Research Article

Temporally Nonadjacent Nonlinguistic Sounds Affect Speech Categorization

Lori L. Holt

Carnegie Mellon University

ABSTRACT—*Speech perception is an ecologically important example of the highly context-dependent nature of perception; adjacent speech, and even nonspeech, sounds influence how listeners categorize speech. Some theories emphasize linguistic or articulation-based processes in speech-elicited context effects and peripheral (cochlear) auditory perceptual interactions in non-speech-elicited context effects. The present studies challenge this division. Results of three experiments indicate that acoustic histories composed of sine-wave tones drawn from spectral distributions with different mean frequencies robustly affect speech categorization. These context effects were observed even when the acoustic context temporally adjacent to the speech stimulus was held constant and when more than a second of silence or multiple intervening sounds separated the nonlinguistic acoustic context and speech targets. These experiments indicate that speech categorization is sensitive to statistical distributions of spectral information, even if the distributions are composed of nonlinguistic elements. Acoustic context need be neither linguistic nor local to influence speech perception.*

Understanding context dependence is a central challenge in theories of perception because perception of identical physical signals varies considerably with changing context. Speech perception is an excellent example of the profound influence context has on perception of complex physical signals. Context effects whereby a preceding speech sound influences categorization of a subsequent speech target have been widely observed (e.g., Lindblom & Studdert-Kennedy, 1967; Mann, 1980; Mann & Repp, 1980, 1981). Interestingly, preceding nonlinguistic acoustic stimuli can also shift categorization of following speech targets (Coady, Kluender, & Rhode, 2003; Holt & Lotto, 2002; Holt, Lotto, & Kluender, 2000; Lotto, Sullivan, & Holt, 2003; Lotto & Kluender, 1998). For both effects of context produced by adjacent speech stimuli and effects elicited by adjacent nonlinguistic sounds, the influence of context is spectrally contrastive (Holt & Lotto, 2002; Holt et al., 2000; Lotto & Kluender, 1998; Lotto, Kluender, & Holt, 1997); when the preceding context possesses relatively high-frequency acoustic energy, listeners tend to categorize subsequent speech sounds as possessing lower-frequency energy.

A classic example is the influence of preceding /al/ or /ar/ on categorization of perceptually ambiguous speech targets as either /ga/ or /da/. Listeners are more likely to categorize speech targets as “ga” (the alternative with greater low-frequency energy) when they are preceded by /al/ (with greater high-frequency energy). They label the same speech targets as “da” (with more high-frequency energy) when they are preceded by /ar/ (with more low-frequency energy; Mann, 1980). High- and low-frequency nonlinguistic sine-wave tones that mimic minimal spectral characteristics of /al/ and /ar/, respectively, produce the same spectrally contrastive influence on speech categorization (Holt & Lotto, 2002; Lotto et al., 2003; Lotto & Kluender, 1998). To date, spectrally contrastive context effects have been reported for the influence of speech contexts on speech targets (e.g., Lindblom & Studdert-Kennedy, 1967; Mann, 1980; Mann & Repp, 1980, 1981), nonlinguistic acoustic contexts on speech targets (Coady et al., 2003; Holt & Lotto, 2002; Holt et al., 2000; Lotto et al., 2003; Lotto & Kluender, 1998), and speech contexts on nonlinguistic acoustic targets (Stephens & Holt, 2003). Observations of context effects across different classes of sounds are indicative of general perceptual interactions in processing acoustic energy of adjacent sounds, whether linguistic or nonlinguistic.

Although previous experiments have demonstrated that nonlinguistic precursors can influence speech perception (e.g.,

Address correspondence to Lori L. Holt, Department of Psychology, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213; e-mail: lholt@andrew.cmu.edu.

P S C I	1 5 3 2	B	Dispatch: 1.2.05 Journal: PSCI	CE: Blackwell
Journal Name	Manuscript No.		Author Received: No. of pages: 8	Saravan/Anand Kumar

Coady et al., 2003; Holt & Lotto, 2002; Holt et al., 2000; Lotto & Kluender, 1998), these precursors have been single acoustic signals temporally adjacent to the speech targets or separated from them only by brief silent intervals. The temporal adjacency of the sounds allows for the possibility that cochlear interactions such as masking (Moore, 2003) account for the pattern of results. Thus, one interpretation of context effects across classes of sound is that higher-order linguistic or articulation-based processing underlies speech context effects, whereas cochlear-level auditory interactions govern the influence of nonlinguistic contexts on speech categorization (e.g., Fowler, Brown, & Mann, 2000).

The present studies introduce a new experimental paradigm to investigate whether the influence of nonlinguistic sounds on speech categorization is limited to interactions between temporally local sounds in auditory processing. The paradigm further tests whether the auditory system exhibits sensitivity to the statistical structure of spectral distributions of energy and, moreover, whether such sensitivity affects listeners' categorization of following speech. The context effects reported here indicate that linguistic and nonlinguistic acoustic information interact at higher levels of auditory processing than has been demonstrated previously.

