentional effect might be concentrated in adaptation processes that operate at a central level—that is, ones tapped by the contralateral condition. This pattern would result in a three-way interaction of distractor task, adaptation, and ear of presentation, an interaction that did not approach significance [$F(1,10) = 0.72$, n.s.].

The results of Experiment 1 are thus quite clear: Doing an attentionally demanding series of arithmetic problems during the adaptor presentation does not diminish the adaptor's effect. This is just as true for the centrally based part of /wa/'s adaptation effect as for its more peripheral aspects. These results provide support for the view that

adaptation effects operate at an early, preattentive level of processing.

EXPERIMENT 2

In Experiment 2, we included two further tests of the automaticity of adaptation. In one condition, we used a variation of Experiment 1's arithmetic task. This new version allowed less chance for adaptor processing to "slip in" during relatively easy parts of the distractor task. In the second condition of Experiment 2, we used a phonetically based rhyming distractor task. This condition provided a more domain-specific test of adaptation: Would adaptation be reduced if the listener was busy doing a phonetic task during the adaptation?

Method

Subjects

Eighteen native English speakers, with normal hearing, participated. They received course credit or were paid for their participation.

Stimuli

The same test series (/ba/-/wa/) and adaptor (/wa/) were used as in Experiment 1.

Procedure and Apparatus

Except as noted, the procedure and apparatus were as in Experiment 1. Each subject took part in three sessions, held at least 24 h apart, with the order of the sessions counterbalanced across subjects.

Figure 2. Identification of the /ba/-/wa/ test syllables, before and after adaptation, in Experiment 2. The left panel shows the no-distractor control case, the middle shows the math condition, and the right panel shows the results from the rhyme distractor task.

Each session included a baseline identification test followed by an adaptation test. There were 30 passes on the identification test (including 6 practice), and 24 adaptation passes. Because Experiment 1 showed that there was no differential attentional effect for centrally based adaptation, all adaptors and test syllables were presented binaurally, rather than monaurally. The three sessions varied only in the nature of the distractor task during adaptation. These three situations will be referred to as the *control*, *math*, and *rhyme* conditions.

Control. The control condition was just that—no distracting task was used. This was comparable to the no-distractor condition of Experiment 1.

Math. The math condition was similar to the distractor condition of Experiment 1. However, in this case, the subjects were engaged in a more *continuous* version of the arithmetic task. In Experiment 1, a fixed interval of 2 sec was given for each problem, regardless of difficulty. It is possible that on relatively easy problems, the subject could solve the problem in well under 2 sec, leaving some time in which attention could be available to process the adapting sound. In Experiment 2, button-board responses to each problem (“correct” or “incorrect”) were required immediately. As soon as the subject responded to one problem, a new problem was put up on the screen. The new problem was displayed immediately before the next adaptor iteration. This continuous version of the arithmetic task was thus individually paced, with subjects urged to respond very rapidly to each math problem.

Rhyme. The rhyme condition followed the same individually paced procedure as did the math. However, rather than an arithmetic problem, each display contained three words presented in a row. Each display either had three words that rhymed (e.g., *such touch hutch*), or two rhyming and one nonrhyming word (e.g., *bear chair near*). Mismatching orthography was used (e.g., *bear chair*, rather than *bare fare*) to force judgments to be phonetically determined. When a nonrhyming word occurred, its position was evenly distributed (left, middle, or right). As in the math condition, the subjects were strongly encouraged to respond rapidly to the rhyme stimuli. Half of the displays called for a “yes” response, and half required “no”; appropriately labeled buttons were present on the response boards. Approximately 300 word sets (in both positive and negative versions) were constructed. Sets were randomly presented without replacement. If an individual responded quickly enough to exhaust the pool, some items could appear a second time. On the average, 70 word sets were used a second time. Given the large number of word sets, the use of both positive and negative versions of each set, and the difficulty of the task, it is unlikely that repetition of a word set would have been noticed.

Results and Discussion

As in Experiment 1, the distractor tasks yielded performance in the desired range. An average of 15.9 math problems were done in each 30-sec adaptation phase, with an average accuracy of 82%. The rhyming task was slightly harder, with 16.1 problems per pass done with 70% accuracy. For both tasks, there was no correlation between the number of problems correctly completed and the observed adaptation effect (math, $r = .08$, $p > .5$; rhyme, $r = .00$, $p > .5$). These results indicate that the two distractor tasks operated as desired, requiring substantial attention.