EXPERIMENT 1

Method

For each stimulus, an "acoustic history" composed of 21 sine-wave tones plus a final standard tone preceded a speech syllable drawn from a series varying perceptually from /ga/ to /da/ (see Fig. 1). Each sine-wave tone had a duration of 70 ms and was separated from the following tone by a 30-ms silent interval.

Natural tokens of /ga/ and /da/ spoken in isolation were digitally recorded from an adult monolingual male English speaker (Computer Speech Laboratory, Kay Elemetrics, Lincoln Park, NJ; 20-kHz sample rate, 16-bit resolution). Onset frequencies of the second (F2) and third (F3) formants were increased in approximately equal steps to create a series of nine stimuli that spanned these endpoint stimuli and varied perceptually from /ga/ to /da/ (Analysis-Synthesis Laboratory, Kay Elemetrics, Lincoln Park, NJ). These speech series members served as targets for speech categorization in each of the experiments reported here.

The composition of the acoustic histories preceding the speech stimuli differentiated experimental conditions. Each acoustic history was composed of 21 sine-wave tones with unique frequencies. Experiment 1 contrasted acoustic histories sampling four distinct spectral distributions (see Fig. 1 for an example stimulus from each condition). Mean frequencies for these distributions were chosen on the basis of the findings of Lotto and Kluender (1998), who demonstrated that single 1824- versus 2720-Hz tones produce a spectrally contrastive context effect on phonetic categorization of a following syllable; these frequencies correspond roughly to the F3 offset frequencies of /al/ and /ar/ precursors that produce a similar effect of context on speech categorization (Mann, 1980). In Experiment 1, *low-mean* acoustic histories were composed of 1300- to 2300-Hz tones (M = 1800 Hz, 50-Hz steps). *High-mean* acoustic histories possessed tones sampling 2300 through 3300 Hz (M = 2800 Hz, 50-Hz steps). Two intermediate conditions shared a

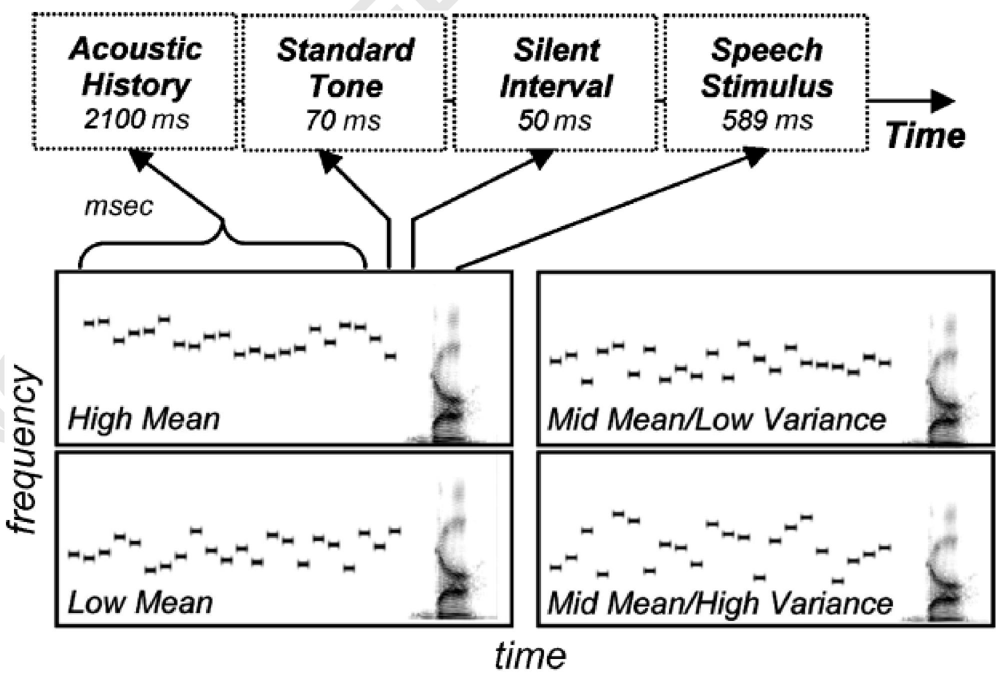

Fig. 1. Representative spectrograms (time×frequency) illustrating stimuli from each of the four conditions of Experiment 1. The schematic illustration at the top shows the durations of the components of the stimuli.

distribution mean. *Mid-mean/low-variance* acoustic histories possessed 1800- to 2800-Hz tones (M = 2300 Hz, 50-Hz steps), spanning a 1000-Hz frequency range, just like the low- and high-mean acoustic histories. *Mid-mean/high-variance* acoustic histories were composed of 1300- to 3300-Hz tones (M = 2300 Hz, 100-Hz steps), spanning the 2000-Hz frequency range sampled by the high- and low-mean acoustic histories jointly.