Figure 2 shows the identification and adaptation results for the control, math, and rhyme conditions. As in Experiment 1, the average identification of the middle four members of the /ba/-/wa/ series was used in the analyses. A two-factor analysis of variance (identification vs. adaptation $\times$ control/math/rhyme) confirmed what is apparent in Figure 2. There were large adaptation effects [$F(1,17) = 200.0$, $p < .001$], but there was no effect of either the math or the rhyme distractor task. This noneffect was observed both for the main effect of task [$F(2,34) < 1$] and in the more meaningful interaction of the two factors [$F(2,34) < 1$]. Indeed, as Figure 2 illustrates, the observed shifts (baseline vs. adaptation %W) were impressively similar, with average effects of 33% (control), 36% (math), and 36% (rhyme). Clearly, the adaptation process was not affected by either continuous arithmetic, or by making a rapid series of phonetic judgments.

EXPERIMENT 3

The results of the first two experiments are extremely clear: Neither the math nor the rhyme tasks interfered with the occurrence of adaptation. We noted previously that this result is particularly telling because these tasks are subjectively very demanding. Experiment 3 was designed to provide a more objective estimate of just how demanding these tasks are: If the math and rhyme tasks demonstrably consume high levels of subjects’ resources, the immunity

of adaptation to these tasks must be due to its not requiring such resources (i.e., adaptation is automatic).

The approach taken in Experiment 3 was to choose a task that involved the encoding of consonant–vowel (CV) information, just as the adaptation phase consisted of a succession of CV presentations. However, the task in Experiment 3 was intended to be one that clearly required resources—in particular, those needed for processing phonetic information. On each trial, a series of five CV syllables (e.g., *bah, rah, nah, tah, gah*) was presented, followed by a “target” syllable. The task was to respond “yes” if the target had been in the set of five, and “no” if it had not.

Experiment 3 included four conditions. The baseline case involved only the task just described. In three other conditions, the subjects simultaneously performed a second task as well. In the rhyme condition, the second task was the rhyme judgment task used in Experiment 2. If the rhyme task does indeed consume substantial phonetic processing resources, we should observe a substantial decrement in performance on the CV recognition task. The same logic applied to the remaining two conditions: The math task from Experiment 2 was one of these, and it tested for more general cognitive resource use. The final condition (see/she) was a version of the task used by Sussman (1993). The results in this condition would allow us to estimate the attentional demands imposed by that task.

Method

Subjects

Twelve individuals from the same population as those in Experiment 2 participated. They received course credit.

Stimuli

Fourteen CV syllables spoken by the first author were digitally recorded (4.8 kHz, low-pass filtered; 12-bit A/D at 10 kHz). The vowel was always /a/, preceded by one of the following consonants: /bdgptkmnfsčwrl/. The duration of the syllables ranged between 292 and 470 msec.

Procedure and Apparatus

Each subject completed four tasks, with order of the four counterbalanced across subjects. The four tasks will be designated by the nature of the load task: *none*, *rhyme*, *math*, and *see/she*.

None. In the baseline condition, the subjects performed only a syllable recognition task; there was no load task. The syllable recognition task consisted of 5 of the 14 CVs being randomly selected and presented over headphones at a rate of one per second (onset-to-onset). Three seconds after the last syllable was presented, a target syllable was played. The word “Present?” was displayed on the screen at this time. The subjects pressed one of two labeled buttons (“yes,” or “no”) to indicate whether the target had been in the list of five. The target was in fact from the list half of the time; it was randomly chosen from the first four list items on these trials. Otherwise, it was a randomly selected syllable from the remaining nine CVs.

Rhyme. In the rhyme condition, the rhyme task from Experiment 2 was added to the syllable recognition task. Each trial consisted of a series of rhyme problems (presented visually, as before), with the syllable recognition task concurrently presented auditorily. The rhyme problems were self-paced, with both speed and accuracy heavily emphasized as in Experiment 2. Buttons labeled “right” and “wrong” were used for the rhyme problem. The first auditory syllable

was presented at the onset of the second rhyme problem, and one rhyme problem followed the presentation of the last syllable. Thus, the 5 sec of auditory syllable presentation were preceded and followed by one rhyme problem; rhyme problems were of course also maintained during the syllable presentation. As soon as the final rhyme response was made, “Present?” was displayed on the screen, and the target syllable was played for recognition.