So that effects elicited by any particular tone ordering would be minimized, acoustic histories were created by randomizing the order of the 21 tones on a trial-by-trial basis. As a result, each trial was unique. Each trial within a condition was distinctive in surface acoustic characteristics, but was statistically consistent with other stimuli drawn from the distribution defining the condition. Observation of an influence of acoustic histories on speech categorization therefore required that listeners be influenced by the long-term spectral distribution of the acoustic histories and not merely by the simple acoustic characteristics of any particular segments of the acoustic histories.

One of the aims of this series of studies was to investigate whether nonlinguistic context must be temporally adjacent to speech targets to influence speech categorization. Therefore, acoustic histories and speech targets were separated by a neutral acoustic stimulus. In Experiment 1, this stimulus was a 70-ms, 2300-Hz standard tone that terminated each acoustic history so that the sound temporally adjacent to the speech targets was constant for all stimuli, across all conditions. This standard tone frequency was chosen as a neutral adjacent context because, given the findings of Lotto and Kluender (1998), it was expected to have little influence on categorization of the speech stimuli. In a pilot study testing this expectation, 10 listeners categorized the speech series for Experiment 1 both in isolation and preceded by a 2300-Hz standard tone plus a 50-ms silent interval. Speech categorization in the two conditions did not differ significantly, $t(9) = 1.35$, $p = .21$, $\eta_p^2 = .169$. Thus, the standard tone serves as an effective neutral adjacent acoustic context. Note, also, that the two mid-mean acoustic histories serve as control contexts because they sample spectral distributions with this neutral mean. The pilot experiment also demonstrated that when presented in isolation, the speech stimuli along the nine-step series were categorized approximately equally often as "ga" and "da" by listeners. Listeners' mean percentage of "ga" responses across the series was statistically indistinguishable from 50%, $t(9) = 0.891$, $p = .433$, indicating that listeners' category boundary between /ga/ and /da/ was centered along the stimulus series. Probit analyses (Finney, 1971) confirmed that listeners' category boundaries were statistically equivalent to the center of the speech series, $t(9) = 0.312$, $p = .762$.

Acoustic history and standard tones were synthesized with 16-bit resolution and sampled at 10 kHz using MATLAB (Mathworks, Inc., Natick, MA). Linear onset and offset ramps of 5 ms were applied to all tones. Speech stimuli were digitally down-sampled to 10 kHz, and both tones and speech tokens were digitally root mean square (RMS) matched to the RMS energy of the /da/ endpoint of the speech series. Next, the standard tone and 50-ms silent interval were appended to each of the 80 unique acoustic histories (20 acoustic histories in each of the four conditions). Each of the resulting precursors was paired with each speech token along the /ga/-to-/da/ series, for a total of 720 unique stimuli.

Seated in individual sound-attenuated booths, 10 adult monolingual English-speaking volunteers who reported normal hearing categorized the speech syllable in each stimulus by pressing the electronic button labeled "ga" or "da." Each listener responded to stimuli from each condition; within the experimental session, the order of stimulus presentation was mixed across conditions. Acoustic presentation was under the control of Tucker Davis Technologies (Alachua, FL) System II hardware; stimuli were converted from digital to analog, low-pass filtered at 4.8 kHz, amplified, and presented diotically over linear headphones (Beyer DT-150, Berlin, Germany) at approximately 70 dB SPL(A). Results were analyzed in terms of average percentage of "ga" responses.

Results and Discussion

As shown in Figure 2, speech categorization was influenced by the acoustic-history distributions differentiating conditions, $F(3, 9) = 29.3$, $p < .001$, $\eta_p^2 = .765$. Planned Bonferroni-

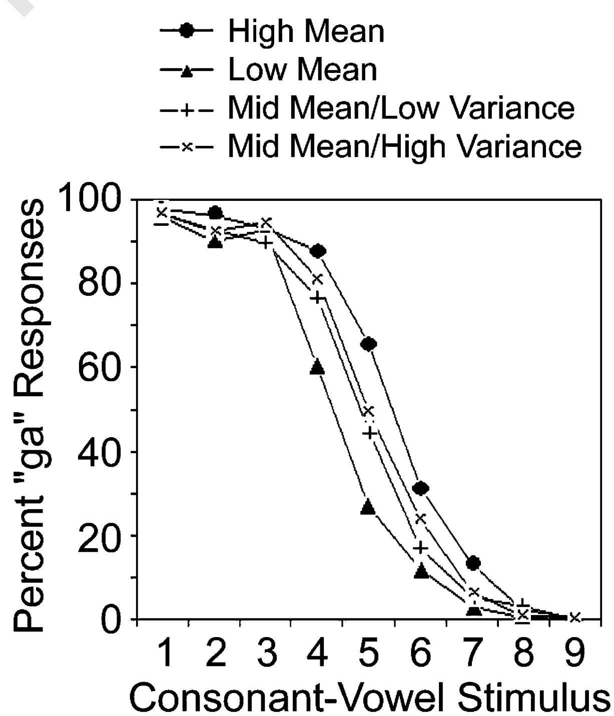

Fig. 2. Results of Experiment 1: mean percentage of ''ga'' responses to each stimulus in the speech series as a function of the spectral distribution characteristics of the preceding acoustic history.

corrected comparisons indicated a marginal difference in speech categorization for the two mid-mean conditions (M = 2300 Hz, differentiated by the range of variability of the tones composing the acoustic histories), $p = .063$, $\eta_p^2 = .333$; listeners were somewhat more likely to identify syllables as "ga" when the acoustic history was composed of tones sampling a larger range. All other planned post hoc cross-condition comparisons were significant, $p < .001$, $\eta_p^2 > .644$. There was also a main effect of stimulus, $F(3, 9) = 317.97$, $p < .001$, $\eta_p^2 = .972$, indicating that listeners differentially categorized the speech tokens across the series, and a significant stimulus-by-acoustic-history interaction, $F(3, 9) = 9.42$, $p < .001$, $\eta_p^2 = .511$, indicating differences in the shape of the categorization curves across conditions.

These patterns of response demonstrate that context that is neither linguistic nor local can have a marked influence on listeners' speech categorization. It is worth noting that the direction of the influence of the precursor acoustic histories was spectrally contrastive. When acoustic histories composed of tones sampling the high-frequency distribution preceded the speech targets, listeners categorized the syllables as "ga" (the alternative with greater low-frequency energy) more often than "da" (the alternative with greater high-frequency energy). Conversely, acoustic histories composed of lower-frequency tones induced listeners to categorize the same speech sounds as "da" more often than "ga."

These results indicate that auditory processing is sensitive to the distribution of spectral information in the sound environment. The composition of the acoustic histories varied on a trial-by-trial basis, so a single tone or an idiosyncratic ordering of a small group of tones within the acoustic histories cannot be responsible for the effect. As a result, this context effect required some degree of memory as the acoustic histories unfolded in time and appears to indicate a degree of abstraction from the precise acoustic signal presented on any given trial. Recall also that the standard tone terminating each acoustic history was constant across conditions and was shown to be a perceptually neutral context in a pilot test. Thus, acoustic histories were always temporally nonadjacent to the speech targets. The effect of acoustic histories on speech categorization under these circumstances makes cochlear masking or other low-level auditory perceptual processes unlikely candidate mechanisms, therefore suggesting that nonlinguistic acoustic signals may interact with speech categorization at higher levels of auditory processing.

EXPERIMENT 2

A means of investigating this possibility further is to examine the time course across which acoustic histories exert an influence on speech categorization. Previous research has demonstrated that the influence of temporally adjacent acoustic contexts (whether speech or nonspeech) on speech categorization diminishes when context and speech target are separated by silent intervals greater than approximately 200 ms (Holt & Lotto, 2002; Lotto et al., 2003). Experiment 2 examined the time course of the effect observed in Experiment 1 by manipulating the silent interval between the final (standard) tone and the speech target.

Method

The context effect observed in Experiment 1 was greatest for the third through seventh speech stimuli in the series, which were perceptually ambiguous. The range used in Experiment 2 was therefore reduced to decrease the number of observations necessary at each silent interval. Stimuli were created as in the high- and low-mean conditions of Experiment 1 except that duration of the silent interval was varied: 100, 150, 200, 250, and 300 ms in Experiment 2a and 500, 700, 900, 1,100, and 1,300 ms in Experiment 2b.

The procedure and apparatus were the same as in Experiment 1. Two groups of 10 adult monolingual English speakers who reported normal hearing served as listeners. Each listener responded to stimuli from each condition; within an experiment, the order of stimulus presentation was mixed across conditions. In all, listeners identified the speech syllables of 500 unique stimuli (10 repetitions of each of the combinations of five speech stimuli, five silent intervals, and two acoustic-history means).

Results and Discussion

The results are shown in Figure 3. For the silent intervals investigated in Experiment 2a (top panels), there was a significant spectrally contrastive effect of the acoustic-history distributions on speech categorization, $F(1, 9) = 160.84$, $p < .001$, $\eta_p^2 = .947$, with listeners labeling speech tokens as "ga" more often when acoustic histories were sampled from the high-mean distribution than the low-mean distribution. Bonferroni-corrected comparisons revealed significant context effects at all silent intervals, $p < .05$, $\eta_p^2 > .702$. There was no main effect of duration of the silent interval, $F < 1$, $\eta_p^2 = .056$. However, the interaction between silent interval and acoustic-history mean was significant, $F(4, 9) = 6.21$, $p < .001$, $\eta_p^2 = .408$. The source of this interaction is surprising: The size of the context effect increased with longer silent-interval durations. There was a significant main effect of stimulus, $F(4, 9) = 140.66$, $p < .001$, $\eta_p^2 = .94$, and a stimulus-by-mean-acoustic-history interaction, $F(4, 9) = 7.23$, $p < .001$, $\eta_p^2 = .445$, but no interaction of silent interval and stimulus, $F(16, 9) = 1.38$, $p = .12$, $\eta_p^2 = .141$. The three-way interaction was significant, $F(16, 9) = 2.71$, $p < .001$, $\eta_p^2 = .231$.