Math. The procedure for the math condition was completely analogous to that for the rhyme, except that the math task of Experiment 2 was used. As with rhyme, visual problems preceded, accompanied, and followed the syllables. Again, the “right” and “wrong” buttons were used for the math problems, with the “yes” and “no” buttons used for the syllable recognition task.

See/she. The visual presentation in this condition consisted of the words *she* and *see*. The task was to count the number of *she* occurrences (Sussman, 1993). As in the math and rhyme conditions, one problem preceded, and one problem followed, the five-syllable list. Thus, the trial started with either *see* or *she* being displayed. The following five *see/she* presentations were timed to coincide with the five syllables (i.e., one per second), and a final *see/she* was presented after the final syllable. Unlike in the rhyme and math conditions (but as in Sussman, 1993), subjects did not respond to the visual items on line. Rather, they were told to count the number of *she* occurrences. After the final *see/she* presentation, a digit was presented (e.g., “4?”). If the digit matched the subject’s count, the “right” button was pushed; if not, the “wrong” button. The digit offered to the subject was correct half the time, and wrong by 1 or 2 the rest of the time. As in rhyme and math, the visual judgment was immediately followed by the “Present?” cue and target syllable, for a “yes/no” reply. In all three load tasks, the time between the final list syllable and the target syllable was 2–3 sec, as in the none baseline. In all four conditions, there were 28 trials, of which the first 4 were practice.

The apparatus was the same as that in Experiment 2.

Results and Discussion

The basic question of Experiment 3 is the degree to which each load task impairs the phonetic coding of the CV syllables. The left side of Figure 3 presents the average recognition accuracy in the four conditions, with a chance level of 50%. Consistent with our subjective impressions, the math and rhyme tasks caused very large performance decrements from the baseline level (83%). Both tasks yielded reliably reduced performance [rhyme (59%) vs. none, $F(1,11) = 20.92$, $p < .001$; math (63%) vs. none, $F(1,11) = 30.00$, $p < .001$]. The *see/she* task also caused a reliable drop in syllable recognition [71%; $F(1,11) = 9.14$, $p < .02$]. However, its effect was reliably smaller than that of both the math [$F(1,11) = 6.75$, $p < .03$] and the rhyme [$F(1,11) = 7.24$, $p < .03$].

The right side of Figure 3 provides some rough estimates of the attentional consequences of the three load tasks, compared with the no-load case. First, each task was scaled relative to the baseline, treating each as a level above the chance rate of 50%. For example, the 71% accuracy for *see/she* is 21% above chance, versus 33% over chance for none; their ratio is .64. The corresponding ratios for rhyme (9% over chance) and for math (13%) were .27 and .39, respectively. These ratios reflect the relative performances on the syllable recognition task. These values were adjusted by the observed accuracy on the load tasks themselves:

Figure 3. Performance in the four conditions of Experiment 3. The left panel shows the recognition accuracy on the syllable recognition task; the right panel presents an adjusted measure of attentional availability.

Each load task involved trivially simple judgments individually (e.g., everyone should correctly solve " $7 + 6 = 13$ "); any errors must be due to a failure to allocate sufficient resources to the task. Put another way, a subject could have chosen to work *only* on the syllable recognition, at the expense of accuracy on the load task. By multiplying the ratios above by the observed accuracy for see/she (94%), rhyme (76%), and math (87%), we get an estimate of the attentional resources available when both tasks are undertaken. This estimate is not extremely precise, of course, but it should be a reasonably good assessment of whether a task is highly demanding, moderately demanding, or not very demanding. The estimate for rhyme (86% resource consumption) certainly suggests that any task that is resource dependent should show major effects of this concurrent task.

Thus, Experiment 3 demonstrates that both of the load tasks used in the present study have a very substantial impact on the ability to briefly encode phonetic information. The complete immunity of adaptation to both the math and rhyme tasks indicates that adaptation does not require any significant amount of processing capacity; adaptation is automatic.