Experiment 2b (see bottom panels of Fig. 3) tested the context sensitivity of speech categorization among a different group of listeners to examine whether the context effect diminishes when much longer silent intervals separate the context and speech targets. Participants again exhibited a significant spectrally

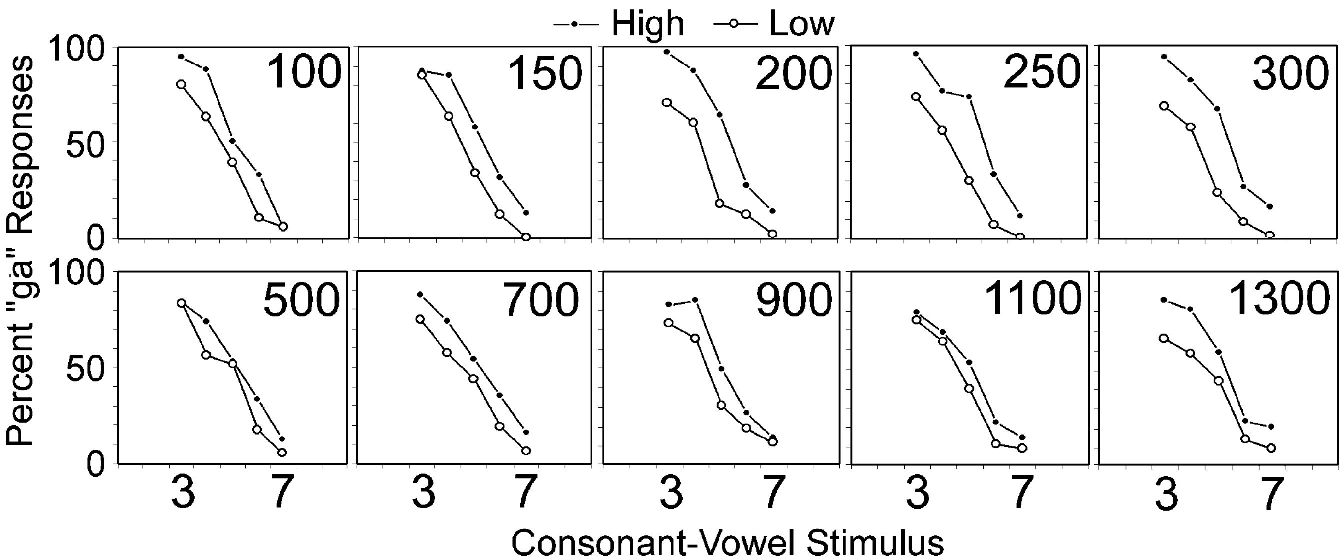

Fig. 3. Results of Experiments 2a (top row) and 2b (bottom row): mean percentage of "ga" responses to each stimulus in the speech series as a function of the spectral distribution characteristics of the preceding acoustic history (high vs. low). Each graph shows the results for a different silent interval separating context and speech target (100 ms–1,300 ms).

contrastive context effect across the acoustic-history means, $F(1, 9) = 46.165$, $p < .001$, $\eta_p^2 = .837$. Bonferroni-corrected comparisons revealed a context effect at all silent intervals, $p < .05$, $\eta_p^2 > .509$. There was a significant effect of stimulus, $F(4, 9) = 105.51$, $p < .001$, $\eta_p^2 = .921$, but no main effect of silent interval, $F < 1$, $\eta_p^2 = .057$. Only the interaction between silent interval and stimulus was significant, $F(16, 9) = 2.07$, $p = .01$, $\eta_p^2 = .187$. The lack of a significant interaction between acoustic-history mean and silent-interval duration, $F(4, 9) = 1.84$, $p = .14$, $\eta_p^2 = .169$, indicates that (unlike in Experiment 2a) the magnitude of the context effect did not vary significantly as a function of the silent interval's duration.

Thus, even when 1.3 s of silence separated the acoustic history and the following speech syllable, there was a significant influence of context on speech categorization. This effect of nonlinguistic stimuli on speech categorization is intriguing, especially given the long durations of silence over which these effects persisted. Although these data do not provide an upper bound for the time course of these context effects, they indicate that the time course is significantly longer than has been previously reported for effects of temporally adjacent precursors on speech categorization. Lotto et al. (2003) reported that adjacent nonlinguistic context effects on speech categorization diminish within several hundred milliseconds. It is of note that the effects of adjacent speech precursors on speech identification also dissipate over several hundred milliseconds (Holt & Lotto, 2002; Lotto et al., 2003). The effect reported here thus appears to be distinct; the influence of temporally nonadjacent acoustic histories on speech categorization is much more robust across time than previously reported effects of temporally adjacent speech or nonlinguistic acoustic contexts on speech categorization.