GENERAL DISCUSSION

The present study included three adaptation tests, with one recurring result: Adaptation is automatic. In Experiment 1, contralateral presentation of the adaptor and test syllables isolated the central component of adaptation, but even this component showed no effect of a simultaneous distractor task. In Experiment 2, a more continuous version of the math task was developed, to minimize the chance that any "stray" adaptor presentations might receive attention. The observed adaptation effect remained unaf-

fected. Even when a distractor task was both continuous and phonetically mediated, the observed adaptation was indistinguishable from shifts in the no-distractor control case. The rhyming task clearly engaged phonetic processing very heavily, indicating that the adaptation process is truly preattentive. Experiment 3 provided an explicit measurement of just how demanding the rhyme task (and the math task) is, reinforcing this conclusion.

The results of the present study have interesting implications in at least three areas. A now substantial literature supports the view that there are three qualitatively different early levels of processing of speech (e.g., Samuel & Kat, 1996; Sawusch, 1977, 1986). The first level is peripheral, being monaurally driven. This level appears to code relatively simple acoustic properties. The second level is a central, binaurally driven set of representations. This level codes more complex acoustic properties, including perhaps certain combinations of formant patterns. The third level is also centrally represented, but appears to code more linguistic, possibly phonetic, information. These three adaptation-identified levels correspond nicely to levels adduced through other procedures, such as categorical perception tests (e.g., Fujisaki & Kawashima, 1970) and studies of binaural fusion (e.g., Cutting, 1976). Recent work (e.g., Samuel & Kat, 1996) has indicated that the three different levels might have different mechanisms underlying adaptation shifts. For example, there is evidence that shifts at the two central levels behave differently: Level 2 shifts typically are accompanied by corresponding reaction time changes (adapted items are not only reported less, they are also identified more slowly), whereas Level 3 shifts do not show any reaction time changes. The results of the present study indicate that despite such underlying differences, the shifts at *all* levels are preattentive. If any level's adaptation effects required attentional resources, we would

have observed an effect in Experiment 1 or 2. Thus, adaptation operates automatically, across whatever differences there might be in its instantiation across levels.

The present results are consistent with prior attempts to assess any attentional role in adaptation. Recall that there have been two such attempts (Mullennix, 1986; Sussman, 1993). In each of these studies, methodological issues made the conclusions only tentative. Nonetheless, each reached the same conclusion as we have in the present study: Adaptation operates at an early, preattentive level. The three tests in the present study thus converge with the two prior examinations.

This conclusion has important implications for the ongoing debate over whether or not lexical activation can affect lower level, phonetic processing. This dispute pits those who reject such top-down influences (e.g., Cutler et al., 1987; Eimas, Hornstein, & Payton, 1990; Massaro, 1989; McQueen, 1991; Norris, 1993) against those who believe that they exist (e.g., Connine & Clifton, 1987; Marslen-Wilson & Welsh, 1978; McClelland, 1991; McClelland & Elman, 1986; Samuel, 1981, 1996). Samuel's (1997) recent demonstration that perceptually restored phonemes can produce significant adaptation effects provides some of the strongest evidence available in favor of top-down lexical influences. As described previously, this effect is not subject to an account at a postperceptual decision stage. The data from the present study provide important support for the top-down explanation of Samuel's results. That study included evidence that adaptation does not occur at the lexical level itself; it is limited to lower level representations. The present study takes this one step further: The results here indicate that adaptation is an automatic, low-level, preattentive process. Because the critical adapting phonemes in Samuel's study were *lexically* determined (e.g., /b/ was restored in *alpha*et*, whereas /d/ was heard in *arma*illo*), a sublexical locus for adaptation necessarily implies a top-down effect: Lexical activation (coupled with an appropriate masking noise) *creates* a true, perceived phoneme, with only this lower level representation being able to be involved in adaptation.

The results of the present study thus solidify the empirical evidence for a perceptual architecture with top-down, lexical to sublexical information flow. In addition, they clearly indicate that the mechanisms underlying adaptation operate at early, preattentive levels: Adaptation is automatic.

REFERENCES

- CONNINE, C. M., & CLIFTON, C. (1987). Interactive use of lexical information in speech perception. *Journal of Experimental Psychology: Human Perception & Performance*, **13**, 291-299.
- CUTLER, A., MEHLER, J., NORRIS, D., & SEGUL, J. (1987). Phoneme identification and the lexicon. *Cognitive Psychology*, **19**, 141-177.
- CUTTING, J. E. (1976). Auditory and linguistic processes in speech perception: Inferences from six fusions in dichotic listening. *Psychological Review*, **83**, 114-140.
- EIMAS, P. D., & CORBIT, J. D. (1973). Selective adaptation of linguistic feature detectors. *Cognitive Psychology*, **4**, 99-109.
- EIMAS, P. D., HORNSTEIN, S. M., & PAYTON, P. (1990). Attention and the role of dual codes in phoneme monitoring. *Journal of Memory & Language*, **29**, 160-180.
- ELMAN, J. L., & MCCLELLAND, J. L. (1988). Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. *Journal of Memory & Language*, **27**, 143-165.
- FUJISAKI, H., & KAWASHIMA, T. (1970). Some experiments on speech perception and a model for the perceptual mechanism. *Annual Report of the Engineering Research Institute*, **29**, 207-214.
- KAT, D., & SAMUEL, A. G. (1984). More adaptation of speech by non-speech. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 512-525.
- KLATT, D. H. (1980). Software for a cascade-parallel formant synthesizer. *Journal of the Acoustical Society of America*, **67**, 971-995.
- MARSLÉN-WILSON, W. D., & WELSH, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, **10**, 29-63.
- MASSARO, D. W. (1989). Testing between the TRACE model and the fuzzy logical model of speech perception. *Cognitive Psychology*, **21**, 398-421.
- MCCLELLAND, J. L. (1991). Stochastic interactive processes and the effect of context on perception. *Cognitive Psychology*, **23**, 1-44.
- MCCLELLAND, J. L., & ELMAN, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, **18**, 1-86.
- MCQUEEN, J. M. (1991). The influence of the lexicon on phonetic categorization: Stimulus quality in word-final ambiguity. *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 433-443.
- MULLENNIX, J. W. (1986). *Attentional limitations in the perception of speech*. Unpublished doctoral dissertation, State University of New York, Buffalo.
- NORRIS, D. G. (1993). Bottom-up connectionist models of "interaction." In G. Altmann & R. Shillcock (Eds.), *Cognitive models of speech processing: The second Sperlonga meeting* (pp. 211-234). Hillsdale, NJ: Erlbaum.
- SAMUEL, A. G. (1981). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General*, **110**, 474-494.
- SAMUEL, A. G. (1986). Red herring detectors and speech perception: In defense of selective adaptation. *Cognitive Psychology*, **18**, 452-499.
- SAMUEL, A. G. (1988). Central and peripheral representation of whispered and voiced speech. *Journal of Experimental Psychology: Human Perception & Performance*, **14**, 379-388.
- SAMUEL, A. G. (1996). Does lexical information influence the perceptual restoration of phonemes? *Journal of Experimental Psychology: General*, **125**, 28-51.
- SAMUEL, A. G. (1997). Lexical activation produces potent phonemic percepts. *Cognitive Psychology*, **32**, 97-127.
- SAMUEL, A. G., & KAT, D. (1996). Early levels of analysis of speech. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 676-694.
- SAMUEL, A. G., & NEWPORT, E. L. (1979). Adaptation of speech and non-speech: Evidence for complex acoustic cue detectors. *Journal of Experimental Psychology: Human Perception & Performance*, **5**, 563-578.
- SAWUSCH, J. (1977). Peripheral and central processes in selective adaptation of place of articulation in stop consonants. *Journal of the Acoustical Society of America*, **62**, 738-750.
- SAWUSCH, J. (1986). Auditory and phonetic coding of speech. In E. Schwab & H. Nusbaum (Eds.), *Pattern recognition by humans and machines* (pp. 51-88). Orlando, FL: Academic Press.
- SUSSMAN, J. E. (1993). Focused attention during selective adaptation along a place of articulation continuum. *Journal of the Acoustical Society of America*, **93**, 488-498.
- TREISMAN, A. M., & GELADE, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, **12**, 97-136.
- WARREN, R. M. (1970). Perceptual restoration of missing speech sounds. *Science*, **167**, 392-393.
- WOLFE, J. M., CAVE, K. R., & FRANZEL, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 419-433.