EXPERIMENT 3

In the previous experiments, the standard tone served as a constant neutral sound context separating the acoustic history and the speech syllable. It might be argued that the presence of a single 70-ms standard tone was not sufficient to accomplish the experimental aim of temporal nonadjacency between context and target. Experiment 3 examined the influence of repetition of the standard tone in order to further probe the degree to which nonadjacent nonlinguistic context may affect speech categorization.

Method

The stimuli in this experiment were created in the same way as the high- and low-mean stimuli in Experiment 2 except that the number of repetitions of the standard tone preceding the silent interval varied and the silent interval was 50 ms long. Stimuli with 1, 3, 5, and 7 repetitions of the standard tone between the acoustic history and the speech target were created for Experiment 3a. Experiment 3b stimuli had 1, 9, 11, or 13 repetitions of the standard tone. As in the other experiments, standard tones had a duration of 70 ms; they were separated by 30-ms silent intervals. Thus, the overall duration of the stimuli varied across conditions.

The procedure and apparatus were the same as in Experiment 1. Two groups of 10 adult monolingual English speakers who reported normal hearing served as listeners. Each listener responded to stimuli from each condition; within an experiment, the order of stimulus presentation was mixed across conditions. Listeners identified the speech syllable of 400 stimuli (10 repetitions of each combination of five speech stimuli, four standard-tone-repetition conditions, and two acoustic-history means).

Results and Discussion

Figure 4 shows the results as mean percentage of "ga" responses by acoustic-history condition and number of standard-tone repetitions. For Experiment 3a (top panels), there was a significant influence of acoustic-history distribution mean on speech categorization, $F(1, 9) = 130.91$, $p < .0001$, $\eta_p^2 = .936$ (all Bonferroni-corrected comparisons, $p < .002$, $\eta_p^2 > .759$). Preceding acoustic histories had a spectrally contrastive influence on speech categorization even across seven repetitions (700 ms) of an intervening standard tone. There were also significant main effects of repetition, $F(4, 9) = 4.41$, $p = .01$, $\eta_p^2 = .329$, and stimulus, $F(4, 9) = 223.30$, $p < .0001$, $\eta_p^2 = .961$. There was no interaction between repetition and acoustic-history mean, $F < 1$, $\eta_p^2 = .095$, indicating that the context effect was statistically equivalent across the standard-tone-repetition conditions. The interaction of acoustic-history distribution and stimulus, $F(4, 9) = 23.21$, $p < .0001$, $\eta_p^2 = .721$, and three-way interaction, $F(12, 9) = 2.58$, $p = .005$, $\eta_p^2 = .223$, were significant, whereas the repetition-by-stimulus interaction was not, $F(12, 9) = 1.04$, $p < .42$, $\eta_p^2 = .095$.

Experiment 3b (see the bottom panels of Fig. 4) tested the context sensitivity of speech categorization among a separate group of listeners who heard up to 13 standard-tone repetitions. As in the previous experiments, there was a main effect of acoustic-history mean, $F(1, 9) = 53.54$, $p < .0001$, $\eta_p^2 = .856$. Bonferroni-corrected comparisons revealed a spectrally contrastive context effect for all levels of standard-tone repetition, $p < .002$, $\eta_p^2 > .656$. There was no interaction of repetition and acoustic-history mean, $F(3, 9) = 2.14$, $p = .12$, $\eta_p^2 = .193$, indicating that the influence of acoustic histories on speech categorization did not vary across different numbers of repetitions of the standard tone. Whether the acoustic histories were followed by a single 70-ms standard tone or by 13 repetitions of the same tone across 1.3 s, the spectral distribution of the tones composing the preceding acoustic history influenced categorization of speech in a contrastive manner. There were also main effects of standard-tone repetition, $F(3, 9) = 4.46$, $p = .01$, $\eta_p^2 = .332$, and stimulus, $F(4, 9) = 123.76$, $p < .0001$, $\eta_p^2 = .932$, and significant interactions of stimulus with number of repetitions, $F(12, 9) = 1.85$, $p = .05$, $\eta_p^2 = .171$, and stimulus with acoustic-history mean, $F(4, 9) = 11.12$, $p < .0001$, $\eta_p^2 = .553$.

In sum, the influence of the acoustic histories on speech categorization persisted even when 13 repetitions of the intervening standard tone terminated each acoustic history. The persistence of the context effect across intervening acoustic stimuli further distinguishes the influence of these acoustic histories from previously reported effects of temporally adjacent nonlinguistic context on speech categorization.

GENERAL DISCUSSION

These results demonstrate that context need be neither linguistic nor local to influence speech categorization. Quite sur-

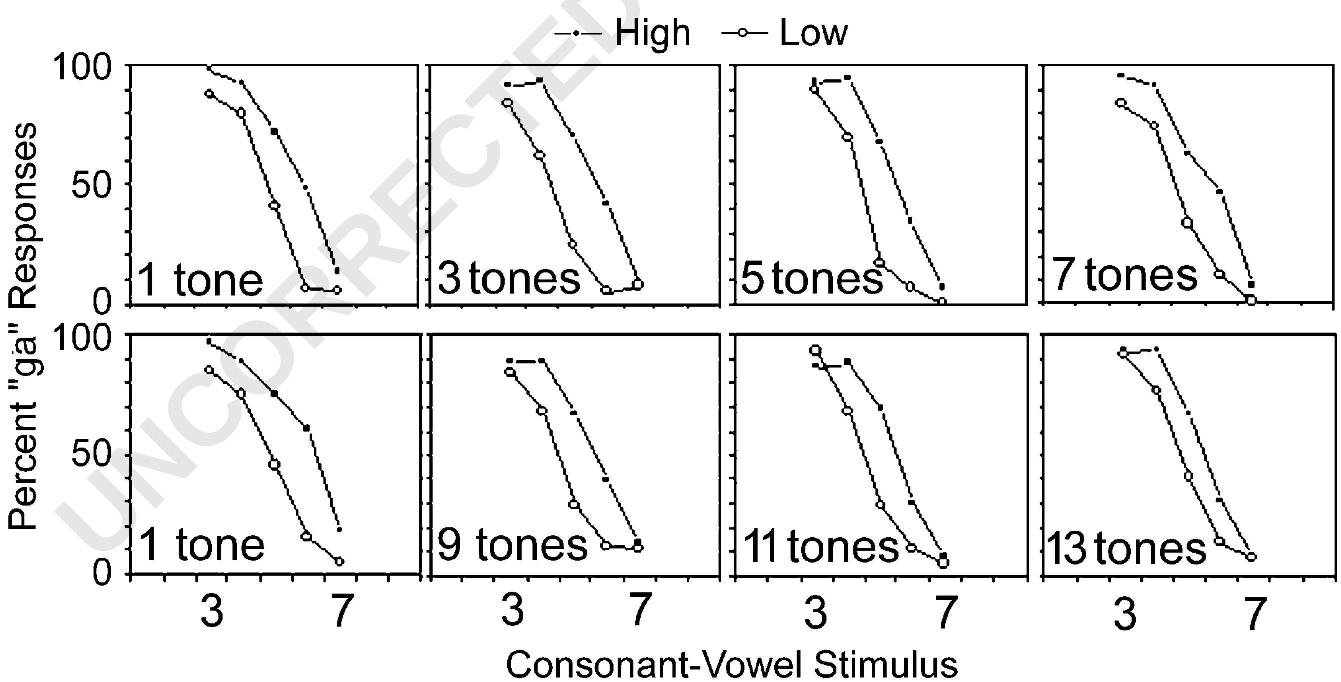

Fig. 4. Results of Experiments 3a (top row) and 3b (bottom row): mean percentage of ''ga'' responses to each stimulus in the speech series as a function of the spectral distribution characteristics of the preceding acoustic history (high vs. low). Each graph shows the results for a particular number of standard-tone repetitions separating acoustic histories and speech targets.

prisingly, hearing a randomly ordered sequence of sine-wave tones sampling a region of the acoustic spectrum greatly influences listeners' categorization of a following speech syllable. The direction of influence of such acoustic histories is spectrally contrastive, a finding that is in agreement with previous results demonstrating that single adjacent nonlinguistic sounds shift speech categorization contrastively (e.g., Holt et al., 2000; Lotto & Kluender, 1998).

However, whereas the effect of temporally adjacent context diminishes when hundreds of milliseconds separate context and speech target, the present results show a considerably more resilient effect, persisting across at least 1.3 s of silence or intervening acoustic stimuli. There remains the possibility that the influence of acoustic histories and that of temporally adjacent stimuli lie at ends of a continuous range of contextual influence on speech categorization. However, from what is presently known, the effect of acoustic histories on speech categorization appears, from its time course, to be distinct from the effect of temporally adjacent contexts. Thus, these results suggest that perceptual contrast may exist at multiple scales in auditory processing.

Some accounts have differentiated context effects elicited by nonlinguistic and speech precursors, with nonlinguistic context effects presumed to originate from processes, such as cochlear masking, that are distinct from the processes that produce phonetic context effects (e.g., Fowler, Brown, & Mann, 2000). This interpretation was challenged recently by the observation that temporally adjacent nonlinguistic contexts influence speech categorization even when context and speech target are presented to opposite ears, and thus do not have an opportunity to interact within the cochlea (Holt & Lotto, 2002; Lotto et al., 2003). The present results further blur the distinction that has been made in explaining the relative influence of speech versus nonlinguistic acoustic contexts on speech categorization. The protracted time course of the effect of acoustic histories on speech categorization and the persistence of the effect across intervening acoustic stimuli render a purely cochlear interpretation untenable. Moreover, the fact that the effect was obtained when acoustic histories were defined statistically, so that the specific acoustic characteristics of the stimuli varied on a trial-by-trial basis, indicates that the effect involves a degree of abstraction from the precise acoustic characteristics of the stimulus that is inconsistent with a solely cochlear account.

The temporal window across which the effect was found makes tempting an analogy to context-sensitive speech processing at the sentence level. Considerable shifts in speech categorization occur when the spectral or temporal characteristics of a preceding sentence are manipulated (Ladefoged & Broadbent, 1957; Summerfield, 1981). These findings have been related to perceptual accommodation to the acoustic variability arising from differences among talkers and their rates of speech. Moreover, they have been classically described as examples of normalization, implying (in many models) a speech-specific translation from the surface acoustic variability to a more canonical representation. The context-sensitive shifts elicited by changes in the spectra or rate of a precursor sentence are contrastive. Thus, the nonlinguistic nature of the present context effects, coupled with their sentence-level time course, suggests the possibility that context effects that have been attributed to speech-specific normalization may arise when spectral or temporal regularities in the acoustic environment produce contrastive effects on speech categorization. Further tests will be needed to determine the extent to which effects like those elicited by the acoustic histories of the present experiments mirror the effects described as instances of normalization in the speech perception literature. It is possible that the sensitivity to spectral distributions observed here may contribute to maintaining perceptual constancy across the variability introduced by, among other things, nonnative accented speech and variability in talkers' characteristics. The experimental paradigm of the present studies provides a means by which to manipulate bottom-up acoustic information to investigate this possibility.

Most generally, the present results indicate that speech categorization is not independent of the extended temporal context of sounds in the acoustic environment. Nonlinguistic acoustic stimuli can influence speech categorization even when these stimuli are nonadjacent, statistically defined contexts that preclude solely cochlear mechanisms. These context effects indicate higher-level interaction across sound classes than has been observed previously and are consistent with the hypothesis that speech perception derives from general characteristics of auditory perception and cognition.

Acknowledgments—The author thanks A.J. Lotto, R.L. Diehl, and L.A. Gonnerman for helpful critiques of the manuscript and Christi Adams for her essential role in conducting the research. This work was supported by grants from the James S. McDonnell Foundation and the National Institutes of Health.

REFERENCES

Coady, J.A., Kluender, K.R., & Rhode, W.S. (2003). Effects of contrast between onsets of speech and other complex spectra. *Journal of the Acoustical Society of America*, *114*, 2225–2235.

Finney, D.J. (1971). *Probit analysis*. Cambridge, England: Cambridge University Press.

Fowler, C.A., Brown, J.M., & Mann, V.A. (2000). Contrast effects do not underlie effects of preceding liquids on stop-consonant identification by humans. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 877–888.

Holt, L.L., & Lotto, A. (2002). Behavioral examinations of the level of auditory processing of speech context effects. *Hearing Research*, *167*, 156–169.

Holt, L.L., Lotto, A.J., & Kluender, K.R. (2000). Neighboring spectral content influences vowel identification. *Journal of the Acoustical Society of America*, *108*, 710–722.

Ladefoged, P., & Broadbent, D.E. (1957). Information conveyed by vowels. *Journal of the Acoustical Society of America*, *29*, 98–104.

Lindblom, B.E.F., & Studdert-Kennedy, M. (1967). On the role of formant transitions in vowel recognition. *Journal of the Acoustical Society of America*, *42*, 830–843.

Lotto, A.J., & Kluender, K.R. (1998). General contrast effects of speech perception: Effect of preceding liquid on stop consonant identification. *Perception & Psychophysics*, *60*, 602–619.

Lotto, A.J., Kluender, K.R., & Holt, L.L. (1997). Perceptual compensation for coarticulation by Japanese quail (*Coturnix coturnix japonica*). *Journal of the Acoustical Society of America*, *102*, 1134–1140.

Lotto, A.J., Sullivan, S., & Holt, L.L. (2003). Central locus for nonspeech context effects on phonetic identification. *Journal of the Acoustical Society of America*, *113*, 53–56.

Mann, V.A. (1980). Influence of preceding liquid on stop-consonant perception. *Perception & Psychophysics*, *28*, 407–412.

Mann, V.A., & Repp, B.H. (1980). Influence of vocalic context on perception of the [sh]-[s] distinction. *Perception & Psychophysics*, *28*, 213–228.

Mann, V.A., & Repp, B.H. (1981). Influence of preceding fricative on stop consonant perception. *Journal of the Acoustical Society of America*, *69*, 548–558.

Moore, B.C.J. (2003). *An introduction to the psychology of hearing*. San Diego, CA: Academic Press.

Stephens, J.D.W., & Holt, L.L. (2003). Preceding phonetic context affects perception of non-speech sounds. *Journal of the Acoustical Society of America*, *114*, 3036–3039.

Summerfield, Q. (1981). On articulatory rate and perceptual constancy in phonetic perception. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 1074–1095.

(Received 3/10/04; Accepted 4/19/04)